Development of molecular tools in the diatom *Phaeodactylum tricornutum*

Dissertation

Zur Erlangung des akademischen Grades des
Doktors der Naturwissenschaften
- Dr. rer. Nat. -

An der Universität Konstanz
Fachbereich Biologie

vorgelegt von Arne Christian Materna

Tag der mündlichen Prüfung: 21. Dezember 2006

Referent/in: Prof. Dr. Peter G. Kroth

Referent/in: Prof. Dr. Iwona Adamska

Widmen möchte ich diese Arbeit meinem Vater,

von ihm habe ich gelernt, dass man mit harter Arbeit beinahe alles im Leben erreicht,
die Arbeit jedoch nur ein Teil des Lebens sein darf.

Danksagung

Mein ganz besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr. Peter Kroth, nicht nur für die Möglichkeit mich im Rahmen meiner Doktorarbeit an einem EU-Projekt beteiligen zu können, sondern auch für die hervorragende Betreuung und einen freundschaftlichen Umgang. Besonders genossen habe ich unsere „Ausflüge" zu zahlreichen Tagungen und Konferenzen, die mir viele Eindrücke in die Welt der Wissenschaft ermöglicht haben und mir stets in angenehmer Erinnerung bleiben werden.

Vielen Dank auch an Prof. Dr. Iwona Adamska für die Übernahme des Zweitgutachtens, und für die immer offene Tür zu ihrem Büro. Besonders dankbar bin ich ihr für ihre Bestrebungen uns stets die Balance zwischen Arbeit und dem nötigen Ausgleich zu ermöglichen (Obertauern rules!).

Besonders dankbar bin ich ebenfalls Prof. Dr. Wolfgang Hess für sein stetiges Interesse an meiner Arbeit, die vielen zahlreichen Anregungen und für seinen freundschaftlichen Brückenschlag über den Atlantik nach Boston.

Vielen Dank auch an Doris Ballert und Angelika Eckert, die mir im Labor immer zur Seite standen. Ohne den unermüdlichen und kompetenten Einsatz von Doris Ballert – und besonders ihrer moralische Unterstützung – erscheint mir die Arbeit der letzten vier Jahre schier undenkbar. Daher danke ich ihr hier ganz besonders.

Ein weiterer Mensch, der mich als Doktorand von Anfang an begleitete ist Sabine Ng Chin Yue, meine erste Vertiefungskursstudentin und späterer Hiwi, die mir auch während ihrer Diplomarbeit treu blieb und heute selbst als Doktorandin Prof. Dr. Kroths Lehrstuhl bereichert. Durch ihren wertvollen Einsatz an meiner Seite und ihre freundschaftliche Unterstützung in allen Lebenslagen gebührt ihr nicht nur mein aufrichtiger Dank, sondern sie hat sich auch dauerhaft Anteil an dieser Arbeit verdient. So ist sie Co-Autorin des Manuskripts im Kapitel II.3 dieser Dissertation.
Großer Dank gebührt auch Ansgar Gruber für spannende und bereichernde Unterhaltungen und die vielen Stunden am Okular diverser Mikroskope. Unsere abendlichen Fachsimpeleien hatten für mich weit mehr als nur Unterhaltungswert.

Ich danke allen Kollegen und Mitarbeiter auf der Etage 9 für das freundliche Arbeitsklima.

Ich möchte mich bei meiner Familie und ganz besonders bei meiner Mutter dafür bedanken, dass Sie mir das Studium der Biologie ermöglicht und mich stets auf meinem Weg unterstützt haben. Dankbar erwähnen möchte ich hier meine Großmutter, die mich auf ihre Art von frühester Kindheit an zum Forschen ermutigte.

Mein Dank gebührt auch meinen Freunden – saben quién estoy hablando. Muchos gracias Amigos, para todos!!!

Mehr als mein Dank gehört meiner Céline, nicht nur für die Logistische Unterstützung während der letzten Monate sondern ganz besonders dafür, mir treue Begleiterin auf allen Wegen zu sein und zugleich Ruhepol und Rückzugsort.

I Introduction

Zusammenfassung

Kieselalgen (Diatomeen), einzellige, eukaryotische Algen besiedeln die Ozeane und Süßwasserhabitate. Sie gehören zusammen mit Cyanobakterien zu den am häufigsten auftretenden phytoplanktonischen Lebensformen auf der Erde und sind daher von großer ökologischer Bedeutung. Ein besonderes Merkmal der goldbraun gefärbten Algen ist ihre Eigenschaft ornamentierte Schalen (frustules) aus Kieselsäure zu bilden. In der Vergangenheit waren Kieselalgen wegen ihrer faszinierenden Physiologie als auch ihrer ökologischen Relevanz Gegenstand zahlreicher Studien. In letzter Zeit wurde ebenfalls ihr Potential für biotechnologische Anwendungen erkannt. Da intensivere Studien an Diatomeen in jüngster Zeit häufig durch den Mangel an Genomdaten sowie einer limitierten Auswahl an geeigneten molekularbiologischen Methoden und „molekularen Werkzeugen" beeinträchtigt worden waren, erhöhte sich die Nachfrage nach einem Modellorganismus für Kieselalgen sowie nach geeigneten molekularbiologischen Methoden.

Diese Dissertation behandelt *Phaeodactylum tricornutum*, eine pennate Kieselalge, die einfach unter Laborbedingungen kultiviert und erforscht werden kann, weshalb sie bereits als möglicher Modellorganismus für Diatomeen diskutiert worden war. Daher wurde auch in diesen Tagen die Sequenzierung ihres Genoms erfolgreich abgeschlossen, die Annotierung des Genoms geht ebenfalls ihrer Vollendung entgegen. (http://shake.jgi-psf.org/Phatr2/Phatr2.home.html). Um *Phaeodactylum tricornutum* im weiteren zugänglich für postgenomische Anwendungen zu machen, sowie um die Rolle der Alge als Modellorganismus für Diatomeen weiter zu festigen, war es das Ziel dieser Dissertation, verschiedene, bislang nicht zur Verfügung stehenden, molekularbiologische Anwendungen zur Erforschung von Kieselalgen zu entwickeln.

Kapitel II.1 beschreibt drei verschiedene Strategien zur Entwicklung eines stabilen Plastidentransformationssystems für *Phaeodactylum tricornutum*. Die Strategien basieren auf der Insertion des Streptomycinresistenzgens *aadA* entweder als zusätzliches Operon-Gen in das plastidär codierte RUBISCO-Operon, oder in nichtcodierende intergenische Bereiche des Chloroplastengenoms. Eine weitere Strategie basiert auf der Substitution des funktionellen *psbA* Gens durch Versionen die verschiedene Punktmutationen tragen und so zur Herbizidresistenz führen soll. Die Resultate der Transformationsexperimente legen nahe, dass Streptomycinresistenz transient in der Plastide der Kieselalge induziert werden kann, jedoch eine stabile Expression des Markers war nicht möglich. Unklar

bleibt, ob Marker nur temporär in das plastidäre Genom integrieren oder transient von episomalen Plasmiden exprimiert werden.

Kapitel II.2 beschreibt erstmalig einen induzierbaren und wahrscheinlich gerichteten mutagenen Mechanismus der Mutationen im Plastidengenom von *Phaeodactylum tricornutum* erzeugt. Mutagenese konnte in den plastidär codierten Genen für D1 (*psbA*) und für die 16S rRNA gezielt induziert werden. Das Phänomen einer induzierbaren Mutagenese wurde bereits intensiv in *E. coli* studiert und konnte darüber hinaus auch in Eukaryoten beobachtet werde. Die hier vorliegende Studie beschreibt jedoch erstmalig induzierbare Mutagenes in einem Organellengenom. Die Aufklärung des Auslösers für die gerichtete Mutagenese in einem Zielgen könnte neben ihren genetischen Implikationen auch zur gentechnischen Manipulation von Plastidengenomen beitragen. Die Konsequenzen der aktiven Induzierung von Mutagenese im Genom eines Zellorganells für dessen Evolutionsraten und damit einhergehend für unser Verständnis von Plastidenevolution sind beachtlich.

Kapitel II.3 enthält ein Protokoll das Gen-Silencing in *Phaeodactylum tricornutum erlaubt*, eine Anwendung, die bislang nicht zur Erforschung von Kieselalgen zur Verfügung stand. Als Zielgen wurde das Gen für die Diadinoxanthin Deepoxidase (*dde*) gewählt, die essentiell für den Lichtschutzmechanismus NPQ ist. Zwei verschiedene Strategien erzeugten gleichermaßen Transformanden, die einen deutlich NPQ-reprimierten Phänotyp aufwiesen. Die Mehrheit der untersuchten Transformanden zeigte eine Reduktion des NPQ um 30-47% im Vergleich zum Wildtyp. Durch Untersuchungen des *dde*-Transkriptlevels mittels RT-qPCR konnten Unterschiede zwischen Transformanden und Wildtyp, als auch zwischen den aus unterschiedlichen Silencing-Strategien hervorgegangenen Transformanden nachgewiesen werden. Diese Unterschiede legten nicht nur die Anwendbarkeit von Gen-Silencing in Kieselalgen nahe, sondern darüber hinaus das Vorhandensein zweier verschiedener Silencing-Mechanismen in *Phaeodactylum tricornutum*.

Im Gegensatz zu Sequenzanalysen, die sich auf den genetischen Inhalt konzentrieren, der in der Nukleotidabfolge der DNA codiert liegt, beschreibt Kapitel II.4 Techniken, die die Betrachtung und Erforschung der verschiedenen in Organellen und Zellkern liegenden Genome ermöglicht. Kapitel II.4 enthält ein Protokoll das ersmalig die selektive *in vivo* Visualisierung von beiden organellären Genomen im selben Organismus erlaubt. Im Weiteren wird eine Methode präsentiert, die mittels RT-qPCR die exakte Bestimmung der Ploidien in Chloroplast und Mitochondrialem Netzwerk ermöglicht.

Abstract

Diatoms (Bacillariophyta) are unicellular eukaryotic algae that colonize the oceans and freshwater habitats. Together with Cyanobacteria, diatoms belong to the most abundant phytoplanktonic organisms on earth and therefore are of great ecological relevance. A salient feature of these golden-brown coloured microalgae is their ability to build unique siliceous cell walls (frustules). Diatoms have been intensively studied because of their fascinating physiology and ecological relevance. Recently diatoms also entered the focus of biotechnology. Since these studies have been hampered in the past due to insufficient genome data and a limited availability of molecular methods and tools, "diatomists" emphasised more and more the obvious need of a model organism and appropriate molecular tools.

This dissertation focuses on *Phaeodactylum tricornutum*, a pennate diatom which became a convenient laboratory strain and consequentially was discussed as a model organism for diatoms. To date its genome has been fully sequenced (sequence information available on http://shake.jgi-psf.org/Phatr2/Phatr2.home.html), and the annotation is nearly completed. In order to make *Phaeodactylum tricornutum* now accessible to postgenomic applications and to further support its role as model organism, the aim of this dissertation was to develop various molecular tools not yet available for diatoms.

Chapter II.1 describes three independent approaches to establish a system for stable chloroplast transformation in *Phaeodactylum tricornutum* basing on the insertion of the streptomycin-resistance gene *aadA* into the plastid encoded RUBSICO-operon as a third operon gene or within a non-coding, intergenic region. Further vectors were designed to replace the functional *psbA* gene by a slightly altered version of this gene, thus increasing herbicide tolerance. The obtained results indicate that the resistance can be transiently expressed in the diatoms' plastids. However, permanent expression of the marker gene in the plastids did not occur. Since the targeted insertion of the marker could not be verified it is likely that the resistance gene is either transiently expressed from episomal transformation vectors or temporarily integrates into the plastid genome via heterologous recombination.

Chapter II.2 reports an inducible mechanism which generates targeted mutations in a chloroplast genome (plastome). Mutagenesis was induced in *psbA* and the 16S rRNA gene, both encoded in the plastid genome *Phaeodactylum tricornutum*. Induced mutagenesis is a phenomenon intensively studied in *E. coli* and also observed for eukaryotes. This study, however, is the first example of an inducible mutagenesis mechanism in an organellar genome. The elucidation of the trigger to induce

mutagenesis in a specific target gene might contribute to engineering the chloroplast genome. Furthermore the implications of the organelle actively increasing mutation rates, and therewith rates of genome evolution, on our understanding of plastid evolution might not be conceivable to date.

Chapter II.3 describes a protocol for gene silencing in *Phaeodactylum tricornutum*. Silencing techniques were not available for diatoms so far. The diadinoxanthin de-epoxidase (*dde*), which is inevitable for the photoprotective NPQ mechanism to develop, was chosen as target gene. RNA interference was induced by transformation of the cells with plasmids which either allow the transcription of antisense fragments or of a self-complementary hairpin like construct with a 5'-sense-overhang. The silencing approaches generated transformants with a phenotype clearly distinguishable from wildtype cells. The majority of the examined transformants showed even between 30% to 47% reduction in NPQ compared to wildtype. Real-time PCR based quantification of *dde* transcripts showed differences in *dde* transcript levels between AS strains and wildtype cells but also between AS and RNAi strains, thus suggesting the presence of two different gene silencing mediating mechanisms in diatoms.

In contrast to sequence analysis which focuses on the genetic information encoded in the nucleotide composition of DNA, chapter II.4 presents two techniques focusing on the genomes themselves as structural entities, safely maintained and replicated within nuclei and organelles. A protocol was developed for the diatom *Phaeodactylum tricornutum*, which allows for the first time to visualize selectively and in vivo chloroplast or mitochondrial nucleoids in the same organism. Furthermore a high throughput capable method was designed to quantify organellar genomes thus determining ploidies with high accuracy. The method bases on a specially designed quantitative Real-time PCR protocol. Applying these techniques allowed the determination and monitoring of organellar ploidies and of nucleoid numbers per organelle. Further the subcellular localization of nucleoids inside the organelles, as well as their genome contents can be studied. The obtained results revealed new and unique insights into the system of a diatom's chloroplast.

General Introduction

Diatoms

Diatoms (Bacillariophyta) are unicellular eukaryotic algae that colonize the oceans and freshwater habitats and belong to the division of Heterokonts (Stramenopiles). They also appear in soils and on damp surfaces. Diatoms live pelagically in open water or aggregate to live in surface biofilms at the water-sediment interface (Mann 1999). Together with Cyanobacteria, diatoms belong to the most abundant phytoplanktonic organisms on earth and therefore are of great ecological relevance. Diatoms contribute to approximately up to 40% of the world's marine primary production (Falkowski et al. 1998). On a global scale diatoms are estimated to produce around 20% of the annual biomass (van den Hoek et al. 1997; Field et al. 1998) which is equivalent to the productivity of the tropical rain forests. A salient feature of these golden-brown coloured microalgae is their ability to build unique siliceous cell walls (frustules). The biogenic silica is synthesised intracellularly by the polymerisation of silicic acid monomers and subsequently extruded to the cell exterior (Zurzolo and Bowler 2001; Falciatore and Bowler 2002). Diatoms build a remarkable variety of highly patterned and often ornate structured cell walls. The patterns of these cell walls are so precise that they form the basis for diatom taxonomy and systematics (Round et al. 1990). The term diatom has its origins in the Greek language. 'Dia' means 'through' while 'tomos' stands for 'cutting'. Indeed, the frustule comprising the diatom "cuts in half" in terms of consisting of separate shells termed valves, which typically overlap one other like the two halves of a petri dish. Hence the alga is contained in a "box of glass" which consists of an outer valve

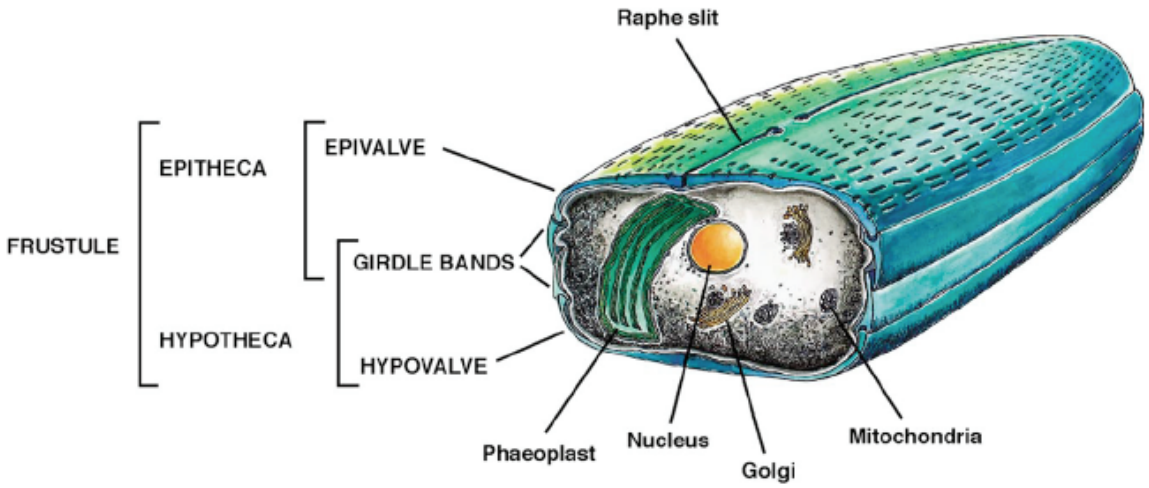


Figure 1: Schematic overview of the general structural features of a pennate diatom.
A. Falciatore & C. Bowler, 2002

(ephitheca) and a smaller inner valve (hypotheca) (Fig. 1). Traditionally, diatoms are grouped into centric and pennate diatoms. Centric diatoms have a radial symmetry of their cell wall patterns, whereas pennate diatoms are bilaterally symmetrical (Round et al. 1990; Kooistra et al. 2003). Furthermore, two forms of pennates occur. Raphid pennate diatoms have a slit (raphe) in the cell wall, required for movement; the araphid pennates lack this slit and are nonmotile. The arrangement of the photosynthetic thylakoid membrane within diatom's plastids displays a structure typical for heterokonts. The thylakoids are grouped into stacks (lamellae) of three, all enclosed by a girdle lamella (van den Hoek et al. 1997). Furthermore the photosynthetic apparatus of diatoms contains a typical set of pigments. The plastids contain the chlorophylls a and c, together with fucoxanthin, the carotenoid responsible for the brown coloration (Owens 1986). There are indications that diatoms may be capable of C4 photosynthesis (Reinfelder et al. 2000). This specialized form of photosynthesis which allows a more efficient utilization of available CO_2 is usually restricted to a few land plants, such as sugar cane and maize. The report by Reinfelder and colleagues is the first description of C4 photosynthesis in a marine microalga, and the data suggest that C4 carbon metabolism may be confined to the cytoplasm, separated from the RUBISCO-dependent reactions within the plastid. C4 photosynthesis in diatoms may provide a further explanation for their ecological success in the world's oceans.

Like a variety of algal cells diatoms generate fatty acids, including the valuable long chain poly-unsaturated fatty acids to store metabolic energy. Fatty acids accumulate in the cell, thus forming oil droplets (Tonon et al. 2002). The principal energy storage polysaccharide of diatoms is chrysolaminaran. The relatively high contribution of chrysolaminaran to marine particulate matter underline this molecule's significant role in the oceanic cycling of carbon (Van Oijen et al. 2005). Between 10 and 20% of the total cellular carbon in exponentially growing diatoms are provided by chrysolaminaran, but it can accumulate to up to 80% of the total cellular carbon in cells stressed by nitrogen limitation (Varum and Myklestad 1984). The structure of chrysolaminaran is fundamentally different from the starches and glycogens as it based on a β-1,3-linked glucan backbone (Chiovitti et al. 2004; Alekseeva et al. 2005; Storseth et al. 2005), which is infrequently branched with β-1,6-linked glucosyl residues. Instead of being stored inside the plastids several diatom species, including *Phaeodactylum tricornutum*, were shown to store chrysolaminaran in the vacuole (Chiovitti et al. 2004).

For the vast majority of diatom species, the Petri-dish nature of the frustule and its unusual mode of biogenesis lead to a reduction in size during successive mitotic divisions in one of the daughter cells. Mitotically dividing diatom populations therefore decrease in size over time. Regeneration of the original size typically occurs via sexual reproduction, followed by auxospore formation. Gametogenesis occurs once cells decrease in size to approximately 30–40% of the maximum diameter. This is known as the critical size

threshold. The resulting male and female gametes combine to create a diploid auxospore that is larger than either parent. This newly created cell then proceeds along the asexual pathway until an appropriate trigger once again elicits gametogenesis. Sexual reproduction in diatoms involves a range of mechanisms (reviewed in (Mann 1993)). In centric diatoms, sex is almost universally oogamous, with flagellated male gametes. Within the pennate diatoms, there is much more variety, including anisogamy, isogamy, and automixis. Only fragmentary information is available because almost all studies are based on microscopic observation of what is a very rare event. Diatom sexuality is in fact limited to brief periods (minutes or hours) that may occur less than once a year in some species and that involve only a small number of vegetative cells within a population (Mann 1993).

Evolution of Diatoms

Chloroplasts have their origin in a process termed primary endocytobiosis: a photosynthetic cyanobacterium was engulfed by a unicellular eukaryotic heterotroph and subsequently converted into a chloroplast (van den Hoek et al. 1997; Delwiche 1999). This primary endocytobiosis event might have happened as early as 1.6 billion years ago (Bhattacharya and Medlin 2004) and was a fundamental step in eukaryotic evolution. Primary endocytobiosis gave rise to three basal lineages of eukaryotic algae - rhodophyta (red algae), glaucophyta and chlorophyta (green algae) from which all land plants are derived (Cavalier-Smith 1998). Secondary endocytobiosis occurred at least twice and led to a drastic increase in algal diversity. While the incorporation of a green alga by a eukaryotic host gave rise to Euglenoids, Dinophytes and Chlorarachniophytes, the incorporation of a rhodophyte by a heterotrophic flagellate (related to the Oomycetes) led to the evolution of all Chromists including Cryptophyta, Dinophytes, Haptophyta and Stramenopiles (Heterokontophyta) (Gibbs 1981; Delwiche and Palmer 1997) (Fig. 2).

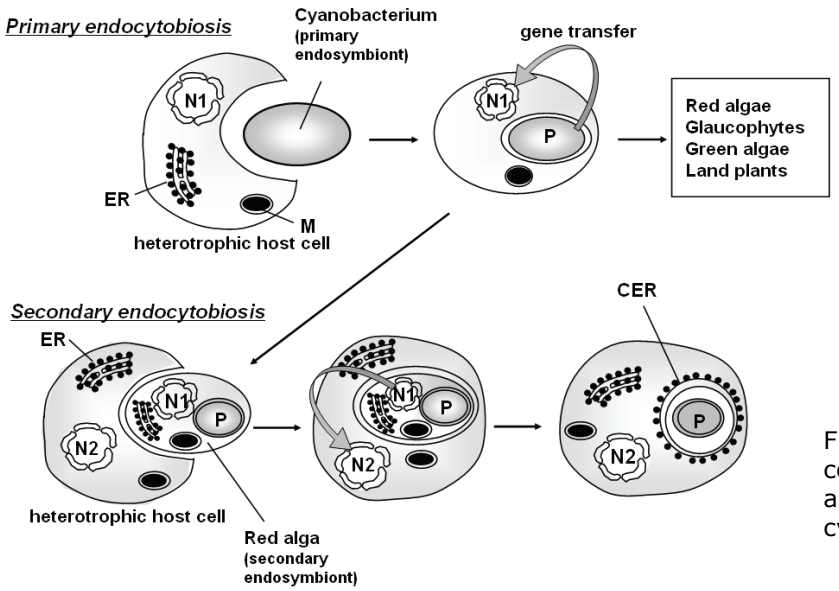


Figure 2: Evolution of complex plastids. Primary and secondary endo-cytobiosis.

Recent studies indicate that the secondary endocytobiosis of a red algae occurred already 1.3 billion years ago (Bhattacharya and Medlin 2004). Diatoms finally arose from the latter lineage, the Stramenopiles. Fossil diatoms are dated back to the Cretaceous, 144–65 million years ago (Falkowski et al. 2004), their evolutionary history may even extend as far back as 550 million years (Cavalier-Smith 2003). Recent diatoms have diversified into a wide range of over 250 genera, with perhaps as many as 100,000 living diatom species (Round et al. 1990; Norton et al. 1996). The complex or secondary plastids of diatoms differ fundamentally from the two major plastid lineages generated by primary encytobiosis, the chloroplasts and rhodoplasts. Green algae and the higher plants, contain chloroplasts featuring stacked thylakoid membranes and the accessory pigments chlorophyll a and b. In contrast, red algae contain rhodoplasts, which use chlorophyll a and phycobilisomes to capture light energy. While both, chloroplasts and rhodoplasts are surrounded by two membranes, the complex plastids of some Chromists, including diatoms, possess four membranes. While the two innermost membranes are considered to represent the plastid envelope membranes of the eukaryotic endosymbiont, the origin of the outer membranes is not yet fully resolved. The two outer membranes might descend from the plasma membrane of the endosymbiont and the host endomembrane system respectively. In cryptophyta and heterokontophyta algae (brown algae, diatoms, and related algae), which all have four-membrane plastids, the outermost membrane may be continuous with the host´s endoplasmic reticulum (ER) membrane system (Gibbs 1979; Ishida et al. 2000) and is therefore named CER (chloroplast endoplasmic reticulum) (Bouck 1965) (Fig.3).

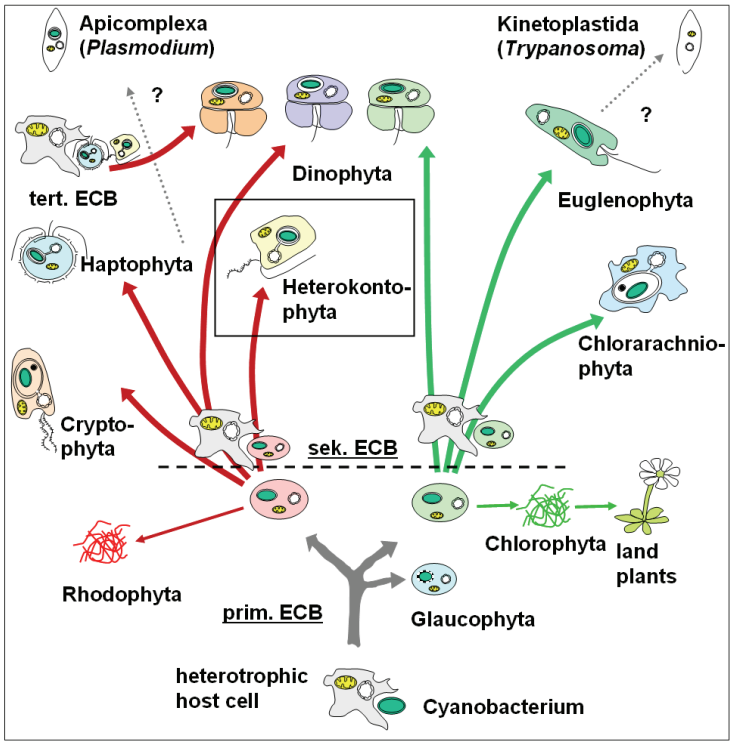


Figure 3: Evolution of algae by primary and secondary endocytobioses. (modified after Delwiche *et al.* 1999) The basal lineages of chloroplasts may be traced back to one primary endocytobiosis (prim. ECB) event in which a cyanobacterium has been taken up by a eukaryotic heterotrophic cell and subsequently was transformed into a chloroplast. The ancestral alga which evolved
by primary endocytobiosis split into the three basal algal lineages: the red algae (rhodophyta), green algae (chlorophyta) and the glaucophyta. At least two secondary endocytobioses (sec. ECB) events led to the evolution of a variety of other eukaryotic algae including nonphotosynthetic eukaryotes.

Phaeodactylum tricornutum – a model organism for diatoms

This dissertation focuses on *Phaeodactylum tricornutum*, a pennate diatom (Fig. 4) which became a convenient laboratory strain and consequentially was discussed as a model organism for diatoms. *Phaeodactylum tricornutum* is a rather atypical diatom in that it is polymorphic. It exists as three different morphotypes (oval, fusiform, and triradiate), which are lacking the typical highly ornamented silicified cell wall (Lewin et al. 1958; Borowitzka et al. 1977; Borowitzka and Volcani 1978; Mann 1993). Previous works on *Phaeodactylum tricornutum* contributed to our understanding of photosynthesis and photoprotection in diatoms (Lohr and Wilhelm 1999; Goss et al. 2006; Lavaud and Kroth 2006), other studies focused on aspects of the carbohydrate metabolism (Kroth et al. 2005; Michels et al. 2005). Furthermore, protein import through the four membranes surrounding the diatoms' complex plastids is intensively studied in *Phaeodactylum tricornutum* (Kroth and Strotmann 1999; Kroth 2002; Kilian and Kroth 2004, 2005).

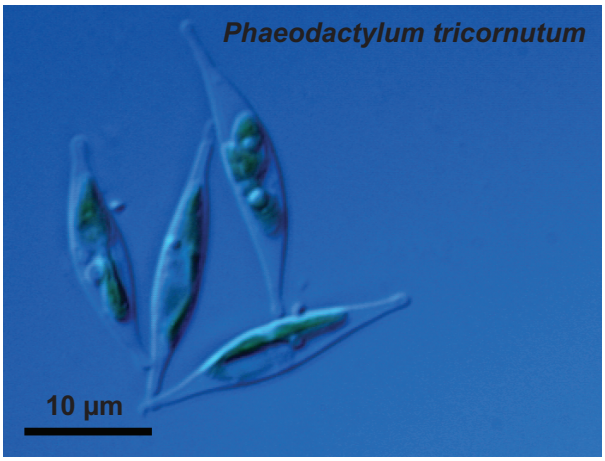


Figure 4: The pennate diatom Phaeodactylum tricornutum.

Systems for the genetic transformation basing on helium-accelerated particle bombardment have been developed for a small number of different diatoms (Dunahay et al. 1995; Apt et al. 1996; Falciatore et al. 1999; Fischer et al. 1999). However, genetic transformation technologies are most advanced for *Phaeodactylum tricornutum* in which a range of antibiotic resistance genes can be used to select for transgenic clones, including phleomycin (zeocin), kanamycin, and nourseothricin (Apt et al. 1996; Falciatore et al. 1999; Zaslavskaia et al. 2000). This convenient transformation technique has been recently applied to perform first but nevertheless significant steps in metabolic engineering. Zaslavskaia et al. (Zaslavskaia et al. 2001) reported the trophic conversion of *Phaeodactylum tricornutum*, which is obligate photoautotroph, into a heterotroph by metabolic engineering. In this work, genes encoding glucose transporters from human erythrocytes (glut1) or from the microalga Chlorella kessleri (hup1) were expressed in *Phaeodactylum tricornutum*. The transgenic cells exhibited glucose uptake and were able to grow in the absence of light. The trophic conversion of *Phaeodactylum tricornutum*

might be useful for large-scale cultivation of diatoms, thus allowing their commercial exploitation. Furthermore, this conversion can facilitate the generation of photosynthetic mutants, which lost the capability of photoautotrophic growth. The great value of photosynthetic mutants to researchers who study photosynthesis has been proved previously in the green alga *Chlamydomonas reinhardtii* (Grossman 2000; Harris 2001).
However, diatoms such as *Phaeodactylum tricornutum* have not only been studied because of their fascinating physiology or ecological relevance. Recently diatoms also entered the focus of biotechnology (Drum and Gordon 2003; Lebeau and Robert 2003a, 2003b). While several biotechnological applications have been discovered for various microalgae – not only diatoms – with respect to both the synthesis of intra or extracellular compounds and biomass, diatoms are especially of interest for nanotechnological application of their capability to build silicious frustules. While most fabrication techniques in nanotechnology involve planar lithographic approaches which have limited 3D capabilities, diatoms are able to generate threedimensional structures right away (Parkinson and Gordon 1999; Zaouk et al. 2006).
Because of this ever growing interest in understanding physiological properties and the ecological importance of diatoms, but also to meet the needs of biotechnological demands, "diatomists" emphasised more and more the obvious need of a model organism and appropriate molecular tools. Lopez et al. complained about studies on diatom biology being 'hampered in the past by the lack of a model species and associated molecular tools' (Lopez et al. 2005). Also Falciatore & Bowler clearly stated in a recent review that 'with increasing interest in understanding the ecological importance of diatoms, it has become clear that more molecular tools must be developed' (Falciatore and Bowler 2002). The authors especially emphasized the need to develop systems allowing 'the specific inactivation of specific genes, such as antisense and sense suppression, and RNA interference (Smith et al. 2000; Zamore et al. 2000).'
Although *Phaeodactylum tricornutum* is merely of global ecological relevance it appeared to be an attractive model due to features such as a barely silicified cell wall or the short generation time which facilitate the experimental handling. More support for the role of *Phaeodactylum tricornutum* as potential model organism provided a phylogenetic analysis performed on 18S rRNA which places it in the middle of the pennate diatom lineage (Falciatore and Bowler 2002). Finally its apparently small genome (Darley 1968; Veldhuis et al. 1997), plenty of insights which are already existing for this particular diatom and the possibility to transform its nuclear genome led to the decision to sequence the genome of *Phaeodactylum tricornutum*. To date its genome has been fully sequenced (sequence information available on http://shake.jgi-psf.org/Phatr2/Phatr2.home.html), and the annotation is nearly completed. In order to make *Phaeodactylum tricornutum* now accessible to postgenomic applications and to further support its role as model

organism this dissertation focuses on the development of different molecular tools not yet available for diatoms.

II Main Chapters

Strategies for stable plastid transformation in the diatom *Phaeodacytlum tricornutum*

Arne C. Materna[1] & Peter G. Kroth[1]

1 Department of Plant Ecophysiology, University of Konstanz, Germany

II.1.1 Abstract

In order to establish a system for stable chloroplast transformation, we focused on the diatom *Phaeodactylum tricornutum* and followed three different approaches: (i) The first approach is based on the insertion of the streptomycin-resistance gene *aadA* into the plastid encoded RUBSICO-operon as a third operon gene. AadA was therefore provided with operon specific ribosome binding sites and a termination sequence. The resistance gene was targeted either within the two operon genes rbcL and rbcS or at the end of the operon. (ii) Another approach relying on *aadA* as marker gene based on the construction of resistance cassettes which were designed to guarantee a flawless expression of *aadA* in the host organelle. The gene *aadA* was flanked by two different host-specific promoters and terminator sequences. In order to avoid impaired expression of functional plastidic genes, the resistance cassettes were targeted into the non-coding intergenic region downstream of rbcS. (iii) For the third strategy, transformation vectors were designed to replace the functional *psbA* gene by a slightly altered version of this gene, which increases herbicide tolerance.

The obtained results indicate that the resistance can be transiently expressed in the diatoms' plastids. However, permanent expression of the marker gene in the plastids did not occur. Since the targeted insertion of the marker could not be verified it is likely that the resistance gene is either transiently expressed from episomal transformation vectors or temporarily integrates into the plastid genome via heterologous recombination.

Key words: biolistic, particle gun, *Phaeodactylum tricornutum*, plastid transformation

II.1.2 Introduction

Plastids are cellular organelles in plants and algae which contain a reduced genome (plastome or ptDNA). Together with mitochondria they have retained numerous eubacterial features, such as gene organization in operons and the prokaryotic mechanism of transcription and translation. The plastome is a highly polyploid circle of double-stranded DNA which varies between 120 kb to 180 kb in size and harbours approximately 120 genes. A salient feature of the plastid genome is a sequence duplication, with a size of approximately 25 kb and in an inverted orientation termed "inverted repeats" (Palmer 1985; Sugiura 1992; Wakasugi et al. 2001). Transformation of plastid genomes has been intensively used in functional genomics by performing gene knock-outs, overexpressions and site-directed mutagenesis of plastid genes. This technique was used, for instance, to investigate the presence of RNA editing in tobacco chloroplasts via introduction of a heterologous editing site from spinach (Bock et al. 1994). Knock-out experiments basing on the disruption or deletion of conserved plastome open reading frames of unknown function (*ycf* genes) helped to reveal the function of *ycf*3 (Ruf et al. 1997), *ycf*6 (Hager et al. 1999), and *ycf*9 (Maenpaa et al. 2000; Ruf et al. 2000; Baena-Gonzalez et al. 2001; Swiatek et al. 2001). To date, more than 20 different plastome encoded genes (listed by (Maliga 2004) have been deleted via plastid transformation. Plastid transformations using fluorescing reporter genes (eGFP fusion proteins) were also developed to monitor e.g. gene expression and regulation (Franklin et al. 2002). These studies have contributed greatly to our understanding of the physiology and biochemistry of biogenetic processes inside the plastid compartment. Furthermore, plastid transformation provides promising opportunities for biotechnological applications. Engineering the plastome allows expression and accumulation of transgenic proteins at high levels due to the high numbers of copies of plastomes present in a plant or algal cell (reviewed by (Maliga 2003). Also in terms of bio-safety, transformation of plastids provides a striking advantage in comparison to nuclear transformation: the chloroplasts in plants are usually maternally inherited and therefore not transmitted by pollen. This allows containment of transplastomic plants. While strict maternal inheritance of plastids was shown for some angiosperms including Zea mays, Glycine max, Oryza sativa, and Arabidopsis thaliana (Avni and Edelman 1991; Maliga 2004), strict maternal inheritance of plastids has been questioned in Nicotiana species (Avni and Edelman 1991). Low frequency pollen transmission of plastids was also reported in a Setaria italica cross (Wang et al. 2004). Transforming plastids of microalgae instead of higher plants might therefore minimize the risk of transgene flow. Large scale clonal cultivation of transplastomic microalgae in bioreactors not only could guarantee safe

handling of the modified organisms but may also meet the requirements of biotechnological applications regarding productivity and scale.

To date transformation of plastid genomes has been performed in *Chlamydomonas reinhartii* (Kindle et al. 1991; Xiong and Sayre 2004), in the red alga *Porphyridium spec*. (Lapidot et al. 2002), in *Euglena gracilis* (Doetsch et al. 2001), tobacco (*Nicotiana tabacum*), and also in industrial relevant plants such as cotton (*Gossypium hirsutum*) (Kumar et al. 2004) or food plants such as tomato (*Lycopersicon esculentum*) (Ruf et al. 2001), potato (*Solanum tuberosum*) (Sidorov et al. 1999), soybean (Dufourmantel et al. 2004), lettuce (*Lactuca sativa*) (Lelivelt et al. 2005) and even in the white poplar (*Populus alba*) (Okumura et al. 2006).

The initial attempts to transform plastids had to deal with the problem of delivering DNA through the physical barrier of at least two membranes surrounding the organelles. While no bacteria or viruses were known to introduce foreign DNA into the organellar genomes, finally, a rather rough method succeeded: the shooting with DNA coated inert metal powder using particle-accelerating devices, nowadays commonly referred to as particle guns. Together with the development of efficient protocols for coating gold or tungsten particles with nucleic acids, this biolistic (biological + ballistic) technique has provided the attractive opportunity to shoot foreign DNA into living cells (Klein et al. 1992). Although the biolistic method is undoubtedly the currently most widespread technology for plastid transformation, stable introduction of cloned DNA into plastomes has also been achieved using two alternative protocols. By agitating a suspension of glass beads and cell wall-deficient *Chlamydomonas* cells in the presence of plasmid DNA transplastomic cells were obtained, however at significantly lower rates than by the biolistic protocol (Kindle et al. 1991). The tobacco plastid genome was alternatively transformed by chemical treatment of protoplasts with polyethylene glycol (PEG) in the presence of vector DNA (Golds et al. 1993; O'Neill et al. 1993; Koop et al. 1996). Finally, femtoinjection techniques also led to transient transformation (van Bel et al. 2001).

Plastid transformation experiments so far rely on three different types of marker genes to select positive transformants: (i) dominant antibiotic-resistance genes which actively confer resistance by mediating detoxification of the selective agent (Goldschmidt-Clermont 1991; Carrer et al. 1993; Bateman and Purton 2000). A frequently used dominant marker is the bacterial *aadA* gene, coding for the aminoglycoside 3'' adenyl transferase conferring Strep/Spec resistance. AadA was the first example of a foreign gene expressed in a chloroplast (Goldschmidt-Clermont 1991). Another dominant marker gene is aphA-6 coding for a aminoglycoside phosphotransferase conferring Kanamycin resistance (Huang et al. 2002). (ii) The second type of markers includes recessive antibiotic-resistance markers which encode antibiotic-insensitive alleles of ribosomal RNA genes (Newman et al. 1990; Svab et al. 1990). Another example for this marker type is the AHAS gene coding for aceto-hydroxy acid synthase which is the target enzyme for

the herbicide sulfometuron methyl (SMM). AHAS$^-$ mutants have been reverted transforming the cells with wild-type sequence, thus re-conferring SMM resistance (Lapidot et al. 2002). (iii) Finally recessive markers restore for instance photoautotrophic growth by complementing non-photosynthetic mutants (Boynton et al. 1988).

Besides selecting positive transformants, another important function of the selective agent is to mediate genome segregation of the integrated marker genes until homoplasmy is achieved. Due to the high plastid genomes copy number, presumably only one to a few genomes integrated the marker successfully after transformation. In order to maintain genetic stability of transplastomic cell lines, homoplasmy is required. Homoplasmy can be achieved by applying constant selection pressure during a sufficient number of cell cycles. For *Chlamydomonas*, this is simply done by re-streaking the growing colonies on fresh culture medium containing antibiotics. For tobacco, plants with a uniform population of transformed genomes are obtained by passing the primary chloroplast transformant through additional cycles of plant regeneration under antibiotic selection (Bock 2001).

Fortunately, plastids have inherited from their cyanobacterial ancestors an efficient RecA-type recombination system (Cerutti et al. 1992). Therefore, integration of the cloned marker genes into the plastid genomes usually occurs via homologous recombination. Constructs for any plastid genome manipulation then require that the sequence to be introduced into the plastid genome is flanked by regions of homology with the target area in the plastid genome (Staub and Maliga 1992; Kavanagh et al. 1999). Although the minimum lengths of sequences which are required for efficient homologous recombination are not very well defined, it is generally assumed that flanking regions of more than 400 bp on each side, chloroplast transformants are obtained at reasonable frequency (Bock 2001).

A stable plastid transformation system in *Phaeodactylum tricornutum* is a highly desirable tool which would allow intensifying genomic and physiological studies. Especially since *Phaeodactylum*'s plastid genome information became available, the application of reverse genetics is required to elucidate function of unknown *ycf* genes. Finally transforming the plastid genome is a crucial prerequisite for making the diatom model organism available to biotechnological applications. The different transformation strategies presented in this chapter are based on the biolistic method for delivering DNA into the plastids. Particle gun bombardment was already proven to be an effective tool for nuclear transformations of *Phaeodactylum tricornutum* (Apt et al. 1996; Zaslavskaia et al. 2000), hence it provides the greatest potential to deliver DNA into its plastid as well. Also regarding the requirement to transport DNA through four membranes which are surrounding the complex plastids of diatoms, the "shotgun" properties of a particle gun provide the most "penetrating" arguments.

II.1.3 Material and Methods

Strains and media

The *Phaeodactylum tricornutum* strain used for all experiments was the wildtype strain 646 from the Bohlin, University of Texas Culture Collection. *Phaeodactylum tricornutum* was grown at 22°C with continuous illumination at 35 µmol photons m^{-2} s^{-2} in Provasoli's enriched seawater (Starr and Zeikus 1993) using ''Tropic Marin'' artificial seawater at 50% concentration, compared to natural seawater. Solid media contained 1.2% Bacto Agar (Difco). *Escherichia coli* strain XL-1 Blue (Stratagene, Cedar Creek, TX, USA) was grown over night at 37°C in Luria Broth medium, using a shaker for liquid cultures. Solid media contained 1.5% Bacto Agar (Difco).

PCR and construction of plasmids

Standard cloning procedures were used (Sambrook et al. 1989). PCR was performed with a Master Cycler Gradient (Eppendorf, Hamburg, Germany) using recombinant Pfu polymerase (Fermentas, Ontario, Canada) or Triple Master Polymerase (Eppendorf, Hamburg, Germany) according to the manufacturer's instructions. For cloning and construction of the transformation vectors the commercial plasmids pCR-Script Amp, pCR TOPO XL, pGEM-T and pF1-A were used. Altering vector sequences via site-directed mutagenesis was performed using Turbo-Pfu Polymerase and DpnI (Stratagene, La Jolla, California, USA).

Biolistic transformation

Cells were bombarded using the Bio-Rad Biolistic PDS-1000/He Particle Delivery System (Bio-Rad Laboratories, Hercules, Canada) fitted with 1350 psi rupture discs. Tungsten particles (1.1 µm, 0.7 µm or 0.38 µm median diameters) or gold particles (1.0 µm median diameter) were coated with 5 µg of plasmid DNA in the presence of $CaCl_2$ and spermidine, as described by the manufacturer. One hour prior to bombardment approximately 10^8 cells were spread in the centre of a plate containing 20 ml of solid medium. The plate was positioned at the second level within the Biolistic chamber for bombardment. Bombarded cells were allowed to recover for 24 h before being suspended in 1 ml of sterile 10% or 50% artificial seawater medium. 250 µl of this suspension were plated onto solid 10% artificial seawater medium containing 200 µg/ml streptomycin or onto solid 50% artificial seawater medium containing $5 \cdot 10^{-6}$ M DCMU. The plates were incubated at 20°C under constant illumination (35 µmol photons$\cdot m^{-2} \cdot s^{-1}$) for four weeks.

Preparation of nucleic acids and sequencing of the mutated target genes

Resistant mutant colonies were scratched from solid media plates and resuspended in 100 µl 10% or 50% seawater medium. After centrifugation the supernatant was removed and 25 µl of CTAB buffer containing 1% β-Mercaptoethanol (Doyle and Doyle 1990) was added followed by crushing the cells with pistils fitting into the 1.5ml reaction tubes. CTAB buffer containing 1% β-Mercaptoethanol was added to a final volume of 150 µl. The suspension was incubated at 65°C for one hour. After cooling down to room temperature one volume of Chloroform / Isoamylalkohol (at the ratio 24:1) was added to the suspension and mixed by inverting the reaction tubes gently. The solution was subsequently centrifuged with 16000 g for ten minutes at room temperature. The aqueous phase was mixed with one volume of Isopropanol. The DNA was allowed to precipitate at room temperature for up to two hours. This was followed by centrifugation with 16000g for 10 min at room temperature, the resulting DNA pellet was washed with 80% Ethanol. After drying the DNA was resuspended in 30 µl Tris-HCl 10 mM, pH 7.5. The DNA preparation was used as template for amplification of *psbA* or the 16S rRNA gene. The resulting amplicons were sequenced by GATC (Konstanz, Germany).

II.1.4 Results and discussion

Finding the right resistance gene for screening putative transformants

Initially several antibiotics were tested with respect to their capabilities as a selectable marker. A crucial prerequisite for the tested selective agents was their more or less exclusive effect on the prokaryotic system of the chloroplast. For this purpose *Phaeodactylum tricornutum* cells were plated on agar plates containing different

antibiotic	media salt concentration	Growth* (1 week)	Growth* (3 weeks)	antibiotic conc. [µg ml^{-1}]**
gentamicin	10%	+	-	200
	25%	-	-	500
	50%	+	+	1000
kanamycin	10%	+	+	1500
	25%	+	-	500
	50%	+	-	2000
spectinomycin	10%	+	+	1000
	25%	+	+	1000
	50%	+	+	1000
streptomycin	10%	-	-	150
	25%	(+)	-	500
	50%	+	+	1000
tetracyclin	10%	-	-	8
	25%	-	-	8
	50%	-	-	8

Table 1: Antibiotic resistance profiles for *Phaeodactylum tricornutum*. * : + = cell growth; **-** = no cell growth i.e. no resistance. ** i.e. the lowest antibiotic concentration on which growth was still prevented or, if a "+" indicates resistance, the highest tested concentration.

concentrations of a variety of antibiotics. First tests revealed that some of the important aminoglycosidic antibiotics are inactivated by high salt concentrations, therefore effects of antibiotics were tested under various salt concentrations. *Phaeodactylum tricornutum* was cultivated on media plates containing 10%, 25% or 50% of seawater salt concentrations. Defined aliquots of these cultures served as inoculum for resistance-experiments using the antibiotics streptomycin, spectinomycin, kanamycin, gentamicin and tetracycline at concentrations ranging from 10-2000 µg ml^{-1}. The results of this antibiotic screen (Table 1) demonstrate that genes conferring resistance to gentamicin (aph(3`) II), kanamycin (aphA-6), streptomycin (*aadA*) and tetracycline (tetR) are suitable selection markers in *Phaeodactylum* plastid transformation when applied at reduced salt concentrations.

Additionally the two urea class herbicides atrazine and diuron (DCMU) have been tested with respect to growth inhibition of *Phaeodactylum tricornutum*. The utilization of herbicides as selection markers is of interest because of the possible usage of plastid transformation for biotechnological purposes. First experiments show that *Phaeodactylum*

tricornutum shows sensitivity to both atrazine and diurone, however at concentrations which are usually used for selecting already resistant Cyanobacteria. The inhibitory effect of the herbicides was not reduced by high salt concentrations (table 2).

herbicide	media salt concentration	Growth* (1 week)	Growth* (3 weeks)	hernicide conc. [µg/ml]]**
atrazine	10%	-	-	30
	25%	-	-	30
	50%	-	-	30
diuron (DCMU)	10%	-	-	5
	25%	-	-	5
	50%	-	-	5

Table 2: Herbicide resistance profiles for *Phaeodactylum tricornutum*. * : + = cell growth; −= no cell growth i.e. no resistance. ** i.e. the lowest tested herbicide concentration.

Further experiments revealing the minimal inhibitory concentrations (MIC) suitable for selection of herbicide resistant *Phaeodactylum tricornutum* cells show that growth was inhibited on $5 \cdot 10^{-6}$ M atrazine and by diuron at concentrations between $1\text{-}5 \cdot 10^{-6}$ M

herbicide	media salt concentration	Growth* (1 week)	Growth* (2 weeks)	hernicide conc. [M]**
atrazine	50%	-	-	$5 \cdot 10^{-6}$
atrazine	50%	**+**	**+**	$1 \cdot 10^{-6}$
diuron (DCMU)	50%			
diuron (DCMU)	50%			

Table 3: Minimal inhibitory concentrations of atrazine and diurone. * : + = cell growth; - = no cell growth i.e. no resistance. ** i.e. herbizide concentration, boulder letters indicate the minimal inhibitory concentration.

(Table 3).

Resistance to both herbicides was found to be conferred to landplant plastids and cyanobacteria by point mutations in the *psbA* gene encoding for the PS II protein D1.

The results of the resistance-experiments demonstrate a high natural tolerance against various different selective agents. *Phaeodactylum tricornutum*'s high tolerance against a variety of further agents was previously also reported by Apt et al. (Apt et al. 1996). Except tetracycline which was shown to be highly effective even at low concentration the other applied antibiotics inhibited growth only at concentrations above 150µg/ml when media salt concentrations were reduced. For establishing a stable plastid transformation system, the effective concentrations of the appropriate selective agent should not be too high, thus allowing convenient dosing of the agent. Furthermore, these rather expensive drugs should be applied in moderated concentrations to keep the established experimental setup also cost efficient. However, very high stringency, as shown for tetracycline, might also be problematic since the antibiotic effect should be rather restricted to the prokaryotic system instead of affecting the whole organism. Therefore

we found that streptomycin appears to be the most suitable selective agent for plastid transformation in *Phaeodactylum tricornutum*. The respective resistance gene is *aadA*, an originally bacterial aminoglycoside 3´´-adenylyltransferase gene, which was the first chloroplast specific antibiotic resistance marker conferring resistance to a number of antibiotics of the aminoglycoside type, including spectinomycin and streptomycin (Goldschmidt-Clermont 1991). The AadA protein catalyzes the covalent transfer of an AMP residue from ATP to spectinomycin, thereby converting the antibiotic into an inactive form (adenylylspectinomycin) that no longer inhibits protein biosynthesis on prokaryotic 70 S ribosomes as present in the chloroplast. AadA is meanwhile routinely used as resistance gene for chloroplast transformation, since aminoglycoside antibiotics feature a high specificity as prokaryotic translational inhibitors and have low side effects on plant cells.

Construction of transformation vectors

In order to establish a stable plastid transformation system for the diatom *Phaeodactylum tricornutum* based on homologous recombination we followed three different strategies to implement modifications in the chloroplast genome: (i) The first strategy based on the insertion of the dominant resistance-gene *aadA* into the Rubisco operon as a third operon gene. The transformation vectors for this approach do not require host-specific promoter or terminator sequences. (ii) In contrast, for the second approach a resistance cassette was designed which consisted of the *aadA* gene provided with a host specific promotor, the 5'-UTR (untranslated region) containing the ribosome binding site (Shine-Dalgarno sequence) and a termination sequence to enable correct expression. The transformation vectors contained different resistance cassettes flanked by linker sequences homologous to the target area in the plastome. (iii) For the third strategy a recessive marker was applied: the transformation vectors were designed to replace the functional *psbA* gene by a modified *psbA* version conferring herbicide resistance.

The transformation vectors were designed with respect to suitable options for screening putative *Phaeodactylum tricornutum* transformants. Therefore, the chosen marker genes confer either streptomycin resistance or herbicide resistance (diurone and atrazine) since *Phaeodactylum* was shown to be strongly inhibited in the presence of both types of agents.

Inserting *aadA* into the Rubisco operon

In order to compensate the initial lack of plastome sequence information, rbcL/rbcS-operon was chosen as target for plastid transformation. The polycistronic transcription unit enables the insertion of a marker gene as additional operon gene (Fig. 1-I).

Streptomycin resistance was supposed to be conferred by inserting the *aadA* gene either in between the two genes of the Rubisco operon rbcL and rbcS (LAS transformation vector) or at the 3´-end of the operon (SAE transformation vector) (Fig. 1+2). In order to insert the resistance gene between rbcL and rbcS a ribosome binding site and a 5´-UTR had to be provided upstream of the resistance gene, while downstream a 3´-UTR was required. For this purpose, the amplified homologous linkers overlap (Fig. 1-II), creating an additional rbcL-3´-UTR and a termination loop at the 5´-end of the linker fragment downstream of the resistance gene. By modifying the sequences of the amplification primers, corresponding restriction sites were inserted at both ends of the homologous linkers and at both ends of the resistance gene (Fig. 1-II+III), thus allowing the construct assembly via subsequently ligating the fragments into the vector pGEM-T. To avoid possible suppression of rbcL/rbcS expression due to integration of the marker, a second transformation vector was designed to insert the resistance gene downstream of the rbcS-gene, so that it is still under control of the rbcL/S-operon (Fig. 2). For this purpose a specially designed primer was used for amplification of the upstream-linker, creating an additional sequence containing a ribosome-binding site and a copy of the rbcS 5´-UTR at the 3´-end of the linker. The downstream linker contains the rbcS 3´-UTR to facilitate marker gene expression (Fig. 2-II). The linkers as well as the resistance gene are equipped with corresponding restriction sites for precise ligation of the single fragments (Fig. 2-III) in pGEM-T. The constructs were named according to the position of the resistence gene (LAS for the construct containig *aadA* within the operon genes rbcL/S, and SAE for the construct which inserts *aadA* at the end of the operon).

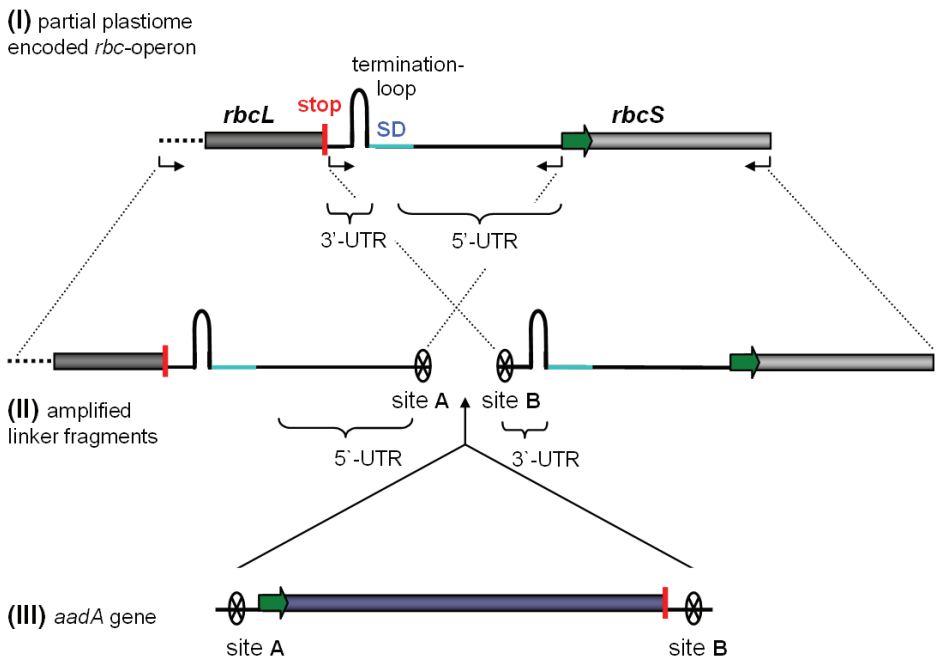


Fig. 1: Construction of the LAS construct for *aadA* insertion into the *rbcL/S*-operon**.** Based on the original sequence of the plastidal *rbc*-operon **(I)** the amplification of two overlapping linker fragments was performed roviding a Shine-Dalgarno site (SD) and operon-typical 5´- and 3´-UTR´s to the insert. These linkers were equipped with different restriction sites (A+B) by using specially designed primers **(II)**. The resistance gene also contains corresponding restriction sites (A+B) thus placing the resistance gene between *rbcL* and *rbcS* serving as linkers **(III)**.

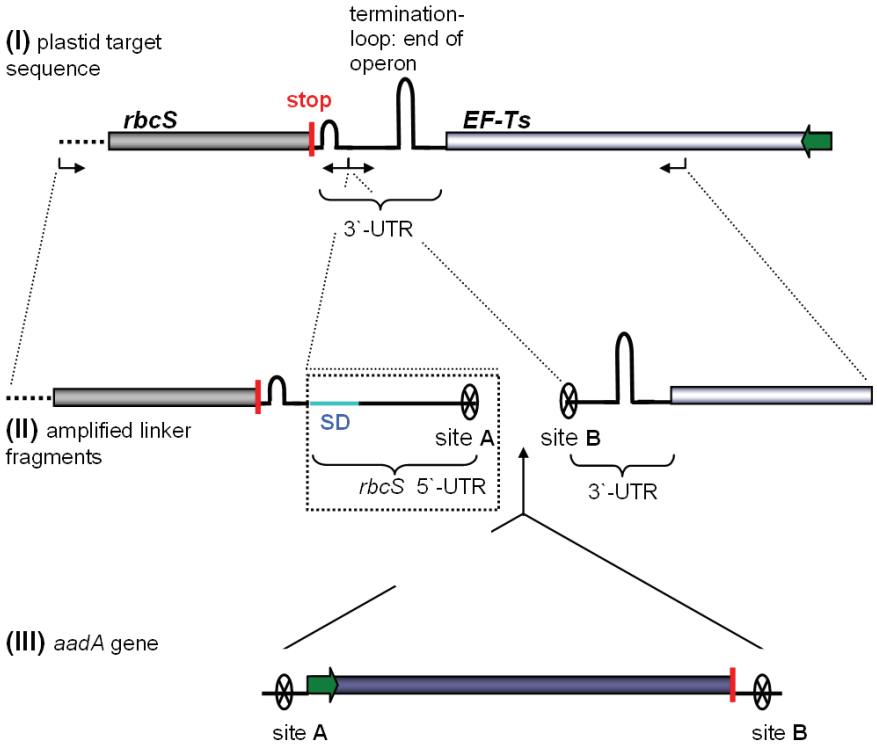


Fig. 2: Construction of a SAE construct for *aadA* insertion downstream of the *rbcL/S*-operon**.** From the original plastidal sequence **(I)** the amplification of two fragments was performed using specially designed primers to alter the linker sequences thus providing a Shine-Dalgarno site (SD) and operon-typical 5´- and 3´-UTR´s to the insert. These linkers are equipped with different restriction sites (A+B) placing the resistance gene downstream of *rbcS* at the end of the operon **(II)**. The resistance gene also contains corresponding restriction sites (A+B) for ligation between the linkers **(III)**.

After assembling the final transformation vectors both constructs have been successfully verified by sequencing. Additionally the constructs' ability to confer resistance to a prokaryotic system was checked by transforming the LAS- and SAE-vector in a non resistant *E. coli* strain (XL-1 blue). After the transformation the *E. coli* cells were able to grow in Streptinomycin while growth of a negative control was inhibited. Both LAS and SAE transformation vectors were used for various transformation attempts. The transformation experiments were performed using a "biolistic" Particle Delivery System (BioRad) utilizing 1.1 µm and 0.7 µm tungsten microcarriers that were delivered to the cells after being coated with vector DNA. Parameters like the cell-to-microcarrier distance and the recovery time (phenotypic expression) were tested. After one or two days of recovery the transformed cells were resuspended and transferred to agar plates (about $5 \cdot 10^7$ cells per plate) containing selective media. The transformation efficiencies at all applied antibiotic concentrations were in the same order. On plates containing up to 200 µg/ml the average transformation efficiency was $0.75 \cdot 10^{-6}$ (cfu´s per bombarded cell) with efficiencies up to $1{,}24 \cdot 10^{-6}$. In all transformation experiments, both the cell-to-microcarrier distance and the different recovery times did not affect the amount of putative transformant colonies after plating on selective media. After four weeks of selection the cfu`s were re-plated on plates containing the same or higher streptomycin concentrations. Most clones obtained from re-plating putative transformant cfu`s were able to grow on the applied streptomycin concentrations, while wild-type control cells did not survive.

After this first re-plating step, cfu's were used for PCR-analysis. In first experiments, the *aadA* gene was amplified by colony-PCR. In almost all cases a band of the appropriate size (0.8 kb) was amplified from the cells as shown in Fig. 3, thus confirming the presence of the resistance gene in the putative transformants. However, correct insertion of *aadA* within the operon could not be verified by PCR using primers binding in *aadA* and beyond the homologous linkers. Clones obtained from re-plating were used for a second re-plating step with the same streptomycin concentrations. Unfortunately, after this second re-plating step all clones lost their capability to grow on the applied streptomycin concentration within five to seven days and finally died within the next two weeks. In a few cases cells survived but growth was limited. This growth behaviour might indicate that *aadA* did not integrate into the plastome and was expressed transiently. The negative PCR verification of correct *aadA* insertion supports transient episomal expression. Although the used transformation vectors didn't feature a host-specific promoter, the prokaryotic T7- or SP6-promoters on the plasmid might allow unspecific expression of the resistance gene. Why the transformants died after several generations remains elusive, however impaired replication of the plasmid during cell- and organelle division might explain the temporary resistance. Furthermore it is also still possible that *aadA* inserted into the Rubisco operon, however, therewith affected drastically the operon's expression, which could be lethal for the photoautotrophic cells. In the latter case *Phaeodactylum tricornutum* cells lost the ability to grow on the selective media during the segregation process.

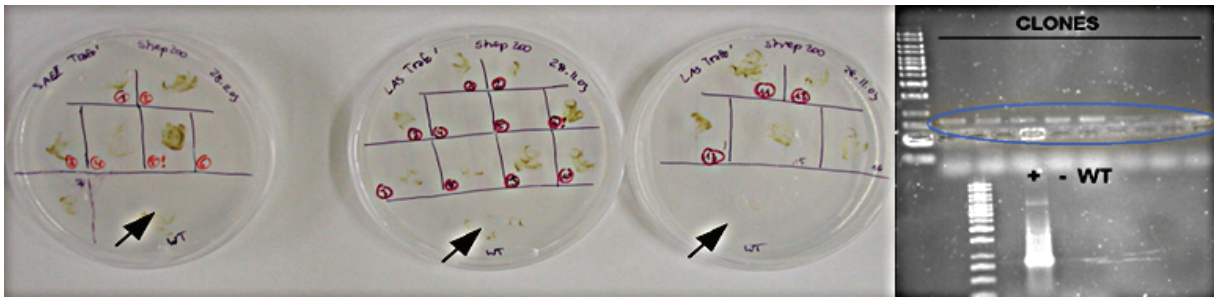


Figure 3: Plastid transformation attempts using the LAS/SAE transformation vectors. After three weeks of selection subsequent to transformation, the surviving cfu's were replated on 200µg/ml streptomycin. Fom both, LAS- and SAE-transformation approaches restreaked cfu'were able to grow on the antibiotic, while wild-type cells (black arrows) died. Colony PCR allowed the amplification of the resistance gene from LAS-/SAE-clones. No bands appeaed in wild-type or negative controls.

Sequencing of *Phaeodactylum tricornutum* plastid genome fragments

In order to obtain more sequence information on the plastid genome of *Phaeodactylum tricornutum* two large fragments have been amplified via long-range PCR. After comparing known plastome sequences of the diatom Odontella sinensis and the red alga *Porphyra purpurea* homologous regions were selected for deriving the respective

degenerated primers. The amplicons were cloned in pCR TOPO XL (Invitrogen, Carlsbad, CA, USA) and subsequently sequenced via primer-walking.
Alltogether 13043 bp of the chloroplast genome of *Phaeodactylum tricornutum* (see NCBI Genbank AY864816, AY819643) were made available this way (Fig. 4). Sequence analysis revealed that the *aadA* gene used as resistance gene contains codons which are not or only rarely used in the plastid genome. This might be an explanation for the instability of the *aadA* insertion. Moreover, the sequencing revealed continuous sequence including genes but also regulating up- and downstream sequences.

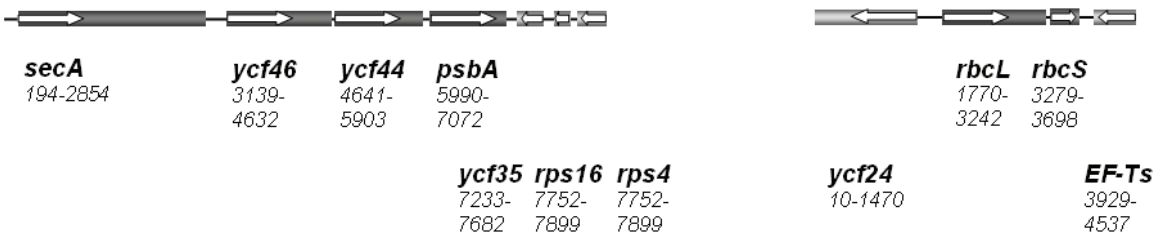


Figure 4: Plastid encoded genes revealed by sequencing two fragments of 8kb and 5kb length.

Inserting *aadA* in an intergenic region

Since the approaches to insert the antibiotic resistance gene *aadA* into the chloroplast genome under control of the Rubisco operon generated apparently only temporarily expressing transformants, further attempts were made to permanently insert *aadA* into *Phaeodactylum tricornutum*'s chloroplast genome. These attempts were based on using a marker gene with an optimized codon usage. This adapted marker gene was in addition targeted into a non-coding intergenic region instead of altering the Rubisco operon structure. The expression of the marker gene should be guaranteed by flanking the gene with host specific promoters and 5'- or 3'- untranslated regions (UTR's), thus creating a functional resistance cassette (Fig. 5A). Therefore the streptomycin resistance gene *aadA* was cloned and mutagenized, according to the codon usage of the chloroplast genome. In order not to interfere with potentially vital plastid genes the resistance cassette was inserted into intergenic regions. Further host specific upstream and downstream fragments containing both promoter and terminator are required to drive expression of the resistance gene. A variety of fragments containing the promoter and 5'-UTR or downstream sequences of Rubisco and *psbA* were cloned and assembled with the modified *aadA* gene to form complete resistance cassettes (Fig. 5A). A 0.3 kb or a 0.7 kb fragment of *psbA* upstream sequence as well as a 0.7 kb fragment of rbcL upstream sequence were ligated to the 5'-end of the modified streptomycin resistance gene. In order to complete the resistance cassette downstream of the marker gene ~ 0.2 kb of rbcS downstream sequence were added. For those constructs containing the rbcL

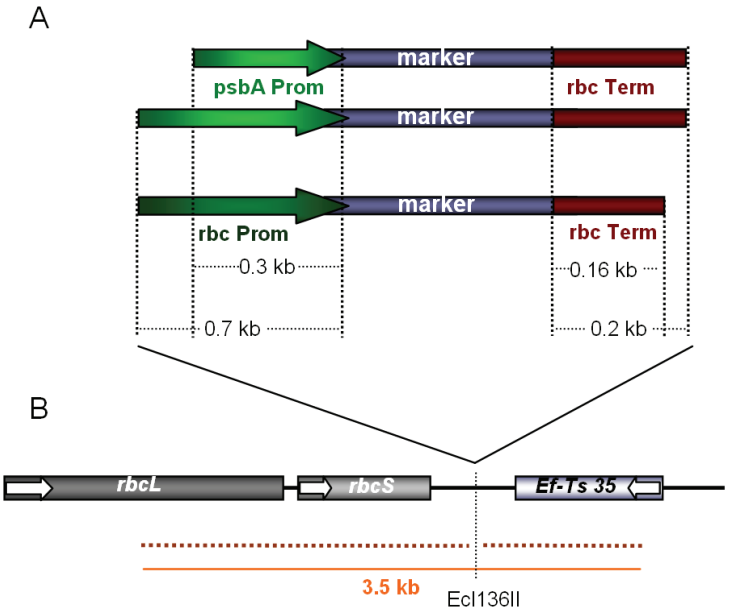


Figure 5A: Resistance cassettes for plastid transformations in *Phaeodactylum tricornutum*. The cassettes consist of 0.3kb or 0.7 kb of regulatory *psbA* upstream sequence (*psbA* Prom) and 0.2 kb of RuBisCO downstream sequence (*rbc*Term) in order to provide expression of the marker gene. When 0.7 kb of regulatory RuBisCO upstream sequence (*rbc*Prom) is used the *rbc*Term fragment is reduced to 0.16 kb to minimize unwanted recombination effects during transformation. As marker gene served the modified *aadA* gene.

Figure 5B: Homologous linkers for the resistance cassettes. To insert the resistance cassettes via homologous recombination into the chloroplast genome linker fragments around the intergenic regions downstream of the RuBisCO operon were cloned into pF1-A. For inserting the resistance cassettes a EcI136II site was added marked by the thin dotted line. Bold dotted lines indicate region of homology of the linker fragments. Both corresponding linker fragments together add up to a length of 3 to 3.5kb.

upstream sequence, only 0.16 kb of rbcS downstream sequence were added. In this case, the rbcS downstream fragment was reduced in size in order to minimize unwanted recombination effects during the transformation procedure. Transforming plastid genomes generally bases on homologous recombination of a DNA construct with the plastome region of interest. In order to target the streptomycin resistance cassettes fragments surrounding the intergenic regions downstream of the Rubisco operon were used as homologous flanking linkers (Fig. 5B). An EcoI136II restriction site was added in the center of the intergenic region to allow the insertion of the resistance cassettes. After assembly of the flanking linkers and resistance cassettes the resulting plastid transformation vectors were used for transformation experiments as described above. However, in addition to the tungsten particles, 0.1 µm gold microcarriers were utilized for the particle gun bombardment as well. After four weeks of regeneration on selective media containing 200 µg/ml Streptomycin colonies appeared with frequencies of ~1 cfu/plate (corresponding efficiency = $2 \cdot 10^{-8}$ cfu's/ cell). The cfu's were restreaked on fresh selective media plates. Replating of these cfu's was performed regularly every week. In order to verify and monitor the uptake and integration of the resistance cassettes into the chloroplast genome a part of the cassettes was amplified from CTAB-DNA extractions of the respective colonies before replating them (Fig. 6). Controls of correct insertion of the marker gene via PCR again could not verify insertion of *aadA* in the target area. As already observed for the LAS/SAE transformation attempts, once more the resistance cassettes were only temporarily detectable. The number of cfu's

allowing amplification of the resistance cassettes was decreasing with every replating step. After the third replating only one out of the eight monitored transformants still revealed a positive signal in the PCR. Accordingly, only 17% of the cfu's surviving the transformation were still alive after the third replating step (table 4). After two months of continuous segregation experiments, amplification of *aadA* from the very few surviving colonies was not possible anymore.

These findings suggest that this transformation approach generated again transformants transiently expressing the resistance gene. Stable insertion of the resistance cassette into the chloroplast genome did not occur. The fact that resistant cfu's were generated by the plastid transformation procedure may hint to transient expression of *aadA* from episomal plasmids. Since the plasmids contain functional resistance cassettes and host specific promoters the expression of the marker gene is possible independent from correct insertion. Further the utilized plasmids contain a prokaryotic origin of replication which, at least theoretically, allows replication of plasmid in dividing organelles. Heterologous recombination of the marker into the plastome might provide another explanation for the unstable phenotype which would further explain why the presence of *aadA* could be detected via PCR, however, not its correct insertion. The dense architecture of the chloroplast genome hardly contains non-coding regions. If *aadA* inserted randomly into the plastome potential negative side-effects of the marker integration would explain a loss of the marker during segregation. The fact that all transformants die during the segregation attempts indicates that (i) permanent integration does not occur, (ii) but also that the dividing diatom cells cannot replicate episomal transformation vectors in a sufficient manner. The fact that the obtained transformants survived more plating steps than the LAS-/SAE-transformants

after 4 weeks regeneration	cfu's per plate
average cfu count	1
replating No. 1 *average cfu count*	0.5
replating No. 3 *average cfu count*	0.17

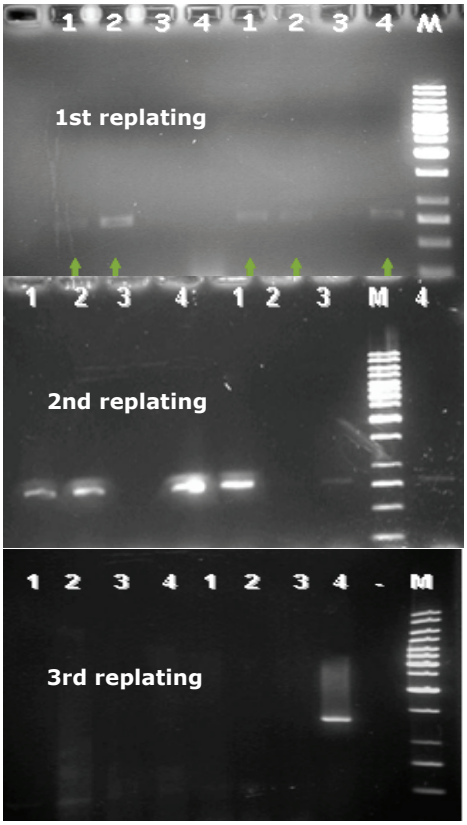


Table 4 and Figure 6: Stability of the resistance cassettes.
The relative number of cfu's obtained per plate after transformation was monitored. After every replating step the surviving cfu's were counted and compared to the initial number of cfu's per plate.
Further presence or absence of the resistance cassette in the cells was monitored for eight putative transformants by amplification of a cassette-fragment previous to every replating.
Not only the number of cfu's surviving the replating decreased but also the number of positive amplification signals thus indicating the instability of the transformation procedure.

might be due to host specific regulatory units which allow an efficient marker gene expression. In contrast, the bacterial T7/SP6 promoters on the LAS-/SAE plasmids would eventually allow only weak *aadA* expression in the prokaryotic system of the chloroplast. Unfortunately, cultivation of the putative transformants in larger volumes was not possible, since the monitored cfu's died after several replating steps. Hence, neither enough DNA nor RNA could be extracted to examine the integration of *aadA* into the plastome, nor its transient expression from episomal transformation vectors.

Insertion of point mutations in D1

The strategy of the third approach to establish a stable plastid transformation system was to avoid any negative effects potentially caused by inserting a marker gene into the plastome. This alternative approach attempted to replace a functional endogenous gene by a mutated version. Therefore a fragment containing the *psbA* gene coding for the DI protein was ligated into the pCR-Script Amp vector. Via site directed mutagenesis different point mutations were introduced into *psbA*, thus altering the cloned sequence to confer resistance against Urea and Triazine class herbicides (all introduced point mutations are listed in ChapterII). In addition to the single amino acid exchange inducing resistance a conservative mutation was inserted in close proximity. This silent mutation does not change the amino acid sequence of DI but the restriction pattern of *psbA* which was thought to allow uncomplicated screening of putative transformands. In order to find ideal conditions for homologous recombination a series of vectors (D1 transformation vectors) was designed containing a short (0.8 kb) and an elongated (3.5 kb) *psbA* fragment harboring the different point mutations (Fig. 7). These homologous inserts are designed to simply replace a fragment of wild-type sequence by slightly altered but functional sequence – no additional genes are inserted into the plastome. Altogether, the D1 vectors provide ideal conditions for introducing the inserts into *Phaeodactylum tricornutum*'s plastid genome via homologous recombination. *Phaeodactylum tricornutum* was transformed via particle gun bombardment using six different D1 vectors. After recovery from transformation the cells were suspended and transferred to plates containing selective media (about $5 \cdot 10^7$

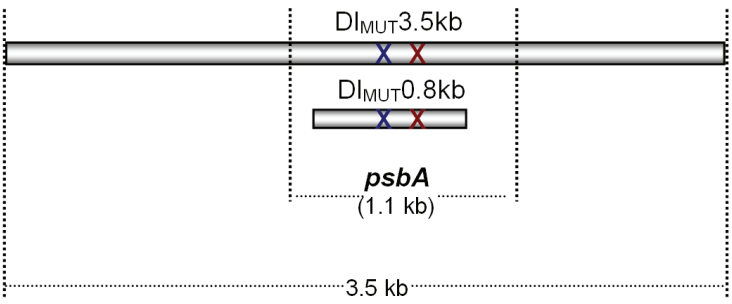


Fig. 7: Transformation vectors designed for inserting mutations into *psbA* leading to herbicide resistant variants of the DI protein. Fragments of homologous sequence (indicated by grey bars) of different length (3.5kb and 0.8kb) have been cloned. These fragments contain one resistance conferring point mutations (indicated by the blue **X**) and conservative mutation (indicated by the red **X**) altering the restriction pattern for screening of putative transformants.

cells per plate, $5 \cdot 10^{-6}$mM DCMU). Four weeks of incubation under constant light and temperature conditions yielded cfu's at an average efficiency of $\sim 10^{-8}$ (cfu's/cell). For the first replating step, the same herbicide concentrations were used. For second replating after one week of growth the herbicide concentrations were raised to $C_{DCMU} = 10^{-5}$ M and 10^{-4} M. All putative transformants were able to grow on DCMU up to 10^{-5} M, in some cases (strain DI-5) even at 10^{-4}M. All cfu's survived further replating steps and were even able to grow in liquid media containing up to $5 \cdot 10^{-4}$ M, while wild-type cells did not survive 10^{-6} M of DCMU (Fig. 8). Restriction analysis and subsequent sequencing of amplified *psbA* fragments, however, revealed that none of the monitored putative transformants carried the point mutations of the respective transformation vectors. The sequenced psbA fragments displayed instead a variety of other nucleotide substitutions in the same area. These nucleotide substitutions led to amino acid exchanges in D1, but in all cases were not identical to those which supposed to be conferred by the transformation vectors. The appearance of mutations in within the target sequence was directly caused by the transformation. All obtained mutants are listed in detail in Chapter II. In accordance with the previous results the D1 transformation attempts to indicate that DNA fragments can be delivered into the chloroplast, marker genes within these fragments are most probably expressed. Although the D1 vectors feature ideal conditions for homologous recombination insertion was never observed, indicating that targeted insertion of a marker via homologous recombination might be not feasible in this diatom's chloroplast.

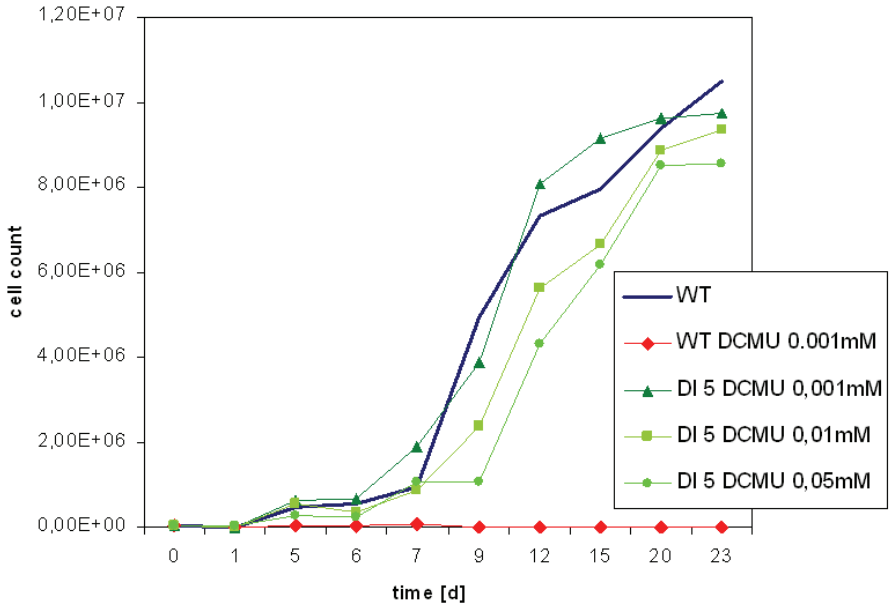


Figure 8: Growth of the *Phaeodactylum tricornutum* mutant D1-5 on various DCMU concentrations. Growth of wildtype cells without herbicide (—) compared to DI-5 cells growing in media containing 10^{-6} M (▲) , 10^{-5} M (■) and $5 \cdot 10^{-4}$ M of DCMU (●). Wildtype cells in 10^{-6} M of DCMU do not grow (♦).

Optimizing the transformation procedure

In order to test a variety of transformation parameters, transformation vectors were designed containing the modified *psbA* which was approved to confer DCMU resistance in mutant D1-5 without effecting growth at selective conditions. Fragments of 0.8 kb and 3.5 kb containing *psbA* were then ligated in pCR-Script Amp. Subsequently the point mutation found in the mutant D1-5 – encoding the substitution of Serin to Alanin at D1 amino acid position 264 – was introduced into the cloned *psbA* via site directed mutagenesis.

The resulting vectors were used for various transformation attempts under different conditions. During these optimization experiments, parameters like microcarrier size and substance (tungsten with diameters: 0.38 µm, 0.7 µm, 1.1 µm; gold with diameter: 1.0 µm) were tested with respect to their influence on the cells' ability to survive the procedure. Furthermore the recovery of the cells after transformation was tested on selection media with DCMU concentrations ranging from $5 \cdot 10^{-6}$ M to 10^{-6} M and recovery time was extended from one to two days. Additionally, the selection was performed under different light conditions ranging from low (7 and 20 µmol photons$\cdot m^{-2} \cdot s^{-1}$) to normal (45 µmol photons$\cdot m^{-2} \cdot s^{-1}$) light intensities. Finally, in order to exclude any kind of interference of the pGEM-T vector backbone during recombination, the inserts were cloned also into pBS KSII^{-}.

The optimization experiments generated resistant cfu's with an average efficiency of 10^{-8}/cell when using 1.0 µm gold microcarriers. The diameter of the tungsten particles did not influence the number of cfu´s. Using tungsten particles significantly decreased the rates, or in some cases did not yield any resistant cfu's. Changing the duration of phenotypic expression before transferring the cells on selective media plates had no effect. Since mutants with altered photosynthetic genes like *psbA* might show increased light sensitivity the transformed cells were exposed to low and normal light conditions. During these experiments only cells grown at 20-45 µmol photons$\cdot m^{-2} \cdot s^{-1}$ were able to develop cfu´s. Again most obtained cfu's survived subsequent replating steps. Negative controls did never yield any resistant cfu's. Sequencing the *psbA* gene from these colonies revealed once more that the resistance was developed due to mutations different D1-5 mutation that should have been introduced.

Construct	Vector	Insert	Particle	recovery	C_{sel}	cfus	Efficiency [cfu´s/plated cells]	Characterization
LAS*	pGEM-T	*rbcL-aadA-rbcS*	W 0.7µm	1, 2 days	90 – 200 µg/ml	√	$0.75 \cdot 10^{-6}$ - $1.24 \cdot 10^{-6}$	transient expression
SAE*	pGEM-T	*rbcS-aadA-EF-Ts*	W 0.7µm	1, 2 days	90 – 200 µg/ml	√	$0.75 \cdot 10^{-6}$ - $1.24 \cdot 10^{-6}$	transient expression
PAT-R	pF1-A	*rbcProm-aadA-rbcTerm*	W 0.7µm Au 1.0µm	1 day	200 µg/ml	√	$2 \cdot 10^{-8}$	transient expression
PAT-P (600nt *psbA* Prom)	pF1-A	*psbAProm-aadA-rbcTerm*	W 0.7µm Au 1.0µm	1 day	200 µg/ml	√	$2 \cdot 10^{-8}$	transient expression
PAT-P (300nt *psbA* Prom)	pF1-A	*psbAProm-aadA-rbcTerm*	W 0.7µm Au 1.0µm	1 day	200 µg/ml	√	$2 \cdot 10^{-8}$	transient expression
DI 0.8 $_1$*	pGEM-T	$psbA_{Mut}$ 0.8kb	W 0.7µm	1, 2 days	$5 \cdot 10^{-6}$ M	X	$1.06 \cdot 10^{-6}$	
DI 0.8 $_2$*	pGEM-T	$psbA_{Mut}$ 0.8kb	W 0.7µm	1, 2 days	$5 \cdot 10^{-6}$ M	√	$1.06 \cdot 10^{-6}$	unspecific mutations
DI 0.8 $_3$*	pGEM-T	$psbA_{Mut}$ 0.8kb	W 0.7µm	1, 2 days	$5 \cdot 10^{-6}$ M	√	$1.06 \cdot 10^{-6}$	unspecific mutations
DI 0.8 $_4$*	pGEM-T	$psbA_{Mut}$ 0.8kb	W 0.7µm	1, 2 days	$5 \cdot 10^{-6}$ M	X	$1.06 \cdot 10^{-6}$	
DI 0.8 $_5$*	pGEM-T	$psbA_{Mut}$ 0.8kb	W 0.7µm	1, 2 days	$5 \cdot 10^{-6}$ M	√	$1.06 \cdot 10^{-6}$	unspecific mutations
DI 0.8 $_6$*	pGEM-T	$psbA_{Mut}$ 0.8kb	W 0.7µm	1, 2 days	$5 \cdot 10^{-6}$ M	√	$1.06 \cdot 10^{-6}$	unspecific mutations
DI 0.8 $_2$ * (optimization) various I_{light}	pGEM-T pBS KSII$^-$	$psbA_{Mut}$ 0.8kb	W 1.1µm; 0.7 µm; 0.38µm Au 1.0µm	1, 2, days	$5 \cdot 10^{-5}$ – 10^{-6}M	√	$3 \cdot 10^{-8}$ - $1.7 \cdot 10^{-7}$	- all W particle sizes -> equal rates - Au particles –> 6 fold rates compared to W particle-rates - Only at ~30 µE / no cfu´s at lower I_{Light} - various unspecific mutations in DI
DI 3.5 $_2$*	pCR Script	$psbA_{Mut}$ 3.5 kb	W 0.7µm Au 1.1µm	1 day	$5 \cdot 10^{-6}$ M	√	$1 \cdot 10^{-8}$	unspecific mutation
DI 3.5 $_3$*	pCR Script	$psbA_{Mut}$ 3.5 kb	W 0.7µm Au 1.0µm	1 day	$5 \cdot 10^{-6}$ M	√	$1 \cdot 10^{-8}$	unspecific mutations
DI 3.5 $_4$*	pCR Script	$psbA_{Mut}$ 3.5 kb	W 0.7µm Au 1.0µm	1 day	$5 \cdot 10^{-6}$ M	√	$1 \cdot 10^{-8}$	unspecific mutations
DI 3.5 $_5$*	pCR Script	$psbA_{Mut}$ 3.5 kb	W 0.7µm Au 1.0µm	1 day	$5 \cdot 10^{-6}$ M	√	$1 \cdot 10^{-8}$	unspecific mutations

Table 2: Summary of transformation vectors used for plastid transformation attempts. LAS, SAE, PAT-R and PAT-P indicate the *aadA* containing vectors; PAT-R and PAT-P contain resistance cassettes with host specific promoter and terminator. DI indicates vectors containing mutated *psbA* genes or fragments of this gene. The numbers 0.8 or 3.5 represent the length of homologous sequence in kb, the small number refers to the mutation the vector are supposed to confer (see chapter II.2). Furthermore are listed the used cloning vectors , the transformable constructs harboured by these vectors, further size and matter of the particles used as microcarriers, time of recovery, and the applied selection pressure. This chart also shows whether cfu´s appeared, the rates of appearance and the results of the respective transformation attempt. All constructs indicated with an asterisk (*) have been shot as circular and linear molecules.

Conclusion

The different attempts to establish a stable plastid transformation system in *Phaeodactylum tricornutum* allowed us to generate temporary resistant transformants. In all cases, however, targeted introduction of the marker into the plastid genome via homologous recombination was not achieved. The attempts to insert *aadA* into the Rubisco operon might suffer from an impaired expression of *aadA* since the gene contained one codon which was later shown to be rarely found in *Phaeodactylum tricornutum*'s plastome. Moreover, the insertion of a foreign gene into a functional operon might be deleterious by impairing the operon's expression. A second series of vectors designed to insert *aadA* into the plastome should exclude some of the previous potential problems. The (overall) length of the homologous linkers was increased from 2kb (LAS-/SAE- vectors) to 3.5kb. AadA was not targeted into a functional operon but into non-coding intergenic regions in between and downstream of two genes of opposite direction. Also the rare codon in *aadA* was substituted by a common one and the resistance gene was flanked by host specific promoters and untranslated regions to guarantee efficient expression. Despite of these improvements all transformation attempts did again not allow targeted insertion of the marker gene via homologous recombination. However, the fact that both approaches to introduce *aadA* into the plastome generated transiently resistant cfu's suggests that a marker gene can be successfully expressed in the chloroplast. PCR experiments could not localize *aadA* in the target region, suggesting that the marker gene is expressed from an episome. In order to test whether insertion of a DNA fragment into the plastome can occur at all, transformation attempts were performed using the D1-transformation vector set. These vectors do not introduce any foreign or additional genes but were rather designed to substitute a functional gene by a slightly altered version conferring herbicide resistance. The cfu's generated by this approach showed high and permanent resistance. Sequencing of these putative transformants, however, revealed unspecific mutations in *psbA*, thus clearly demonstrating that substitution of the wild-type *psbA* via homologous recombination did not occur in any of the herbicide resistant strains. Varying the parameters of the particle gun bombardments during the optimization experiments had no effect on these results. All presented results have in common that transient or permanent resistance could not be attributed to a stable insertion of the selection marker; the results of the D1 approach even suggest that insertion via homologous recombination does not occur at all.

AadA appears to be transiently expressed in the chloroplast of *Phaeodactylum tricornutum*, however, whether the resistance gene is either expressed from episomal transformations vectors or if it integrates temporarily into the plastome via heterologous recombination remains elusive. The fact that both resistance and traceability of *aadA* were not permanent could be attributed to an insufficient replication of episomal

transformation vectors within dividing cells. With every division the amount of expressed *aadA* genes would dilute until streptomycin resistance no longer was provided. Another explanation for the transient resistance might be provided by a random integration of the resistance cassettes into the plastome which could disturb the expression of other potentially vital plastidic genes. In that case, the clear disadvantage of maintaining the marker insertion during segregation would also explain the transient resistance.

Although expression of the selection marker from an episomal plasmid inside the chloroplast is uncommon, it has been reported previously for plastid transformation experiments in *Nicotiana tabacum* (Staub and Maliga 1994, 1995) and *Euglena gracilis* (Doetsch et al. 2001). While homologous recombination can be achieved in tobacco plastids, unknown circumstances prevent the integration of transforming DNA into the plastid genome of *Euglena gracilis* (Doetsch et al. 2001). In *Phaeodactylum tricornutum* we achieved transient plastid transformation. However, the presented results suggest that the capability of the diatom to integrate transforming DNA into the plastome via homologous recombination is absent as well.

Induced and targeted mutagenesis in the chloroplast genome of the diatom *Phaeodactylum tricornutum*

Materna, A.[1] and Kroth, P.G.[1]

1 Department of Plant Ecophysiology, University of Konstanz, Germany

II.2.1 Abstract

We report an inducible mechanism which generates targeted mutations in a chloroplast genome (plastome). Mutagenesis was induced in *psbA* (encoding for the D1 protein of photosystem II) and the 16S rRNA gene, both encoded in the plastid genome of the diatom *Phaeodactylum tricornutum*. To trigger mutagenesis in the gene of interest the cells were transformed via particle gun bombardment utilizing plasmids which contain a sequence fragment homologous to the target gene. The type of microcarrier used to deliver the plasmid DNA into the cell was crucial for an effective mutagenesis induction. Utilizing gold particles with a median diameter of 1.0µm increased the rates 5fold compared to other microcarriers.

Induced mutagenesis is a phenomenon intensively studied in *E. coli* and also observed for eukaryotes. This study, however, is the first example of an inducible mutagenesis mechanism in an organellar genome. Elucidating the trigger to induce mutagenesis in a specific target gene might contribute to engineering the chloroplast genome. Furthermore the implications of the organelle actively increasing mutation rates, and therewith rates of genome evolution, on our understanding of plastid evolution might not be conceivable to date

Key words: 16S rRNA, chloroplast, diatom, homology, induced mutagenesis, particle gun, *Phaeodactylum tricornutum*, plastome, *psbA*, transformation

II.2.2 Introduction

Mutations conventionally are considered being formed spontaneously and independent of interaction with the environment (Luria and Delbruck 1943; Newcombe 1949; Lederberg and Lederberg 1952) as an inevitable consequence of imperfect DNA replication and repair. Together with chromosomal rearrangements those spontaneous mutations are important cellular processes that lead to alteration of the genome structure and act as engines driving evolution. Furthermore studies on the fruitfly *Drosophila melanogaster* published as early as 1927 initially demonstrated that mutations in living organisms can be caused by external agents as well (Muller 1927). More recently the molecular mechanisms of spontaneous mutagenesis have been elucidated by genetic analyses of mutator and antimutator genes of Escherichia coli and by biochemical studies of their gene products (Miller 1996). It has been also clearly demonstrated that the underlying mechanisms are evolutionarily well conserved among various organisms (Reenan and Kolodner 1992; Sakumi et al. 1993; Slupska et al. 1996; van der Kemp et al. 1996; Radicella et al. 1997).

Every mutation is derived from premutagenic damage of DNA. Subsequently misreplication of the damaged DNA by either the normal replicative apparatus or by involving a special type of DNA polymerases participating in DNA repair results in a mutagenic intermediate (Friedberg et al. 1995). The mutagenic intermediate, termed a premutation, is finally converted to a mutation at the next round of DNA replication. A major source of premutagenic damages are different errors occurring during normal DNA replication (Maki 2002): (i) Misinsertions of nucleotides can initiate the formation of a mispair. (ii) Misalignment of the growing chain with the template sequence may occur via simple slippage of the terminus, thus promoting the formation of a single-base bulge that can cause a single-base frameshift (10- 46). (iii) Template switching, during the replication potentially generates multiple-base mismatch. This replication error potentially leads to sequence substitution, a particular class of spontaneous mutation in which a segment of DNA ranging from 2 to about 20 nucleotides is replaced by a completely different sequence (Mo et al. 1991; Yoshiyama et al. 2001). However, DNA damage can be also caused by abiotic factors including UV-radiation leading to thymin dimer formations (Friedberg 2002) or other ionizing radiation, but also mutagenic chemicals can seriously affect the integrity of the DNA molecule (Friedberg et al. 1995). An endogenous source of mutagenic base lesions, which is ubiquitous in aerobic organisms, is the exposure to reactive oxygen species (ROS). Finally spontaneous decomposition of DNA bases can occur.

In contrast to spontaneously occurring mutations the discovery of certain, apparently induced type(s) of mutagenesis was rather surprising. Intensively studied mechanisms in *E. coli* were reported to act as a response to stress-inducing, growth-limiting environments by occasionally generating "adaptive mutations" that allow bacterial growth despite growth limiting conditions. (Friedberg et al. 1995; Tang et al. 1999; Maliszewska-Tkaczyk et al. 2000; Rosenberg 2001; Friedberg et al. 2002; Yeiser et al. 2002; Foster 2005). *E. coli* adaptive mutagenesis requires specific components including the specialised error-prone DNA polymerase Pol IV (DinB) (McKenzie et al. 2000; McKenzie et al. 2001) and proteins for homologous recombination and double strand break repair (DSBR) (Cairns and Foster 1991; Harris et al. 1994; Foster et al. 1996; Harris et al. 1996; McKenzie et al. 2000). Point mutations are generated as a consequence of so called "error-prone DSBR" which involves switching from high-fidelity DNA replication to replication via the error-prone Pol IV. Other mechanisms generating mutations in both, prokaryotes and eukaryotes involve induced derepression of Pol IV (DinB) and Pol V (UmuCD) or their homologues during DNA double strand-break repair (DSBR) or translesion DNA synthesis (reviewed by (Friedberg et al. 2005)). These specialised polymerases are Y family polymerases, they display low fidelity and lack the exonucleolytic proofreading function (Kunkel et al. 2003). Controlled switching from the accurate replicative polymerase to error-prone Y family polymerases initiates incorporation of mismatches into the synthesized DNA strand. Switching to error-prone polymerases during DNA repair under stressful conditions has been reported to induce mutagenesis in bacteria and yeast until cells adapt to the stressful environment (Foster 1999; Rosenberg 2001; Hersh et al. 2004; Tenaillon et al. 2004). For error prone DSBR the switching is induced by the transcriptional-activator protein (sigma) RpoS (He et al. 2006) which is expressed under various environmental stress conditions (Hengge-Aronis 2002).

Diatoms are unicellular microalgae which inhabit the world's oceans and fresh water habitats (Mann 1999). Due to their enormous abundance diatoms are of global ecological relevance and greatly contribute to the world's primary production (Falkowski et al. 1998; Field et al. 1998). As more genetic tools for diatoms are required which allow us to study these unique organism, a plastid transformation system would be highly desirable. Therefore we performed a variety of experiments with the aim to achieve stable plastid transformation in the pennate diatom *Phaeodactylum tricornutum*. Instead of a stable insertion of the applied marker into the chloroplast genome we only obtained transient expression so far. However, as a rather surprising outcome of our experiments we obtained some putative transformants which experienced mutagenesis in the target area instead of inserting the marker. This phenomenon occurred exclusively under certain experimental parameters. The aim of this study was to verify an intrinsic mutagenic mechanism and to uncover the parameters triggering mutagenesis. We performed

several experiments designed to induce potentially targeted mutagenesis in the plastid encoded genes *psbA* and the 16SrRNA gene of the diatom. We selected *psbA* and the 16S rRNA gene as convenient target genes for mutagenesis approaches since mutations in these genes are known to increase resistance to either herbicides or antibiotics (Montandon et al. 1985; Ajlani et al. 1989; Ohad and Hirschberg 1992; Smeda et al. 1993; Dalla Chiesa et al. 1997; GuhaMajumdar and Sears 2005) thus allowing a straightforward selection of resistant mutants.

II.2.3 Material and Methods

Strains and media

The *Phaeodactylum tricornutum* strain used for all experiments was the wildtype strain 646 (Bohlin), University of Texas Culture Collection. P. tricornutum was grown at 22°C with continuous illumination at 35 µmol photons m^{-2} s^{-1} in Provasoli's enriched seawater (Starr and Zeikus 1993) using ''Tropic Marin'' artificial seawater at 50% concentration, compared to natural seawater. Solid media contained 1.2% Bacto Agar (Difco). Escherichia coli strain XL-1 Blue (Stratagene, Cedar Creek, TX, USA) was grown over night at 37°C in Luria Broth medium, using a shaker for liquid cultures. Solid media contained 1.5% Bacto Agar (Difco).

Prediction programs

For chloroplast transit peptide prediction, the programs ChloroP (Emanuelsson et al. 1999), TargetP (Emanuelsson et al. 2000) were used. The program SignalP (Nielsen et al. 1997, Nielsen et al. 1999) was used for predicting the thylakoid signal peptide. The BLAST algorithm (Altschul et al. 1997) was used for sequence homology searches.

PCR and construction of plasmids

Standard cloning procedures were used (Sambrook et al. 1989). PCR was performed with a Master Cycler Gradient (Eppendorf, Hamburg, Germany) using recombinant Pfu polymerase (Fermentas, Ontario, Canada) for amplification of 0.8kb fragments or Triple Master Polymerase (Eppendorf, Hamburg, Germany) for the amplification of the 3.5kb fragments according to the manufacturer's instructions.

For the construction of transformation vectors containing 0.8kb of *psbA* and 3.5kb of sequence including *psbA* the inserts were amplified using the primers D1 196-5' and D1 990-3' for the 0.8kb insert or *psbA* up-5' and *psbA* down-3' for the 3.5kb insert respectively. The 0.8kb amplicons were cloned into pGEM-T (Promega, Madison, WI , USA) and pCR-Script Amp according to the manufacturer's instructions. The 3.5 kb inserts were cloned into pGEM-T and pCR-Script Amp (Invitrogen, Carlsbad, CA, USA) via blunt-end ligation after digestion of the plasmids with EcoRV. These cloning steps revealed the vectors pGEM-T D1 0.8, pCR-Script D1 0.8, pGEM-T D1 3.5 and pCR-Script 3.5. A series of plastid transformation vectors was designed with the intention to alter the *psbA* gene (encoding for the PS II protein D1) in a way known to induce resistance to triacine and urea herbicides (Ajlani et al. 1989; Ohad and Hirschberg 1992; Smeda et al. 1993; Dalla Chiesa et al. 1997; GuhaMajumdar and Sears 2005) (table 1). Therefore six

different point mutations were introduced in pGEM-T D1 0.8 while giving raise to the vectors D1-1 to D1-6 (table 1). Only the point mutation for D1-1 ($S_{264}A$, table 1) was also introduced into pGEM-T D1 3.5 thus generating pGEM-T D1 3.5 $S_{264}A$.

vector	aminoacid position	amino acid exchange*	conferring resistance to**	resistance strength (R_{MUT}/R_{WT})***	photosynthetic activity [%]	Literature
D1-1	264	S - A	A./D.	60-100 fold	100	Ohad N. & Hirschberg J. The Plant Cell 1992
D1-2	264	S - G	A.	2000 fold	100	Ohad N. & Hirschberg J. The Plant Cell 1992
D1-3	264	S - P	A.	10000 fold	60	Dalla Chiesa K. et al. Eur. J. Biochem. 1997
D1-4	264	S - T	D.	250 fold	110	Smeda R. J., *et al*. Plant Physiology 1993
D1-5	801	N - Y	A.	330 fold	67	Narusaka Y., *et al*. Plant Cell Physiol. 1998; Sajjaphan K., *et al*. App. and Env. Microbiol. 2002
D1-6	264 + 801	S–T / N-Y	D. + A.	see Mut. 4 and 5	see Mut. 4 and 5	see Mut. 4 and 5

Table 1: Nucleotide substitutions conferred by D1-transformation vectors and the resulting amino acid exchanges. Six different vectors were constructed harbouring 0.8kb of psbA sequence. Point mutations were inserted within the psbA sequence which were reported to confer herbicide resistance in other organisms with impairing photoautotrophic growth. *: amino acids are shown in the single letter code. **: A. = atrazine, D. = diuron(DCMU). ***: resistance strength is calculated by division of the herbicide tolerance of the mutant by the tolerance of the wildtype.

In all seven vectors except DI-6 a second conservative point mutation was introduced, not altering the amino acid sequence of D1 (PsbA) but deleting a BssSI restriction site. Introduction of the respective point mutations into the plasmid DNA was performed via site directed mutagenesis. Therefore the pGEM-T DI 0.8 and pGEM-T D1 3.5 vectors served as template in round circle PCR amplifications using the D1(A) primer pairs (table 2). For the PCR reaction Turbo-Pfu Polymerase (Stratagene, La Jolla, California, USA) was used. Subsequent DpnI digestion of the PCR product degraded the unmutated methylated plasmid template molecules.

To construct the transformation vectors containing a fragment of the 16SrRNA gene, 0.8kb of the 16S rRNA gene were amplified using the primers P-16S_rRNA-5' and P-16S_rRNA-3' and cloned into pGEM-T, thus giving rise to pGEM-T 16S.

Primer	**sequence 5' - 3'**
D1 196-5'	CCTGTTGCAGGTTCTTTATTATATGG
D1 990-3'	TACTTCCATACCTAAATCAGCACGG
psbA up-5'	TTTATGCACATTTTTACAACAGTTACC
psbA down-3'	TTTGCTAAGAGTCTAGCACAAGCTCG
D1(A) 792 S-A forward	CGTTTAATCTTCCAATACGCT**g**CATTTAACAACTC**a**CGTGC
D1(A) 792 S-A reverse	GCACG**t**GAGTTGTTAAATG**c**AGCGTATTGGAAGATTAAACG
D1(A) 792 S-G forward	CGTTTAATCTTCCAATACGCT**gg**ATTTAACAACTC**a**CGTGC

D1(A) 792 S-G reverse	GCACG**t**GAGTTGTTAAAT**cc**AGCGTATTGGAAGATTAAACG
D1(A) 792 S-P forward	CGTTTAATCTTCCAATACGCT**c**CATTTAACAACTC**a**CGTGC
D1(A) 792 S-P reverse	GCACG**t**GAGTTGTTAAATG**g**AGCGTATTGGAAGATTAAACG
D1(A) 792 S-T forward	CGTTTAATCTTCCAATACGCT**a**C**t**TTTAACAACTC**a**CGTGC
D1(A) 792 S-T reverse	GCACG**t**GAGTTGTTAAA**a**G**t**AGCGTATTGGAAGATTAAACG
D1(A) 801 N-Y forward	CGCTTCATTTAAC**t**ACTC**a**CGTGCTTTACACTTCTTCTTAGC
D1(A) 801 N-Y reverse	GCTAAGAAGAAGTGTAAAGCACG**t**GAGT**a**GTTAAATGAAGCG
D1(A) 792+801 S-T+N-Y forward	CGTTTAATCTTCCAATACGCT**a**C**t**TTTAAC**t**ACTC**a**CGTGCTTTAC
D1(A) 792+801 S-T+N-Y reverse	GTAAAGCACG**t**GAGT**a**GTTAAA**a**G**t**AGCGTATTGGAAGATTAAACG
P-16S rRNA-5'	TTTGAGAGGACGATCAGACACACTGG
P-16S rRNA-3'	ACTTAACCCAACATCTCACGACACG

Table 2: Primers used for the construction of transformation vectors. The D1(A) primers were used to introduce point mutations into the vectors containing 0.8kb or 3.5kb of psbA sequence. Small letters in primer sequences indicate the conferred nucleotide substitutions.

Biolistic transformation

Cells were bombarded using the Bio-Rad Biolistic PDS-1000/He Particle Delivery System (Bio-Rad Laboratories, Hercules, Canada) fitted with 1350 psi rupture discs. Tungsten particles (1.1 mm, 0.7 µm or 0.38 µm median diameter) or gold particles (1.0 µm median diameter) were coated with 5 µg of plasmid DNA in the presence of $CaCl_2$ and spermidine, as described by the manufacturer. One hour prior to bombardment approximately 10^8 cells were spread in the centre of a plate containing 20 ml of solid medium. The plate was positioned at the second level within the Biolistic chamber for bombardment. Bombarded cells were allowed to recover for 24 h before being suspended in 1 ml of sterile 10% or 50% artificial seawater medium. 250 µl of this suspension were plated onto solid 10% artificial seawater medium containing 200 µg/ml streptomycin or onto solid 50% artificial seawater medium containing $5 \cdot 10^{-6}$ M DCMU. The plates were incubated at 20°C under constant illumination (35 µmol photons$\cdot m^{-2} \cdot s^{-1}$) for four weeks.

Preparation of nucleic acids and sequencing of the mutated target genes

Resistant mutant colonies were scratched from solid media plates and resuspended in 100 µl 10% or 50% seawater medium. After centrifugation the supernatant was removed and 25 µl of CTAB buffer containing 1% β-Mercaptoethanol (Doyle and Doyle 1990) was added followed by crushing the cells with pistils fitting into the 1.5 ml reaction tubes. CTAB buffer containing 1% β-Mercaptoethanol was added to a final volume of 150 µl. The suspension was incubated at 65°C for one hour. After cooling down to room temperature one volume of Chloroform / Isoamylalkohol (at the ratio 24:1) was added to the suspension and mixed by inverting the reaction tubes gently. The solution was subsequently centrifuged with 16000 g for ten minutes at room temperature. The aqueous phase was mixed with one volume of Isopropanol. The DNA was allowed to precipitate at room temperature for up to two hours followed by centrifugation with

16000 g for 10 min at room temperature the DNA pellet was washed with 80% Ethanol. After drying the DNA was resuspended in 30 µl Tris-HCl 10 mM, pH 7.5. The DNA preparation was used as template for amplification of *psbA* or the 16S rRNA gene. The resulting amplicons were sequenced using the primers D1 196-5' and D1 990-3' and P-16S rRNA-5' and P-16S rRNA-3', respectively (GATC, Konstanz, Germany).

II.2.4 Results

Background

Initial attempts to transform the chloroplast genome of *Phaeodactylum tricornutum* aimed on stable insertion of a marker gene flanked by homologous linkers via homologous recombination. In order to optimize the linker to marker length-ratio a set of transformation vectors was designed to insert two point mutations within the *psbA* gene, thus inducing herbicide resistance. In contrast to "classic" transformation vectors in this case the insert does not consist of a sequence stretch homologous to the target sequence separated by the marker gene, but rather almost exclusively of homologous sequence, only slightly altered by two point mutations. A set of vectors was designed harbouring 0.8 kb of the *psbA* gene (Fig. 2A). The 0.8 kb insert carried various pairs of point mutations. One of the two mutations was conservative, only silently changing the restriction pattern by deleting a BssSI site in all vectors. The second modification in the *psbA* sequence comprised different alternate point mutations (table 1) which have been reported to induce herbicide resistance in other organisms without severely affecting photoautotrophic growth (Ajlani et al. 1989; Ohad and Hirschberg 1992; Smeda et al. 1993; Dalla Chiesa et al. 1997; GuhaMajumdar and Sears 2005). The resulting vectors D1-1 to D1-6 were precipitated onto microcarriers consisting either of gold (0.1µm diameter) or wolfram (0.7µm) prior to transformation of the cells via particle gun bombardment. The transformation procedure and subsequent selection on $5 \cdot 10^{-6}$M of the urea class herbicide DCMU (3-(3,4-dichlorophenyl)-1,1-dimethylurea) generated resistant colony forming units (cfu's) with rates of 10^{-9} to 10^{-8} (cfu's per cells utilized for the transformation procedure). All obtained resistant cfu's were screened for positive transformants. The correct insertion of the appropriate point mutations was controlled via BssSI restriction analysis of *psbA* amplicons and by sequencing of the target area. The BssSI restriction assay revealed that the BssSI site was not deleted in any of the obtained cfu's. Consequentially, sequencing of amplified DNA fragments revealed that all cfu's contained point mutations however in all cases different from those conferred by the transformation vectors. Instead of inserting specific point mutations into *psbA* via homologous recombination we exclusively obtained apparently unspecific but DCMU resistance conferring point mutations. To exclude that mutations occur either spontaneously or due to DCMU treatment negative controls were performed. These controls - performed by bombarding cells with pure microcarriers lacking vector DNA - did not reveal any resistant cfu's.

Transformation or induced mutations in D1 (PsbA)?

Suspecting to induce mutagenesis instead of achieving homologous recombination we performed several experiments in order to investigate on the potential of the diatom's chloroplast to increase mutation rates rather than integrating the inserts into the plastid genome. Therefore we tested whether integration into the plastid genome occurs at all, or if mutagenesis is induced generally due to the transformation procedure. To guarantee that the utilized transformation vectors possess the potential to confer herbicide resistance in case recombination occurs, a specific point mutation was inserted into the homologous insert that was found to increase drastically herbicide tolerance in some of the already obtained mutants. This substitution of the serine at PsbA amino acid position 264 to an alanine ($S_{264}A$) increased DCMU resistance by a factor of ~3000 compared to wildtype without inhibiting mutants' growth under selective conditions (own unpublished data). The applied transformation vectors contained either 0.8kb of sequence within *psbA* or a 3.5kb sequence fragment including the *psbA* gene. The inserts carried the $S_{264}A$ point mutation, approved to confer effectively DCMU resistance and the conservative mutation deleting a BssSI site, mentioned above (see also Fig. 2). In order to assure that the vector backbone itself is not inhibiting successful recombination of the inserts into the plastome two different vectors were used (pGEM-T and pCR-Script Amp), each carrying both inserts respectively. The four resulting transformation vectors were used for a series of parallel transformation experiments under different conditions listed in table 3: (i) Different microcarrier types varying in size and substance (tungsten particles with a median diameter of 0.38µm or 0.7µm, 1.1µm; gold particles with a median diameter of 1.0 µm) have been tested. (ii) The recovery time after transformation and prior to selection was extended up to two days (iii) and selection was performed on DCMU concentrations ranging from $5 \cdot 10^{-6}$M to 10^{-6}M. (iv) Since mutants with altered photosynthetic genes such as *psbA* might be more sensitive to increased light intensities three different low to normal light intensities were applied during selection ranging from 5 to 45 µmol photons$\cdot m^{-2} \cdot s^{-1}$.

particle (substance)	diameter [µm]
W (tungsten)	1.10
	0.70
	0.38
Au (gold)	1.0

insert	length [kbp]
psbA	0.8kb
psbA	3.5kb

vector backbones	pGEM-T pCR-Script Amp
CDCMU [mol/l]	10x6 – 5 - 10-6

recovery time [d]	1
	2

light intensities	[µmol photons m-2 s-1]
low	5
low	20
normal	45

Table 3: Parameters tested during attempts to achieve insertion instead of inducing mutagenesis**.** Transformation vectors were used harbouring psbA sequence containing the $S_{264}A$ point mutation which was approved to cause DCMU resistance in Phaeodactylum tricornutum.

All experiments yielded resistant cfu's with

average rates of $2 \cdot 10^{-9}$ when tungsten microcarriers were used. The diameter of the tungsten particles did not influence the amount of resistant cfu´s. However, utilizing the 1.0 µm gold particles led to a five-fold increase of the rates to 10^{-8}. All other tested transformation parameters such as applying different vector backbones, different recovery times prior to selection or altering the selection pressure had no effect on the amount of resistant cfu's generated by the procedure. Also selecting cells on low to normal light intensities ranging from 20 and 45 µmol photons$\cdot m^{-2} \cdot s^{-1}$ led to cfu formation at equal rates, only very low light intensities of 5 µmol photons$\cdot m^{-2} \cdot s^{-1}$ did not yield any growth.

Neither of the applied conditions caused recombination of the inserts into the target sequence. BssSI restriction assays showed that the obtained cfu's had failed to introduce the restriction site deletion into *psbA*. Accordingly sequencing of the *psbA* genes of randomly picked resistant cfu's revealed for all experiments that DCMU resistance was due to mutations different to those that should have been introduced.

What triggers mutagenesis?

In bacteria and yeast stress by selective pressure is known to induce adaptive mutations in subpopulations of growth repressed cells. Mutations leading to an adaptation that allows growth again are therefore accumulating among the progeny (Foster 1999; Rosenberg 2001; Hersh et al. 2004; Matic et al. 2004). So far resistant mutants of *Phaeodactylum tricornutum* never occurred during negative controls. To ensure that neither naturally occurring mutations nor mutagenesis induced by herbicidal stress alone are sufficient to generate mutations an extensive set of control experiments were performed. In 28 independent transformation experiments cells were exposed to particle gun bombardment using: (i) bare particles of gold (0.1µm diameter) or tungsten (1.1 µm diameter) without applying transformation vector, (ii) transformation vectors which harbour either 0.8 kb or 3.5 kb of unmutated wildtype sequence containing *psbA* precipitated onto either gold (0.1 µm diameter) or tungsten (1.1 µm diameter) particles, (iii) transformation vectors which harbor either 0.8 kb or 3.5 kb of *psbA* containing sequence that carries the $S_{264}A$ point mutations, precipitated again onto both particle types (see (ii)), (iv) and vector DNA (pBS KS-II) without homologous insert precipitated again onto both particle types. (v) Further cells equal to the amount used for eight transformations were exposed to recovery and subsequent selection only, bypassing the particle gun bombardment itself.

Selection always occurred on plates containing $5 \cdot 10^{-6}$ M DCMU. Resistant cfu's were counted after four weeks of selection. The results of this comprehensive set of experiments are shown in table 4. Utilizing transformation vectors which harbour 0.8kb or 3.5kb of homologous sequence containing the $S_{264}A$ point mutation generated

resististant cfu's at rates of ~10^{-8} as expected from in previous experiments. Remarkably the application of the same vectors containing only wildtype sequence without any point mutations in *psbA* revealed the same results while transformations with pBS KS^-II without insert did not yield any resistant cfu's. Particle gun bombardment with bare microcarrier particles as well as exposing cells to selection pressure only - obviously insufficient to induce mutagenesis - did not lead to growth of cfu's. Sequencing *psbA* genes amplified from randomly picked mutants revealed again various point mutations in *psbA*, however different from the $S_{264}A$ mutation which was supposed to be conferred by the used vectors. These results indicate that transforming the cells with vectors containing plastome homologous sequences may induce mutagenesis (figure 2B). Whether the homologous inserts contained or not a pair of point mutations had obviously no effect on mutagenesis induction. Furthermore the experiments suggest that it is crucial to utilize gold particles with a diameter of 1.0 µm. Using tungsten particles of different diameters led only in a few cases to resistant mutants. No mutants occurred when the largest tungsten particles with a diameter of 1.1 µm were applied (see table 4).

microcarrier / vector	gold (0.1µm)	tungsten (0.7µm)	tungsten (1.1µm)
DI 3.5 WT	*0.6 · 10^{-8}*	*not tested*	-
DI 3.5 $_{S264A}$	*0.5 · 10^{-8}*	*~ 2 · 10^{-9}*	-
DI 0.8 WT	*0.5 · 10^{-8}*	*not tested*	-
DI 0.8 $_{S264A}$	*~ 1 · 10^{-8}*	*~ 1 · 10^{-9}*	-
pBS KS^- II	-	*not testet*	-
-	-	-	-
selection only	-		

Table 4: Induced D1 (PsbA) mutagenesis and negative control experiments. Transforming cells with plasmids containing 0.8 or 3.5 of either wildtype *psbA* sequence or *psbA* sequence equipped with the $S_{264}A$ and the conservative $BssSI^-$ point mutation induced generation of resistant mutants with rates of ~10^{-8} mutants/cell. Rates were lower when 0.7 µm tungsten particles instead of 1.0 µm gold particles were used as microcarriers. Usage of 1.1 µm tungsten particles was insufficient to induce mutagenesis. Neither bombardment of cells with bare microcarriers of different substance and size nor plating of untransformed cells on selective media did yield any resistant colonies. Transformation with an empty pBS KS-II vector was also insufficient to generate resistant colonies.

The importance of the material of the microcarriers on mutagenesis remains unclear. However we hypothesize that the gold particles cause less cell damage during vector delivery to the organelle due to their smoother shape and a lower tendency to form large aggregates. Further tungtsten particles were found intruded into the cells after bombardment (Fig.1).

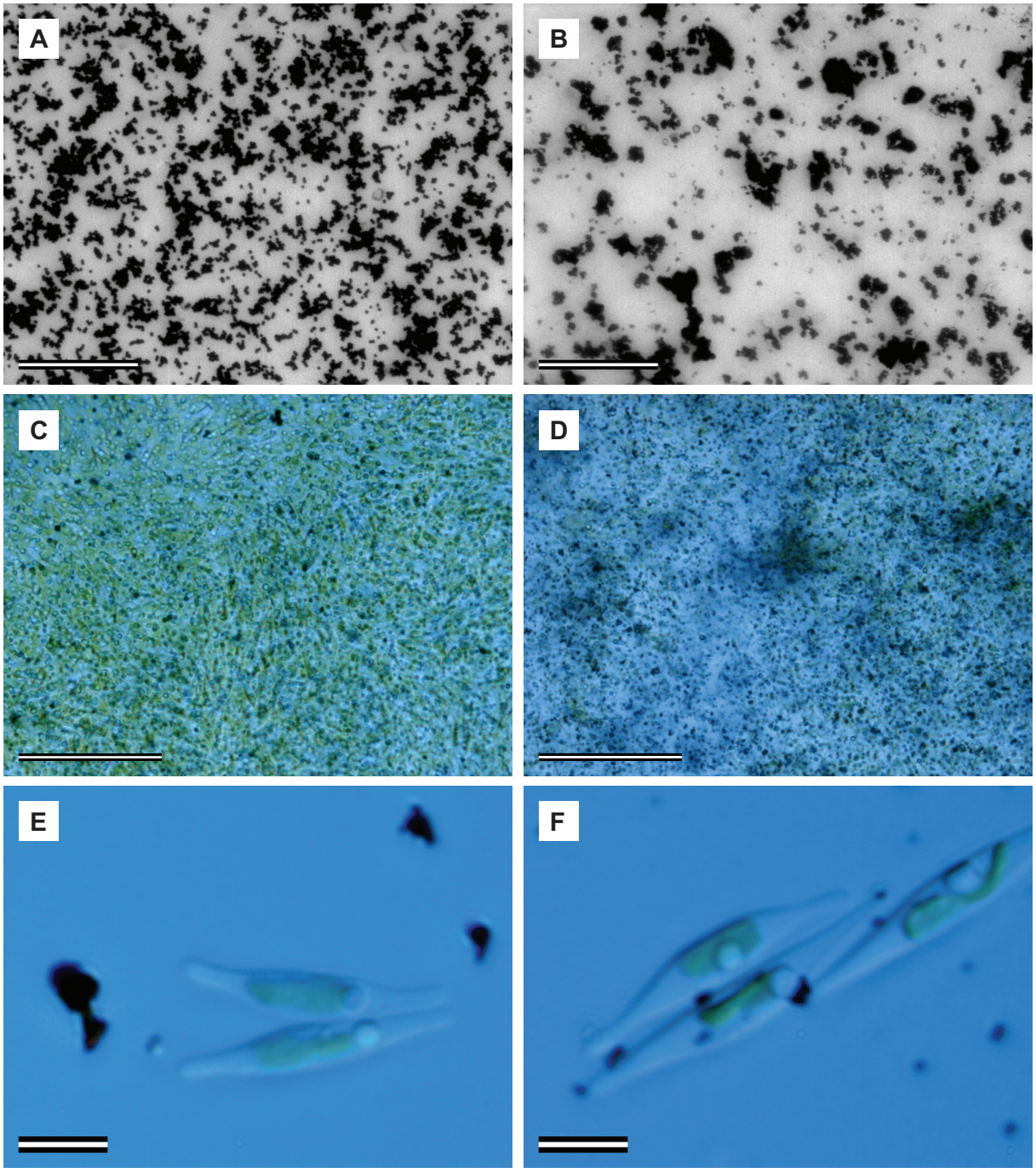


Figure 1: Particles for biolistic transformation. Microparticles dried onto the macrocarrier: **A** 1.0 µm gold particles, **B** 0.7 µm tungsten particles. Surface of an agar plate with *P. tricornutum* cells after bombardment with **C** 1.0 µm gold or **D** 0.7 µm tungsten particles. Scalebar in **A**, **B**, **C**, **D**: 50 µm. *P. tricornutum* cells after bombardment with **E** 1.0 µm gold particles or **F** 0.7 µm tungsten particles. Scalebar in **E** and **F**: 5 µm.

Are mutations inducible in other chloroplast encode genes?

In order to test whether mutagenesis induction is restricted to *psbA* or rather a general feature in *Phaeodactylum*'s plastids we performed a similar approach with the plastid encoded 16S rRNA gene. The 16S rRNA is a target for streptomycin or spectinomycin, aminoglycoside antibiotics which impair the protein biosynthesis machinery of the organism. Mutations in prominent sites altering the secondary structure of the rRNA were previously shown to induce streptomycin resistance in various bacteria as well as in chloroplasts of Chlamydomonas and Euglena gracilis (Montandon et al. 1985; Melancon et al. 1988; Springer et al. 2001; GuhaMajumdar and Sears 2005). In order to induce mutagenesis within the 16S rDNA, a transformation vector was designed analogous to the vectors which induced the mutations in *psbA*. Again the vector contained 0.8kb of wildtype sequence homologous to the 16srRNA gene (figure 2A). In four independent transformation experiments the vectors were delivered to the cells via particle gun bombardment using again 1.0 µm gold particles or 0.7 µm tungsten particles. Negative controls were performed by applying bare microcarriers without vector DNA and by transforming cells with pBS KS$^-$II vectors lacking a homologous insert.

After four weeks of selection on 10% seawater plates containing 200mg/ml streptomycin resistant cfu's were formed. In contrast to selection with DCMU, where no cfu's appeared, negative controls on streptomycin, however, yielded cfu's at low rates (~10^{-9}), probably due to *Phaeodactylum tricornutum*'s high natural resistance towards this class of antibiotics (Apt et al. 1996) (see also chapter II.1). Nevertheless, transforming the cells with the vectors described above generated resistant cfu's with rates of $2.4 \cdot 10^{-8}$ (table 5), therewith rates of obtained resistant cfu's per transformed cell being at average 4.8 fold higher than for the negative controls ($1\text{-}5 \cdot 10^{-9}$). Again mutagenesis was only induced when 0.1 µm gold particles served as microcarriers while usage of 0.7µm tungsten particles never led to an increased amount of resistant cfu's (table 5). Sequencing the 16SrRNA gene amplified from resistant cfu's revealed point mutations within the 16S rRNA fragment in seven out of ten sequenced amplicons. The mutations, both +1 nucleotide insertions and nucleotide substitutions appeared at the nucleotide positions 436, 609, in the sequence region 741-749, and at nucleotides 781, 837, 838, 941, 943 and 1009 (see table 8). The successful approach to induce mutation in the 16S rRNA gene suggests

microcarrier / vector	Gold (1.0µm)	Tungsten (0.7µm)
16SrRNA	***2,4 10^{-8}***	*~1-5 10^{-9}*
pBS KS$^-$ II	*~1-5 10^{-9}*	*~1-5 10^{-9}*
-	*~1-5 10^{-9}*	*~1-5 10^{-9}*

Table 5: Induced mutagenesis of the 16S rRNA gene. Particle gun bombardment with plasmids harbouring 0.8kb of wildtype 16S rDNA sequence yielded rates of $2.4 \cdot 10^{-8}$ resistant colonies/cell. Again rates were 5 times higher when 1µm gold microcarriers where used. Bombardment of cells with bare particles or an empty vector, yielded few resistant colonies since *Phaeodactylum tricornutum* displays high natural resistance against the selective agent, streptomycin. However rates were up to 24 fold higher in transformation experiments that in negative controls.

that mutagenesis can be induced generally in all plastome encoded genes.

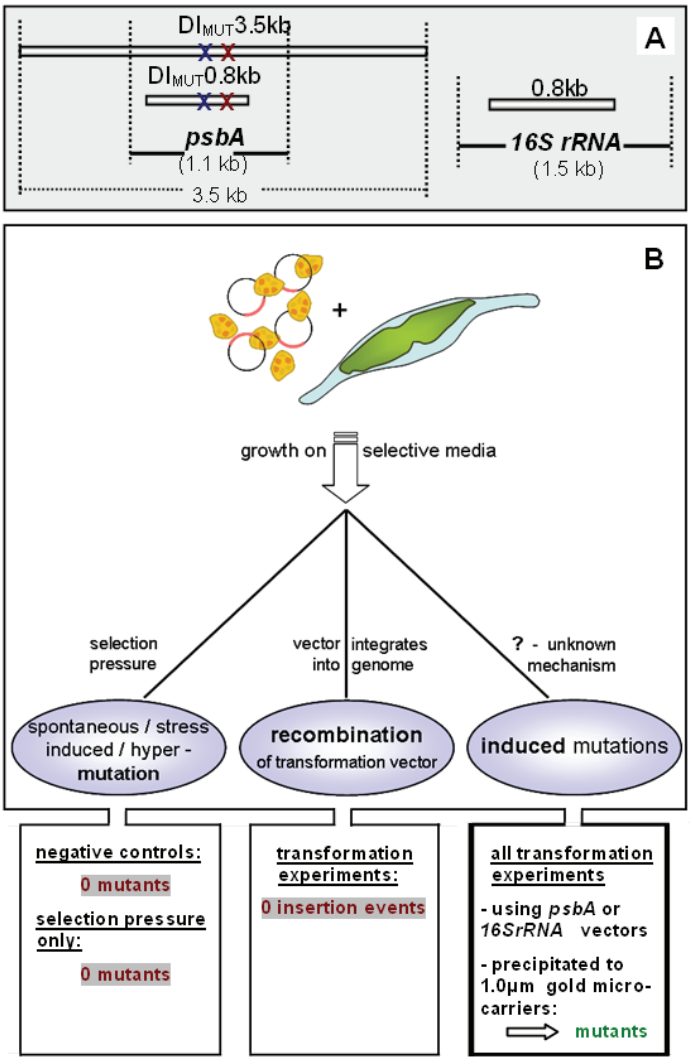


Figure 2: Triggering mutagenesis in a chloroplast genome encoded target gene. **(A)** shows transformation vector inserts (grey bars) homologous to the sequence containing the target genes *psbA* and the 16S rDNA. The inserts were either in wildtype sequence or carried two point mutations (blue and red **X**, also see materials and methods), although not required to induce mutagenesis. **(B)** Plastome mutagenesis was induced by particle gun bombardment delivering plasmids which contain 0.8-3.5kb of sequence homologous to the target gene into the cells. The use of small 1.0µm gold microcarriers was crucial for a more effective mutagenesis induction. The resulting mutants cannot be attributed to spontaneous or stress induced mutations allowing growth on selective media, since neither negative control transformations nor presence of selection media alone were sufficient to generate any resistant mutants. It was further excluded that resistance was induced by correct integration of potentially resistance conferring transformation vectors – homologous recombination of the plasmids into the plastome was never observed under any conditions.

Is induced mutagenesis a gene directed or plastome wide effect?

In order to enlighten whether mutation rates during induced mutagenesis are generally increased in the plastome or only within the targeted sequences homologous to the transformation vector inserts, further sequencing experiments have been performed. Coding and non coding (intergenic) sequences downstream of rbcL and upstream of *ycf*24 were monitored for nucleotide substitutions or other mutations and compared to the wildtype plastome sequences (accession numbers: AY864816.1, AY819643.1). Both monitored regions together span 1859bp of sequence which compares to the lengths of the *psbA* (1083bp) gene or the gene of the 16SrRNA (1482bp). The respective fragments were amplified and sequenced for the wildtype and four randomly picked D1 and 16S rRNA mutants.

Sequencing revealed no nucleotide substitutions within both monitored areas compared to the wildtype sequence. This result might indicate that mutagenesis is not only inducible but also restricted to an area of the plastome with homology to the vector inserts.

Characteristics of induced plastome mutations.

Sequencing a total of 19 D1 mutants reveaed 28 point mutations within the *psbA* gene. Up to five amino acid substitutions were identified in the same mutant. Alltogether, the D1 mutants displayed fourteen different amino acid substitutions (table 6). One particular mutation leading to an amino acid substitution of valine at position 219 to isoleucine was identified in 11 different mutants, thus indicating an apparent hot spot for a certain type of mutation.

Seven out of ten sequenced 16S rRNA mutants displayed different mutation types within the 16S rRNA gene: six +1 nucleotide insertion mutations and 5 nucleotide substitutions were identified. All mutations observed in both genes are listed according to the type of nucleotide exchange in table 7.

D1 Mut*	Frequency
$R_{27}C$	*1*
$A_{51}T$	*1*
$G_{69}C$	*1*
$Q_{113}R$	*1*
$S_{134}G$	*1*
$V_{219}I$	*11*
$F_{255}I$	*4*
$A_{263}V$	*1*
$S_{264}A$	*1*
L_{275}stop	*1*
$L_{275}W$	*1*
$W_{278}L$	*2*
$D_{325}Y$	*1*
E_{329}stop	*1*

16SrRNA Mut**	Frequency
+1	*6*
$C_{436}T$	*1*
$A_{837}G$	*1*
$T_{941}C$	*1*
$A_{943}G$	*1*
$T_{1009}C$	*1*

Transitions	Transversions
26	*7*

Table 6 (above): Frequencies of aminoacid substitutions in in D1 (PsbA) mutants and of nucleotidsubstitutions and +1 nucleotide insertions in 16SrRNA mutants respectively. Observed transition- and transversion events.
*: amino acids are shown in the single letter code. Numbers represent aminoacid positions in D1. **: Nucleotide substitutions. Numbers represent nucleotide positions in the 16S rDNA. +1: +1 nucleotide insertion mutation.

Pyrimidines

Transition C/G - T/A	*Transversion* C/G - A/T	*Transversion* C/G - G/C	*Transiton* T/A - C/G	*Transversion* T/A - A/T	*Transversion* T/A - G/C
3	0	0	6	0	2

Purines

Transversion A/T - T/A	Transition A/T - G/C	*Transversion* A/T - C/G	Transition G/C - A/T	*Transversion* G/C - T/A	*Transversion* G/C - C/G
0	5	0	12	5	0

Table 7: Frequencies of different mutational events observed by sequencing of mutant target genes. Detailed list of observed transitions and transversions generated via induced mutagenesis in the plastome of *Phaeodactylum tricornutum*.

experiment	name	nt pos		codon	aa pos	aa substitution*
D1 0.8 kb (var.)	*DI2-2*	400	A-G	**A**GC-**G**GC	134	S-G
	D12-3	763	T-A	**T**TT-**A**TT	255	F-I
	D13-3 †	824	T-G	T**T**A-T**G**A	275	L-stop
	D13-3	824-825	TA-GG	T**TA**-T**GG**	275	L-W
	D15-1	763	T-A	**T**TT-**A**TT	255	F-I
	D14-3	788	C-T	G**C**T-G**T**T	263	A-V
	D14-3	790	T-G	**T**CA-**G**CA	264	S-A
	DI5-14	790	T-G	**T**CA-**G**CA	264	S-A
D1 0.8kb (S264A / WT)	*G11*	655	G-A	**G**TA-**A**TA	219	V-I
	G11	985	G-T	**G**AA-**T**AA	329	E-stop
	G1	151	G-A	**G**CA-**A**CA	51	A-T
	G1	655	G-A	**G**TA-**A**TA	219	V-I
	G2	97	C-T	**C**GT-**T**GT	27	R-C
	G2	205	G-T	**G**GT-**T**GT	69	G-C
	G2	655	G-A		219	V-I
	G2	762	C-T		254	Y-Y
	G2	763	T-A	**T**TT-**A**TT	255	F-I
	G5	655	G-A	**G**TA-**A**TA	219	V-I
	G6	338	A-G	C**A**A-C**G**A	113	Q-R
	G6	763	T-A	**T**TT-**A**TT	255	F-I
	W5	655	G-A	**G**TA-**A**TA	219	V-I
D1 3.5kb (S264A / WT)	*D12-12*	833	G-T	T**G**G-T**T**G	278	W-L
	DI4-14	655	G-A	**G**TA-**A**TA	219	V-I
	DI4-14	973	G-T	**G**AT-**T**AT	325	D-Y
	DI4-5	660	G-A	**G**TA-**A**TA	219	V-I
	DI5-18	656	G-A	**G**TA-**A**TA	219	V-I
	DI5-18	833	G-T	T**G**G-T**T**G	278	W-L
	DI5-210	657	G-A	**G**TA-**A**TA	219	V-I
	DI5-219	658	G-A	**G**TA-**A**TA	219	V-I
16S rRNA	*16S-32*	741	+1 A	-	-	-
		749	+1 T	-	-	-
	16S-313	759	+1 C	-	-	-
		781	+1 A	-	-	-
	16SS-37	609	+1 A	-	-	-
		831	+1 T	-	-	-
	16S-7	436	C-T	-	-	-
	16S-13	837	A-G	-	-	-
		941	T-C	-	-	-
	16S-16	943	A-G	-	-	-
	16S-1	1009	T-C	-	-	-

Table 8: List of mutants obtained from transformation experiments utilizing plasmids harbouring: 0.8kb of psbA sequence containing various point mutations (table 1) (D1 0.8 kb var.); 0.8kb of psbA sequence either as in the wildtype or containing the S264A point mutation (D1 0.8kb S264A / WT); 3.5kb of sequence including psbA either as in the wildtype or containing the S264A point mutation (D1 3.5kb S264A / WT); 0.8kb of the 16Sr RNA gene (16S rRNA). *: amino acids are shown in the single letter code. †: After four month of cultivation this mutant experienced a spontaneous mutation substitution the stop codon by W.

A majority of 78.8% of all identified nucleotide substitutions were transitions, leading to a rate of transitions to transversions of 26:7 (table 6). All obtained mutants are listed in table 8. One particular D1 mutant is certainly worth mentioning. In mutant D1 3-3 leucin at amino acid position 275 was substituted by a stop codon (TGA). Repeating the amplification and sequencing of the *psbA* gene verified this result. It is remarkable that this mutant was still capable of photoautotrophic growth, albeit at low rates, since it must be assumed that D1 was not properly expressed anymore. However after four month of cultivation this mutant experienced a second, apparently spontaneous mutation substituting the stop codon by the triplet TGG encoding tryptophane.

II.2.5 Discussion

We report induced mutagenesis in the chloroplast genome of *Phaeodactylum tricornutum* leading to a selective advantage for the mutated cells. Mutagenesis was triggered by an experimental setup designed to deliver plasmid DNA into the diatom cells via particle gun bombardment. This approved method for nuclear transformation delivers dependably DNA into the cells. Therefore it can be assumed that DNA bound to microcarriers can also access *Phaeodactylum tricornutum*'s large chloroplast via this procedure. Bombarding cells with vectors carrying sequence fragments homologous to *psbA* or the 16SrRNA gene induced in independent attempts mutagenesis in the target genes, yielding an average of 10^{-8} resistant colonies per bombarded cell. Increasing the length of homologous sequence to 3.5kb had no effect on the mutation rates. However, the microcarrier type applied during the transformation procedure was shown to affect the rate of mutant formation. Using gold particles with a diameter of 1µm instead of tungsten particles further elevated rates of generated mutants by a factor of five. Whether the material of the particles has an influence on the generation of mutants remains elusive, however, it appears plausible that the gold particles cause less damage to cell or organelle due to their smoother shape and a significantly lower tendency to form large particle-aggregates compared to tungsten particles. Sequencing *psbA* or the 16SrRNA genes of randomly picked resistant colonies revealed point mutations within these genes. All sequenced *psbA* genes had experienced nucleotide substitutions, some mutants contained up to five amino acid exchanges in D1. The fact that all sequenced *psbA* fragments contained nucleotide substitutions indicates that selection with DCMU is very stringent, thus promoting efficiently the segregation of the mutations through all plastome copies. Repeating amplification and sequencing of the mutants' *psbA* gene supported this assumption and further demonstrated the stability of the D1 mutations. With 70% also the majority of the 16SrRNA sequences displayed either +1 insertion mutations or nucleotide substitutions. The identification of mutations in only 70% but not all 16S rRNA sequences can be explained with *Phaeodactylum tricornutum*'s high natural tolerance to a range of antibiotics (Apt et al. 1996) (see also chapter II.1) including streptomycin. Hence few colonies were also formed in negative controls of the 16S rRNA-mutagenesis experiments independent of mutations in the 16SrRNA. Another explanation is provided by the fact that the 16S rRNA gene is located within the inverted repeats of the plastid genome. The inverted repeats are two identical ~25kb copies of plastome sequence in inverted orientation. Mutation rates in these sequence repeats is significantly lower than for genes located in the single copy regions (Wolfe et al. 1987; Maier et al. 1995). Furthermore there is strong evidence for a copy correction system which reverts

mutations in the inverted repeats via gene conversion (Svab et al. 1990; Birky and Walsh 1992; Morton and Clegg 1993; Khakhlova and Bock 2006). These facts indicate that the mutations in the 16S rRNA gene may never have segregated completely due to correction mechanisms. It is therefore likely that the most mutated gene copies have been reverted and only a small number of 16S rRNA genes, just enough to provide sufficient Streptomycin resistance, kept the mutation. Hence it seems obvious that sequencing of the 16S rRNA genes could not reveal nucleotide substitutions in all amplicons.

Altogether, the results presented in this study provide clear evidence that mutagenesis in the chloroplast genome can be induced. The crucial factor for inducing resistance appears to be the introduction of a DNA fragment homologous to genes encoding for products which are affected by the selective agent. The results of these experiments – including the experimental parameters and preconditions to induce mutagenesis in a target gene – further allow speculations on the underlying mechanism. Whenever mutagenesis occurs in a chloroplast genome the mutagenic effect of reactive oxygen species (ROS) should be certainly considered as a possible reason. Since the photosynthetic machinery in the chloroplast is fairly prone to generating ROS, the resulting mutagenic effect must be considered to contribute to elevated mutation rates in the plastome. The mutagenic effect of ROS is basing on the generation of 8-oxo-7,8-dihydroxyguanine (8-oxo-dGTP) which has ambivalent base pairing properties and is capable of pairing effectively with both A and C during DNA synthesis. Hence A/T → C/G transversions are generated resulting from the misincorporation of 8-oxo-dGTP in the template opposite adenine (Bridges 2001; Dany and Tissier 2001; Fowler et al. 2003). Error-prevention systems in *E. coli* (mutT, mutY and mutM) and their homologues in eukaryotes protect cells against the effects of 8-oxo-dGTP mediated misincorporation by hydrolyzing 8-oxo-dGTP to prevent its use as a substrate by DNA polymerases (Boiteux et al. 2002). Strains defficient in these error prevention systems thus show an increased number of A/T → C/G transversions (Fowler et al. 2003). However, in *Phaeodactylum tricornutum* mutants we mainly observed transitions, the ratio of transitions/transversions being 3.71. Therefore down-regulating the functional homologues of the error-prevention system in the chloroplast can be ruled out as potential mechanism to increase plastome mutation rates.

Other mutation generating mechanisms known in both prokaryotes and eukaryotes involve induced derepression of the specialised Y family member DNA polymerases PolIV (DinB) and PolV (UmuCD) or their homologues during DNA repair (Friedberg et al. 2005). Recently was demonstrated that a defining feature of induced mutagenesis basing on error-prone DSBR is the requirement for sequence homology between the double strand end and the DNA that acquires the mutation (He et al. 2006). The homology might usually be provided by a sister DNA molecule, a duplicated genome segment, or other

sources of homologous DNA such as episome copies (Foster and Trimarchi 1995; Galitski and Roth 1995; Radicella et al. 1995; Foster 2004) or plasmids (Ponder et al. 2005). A crucial role in this homology recognition certainly plays RecA, a widely conserved DNA repair protein which is essential for most yet studied mutagenesis modells. In prokaryotes and eukaryotes, RecA and its homologues conduct homology recognition and participate in homologous recombinational DNA repair, including double strand break repair (DSBR). In bacteria, RecA also induces the SOS DNA damage response (Cox 2003). During mutagenic translesion synthesis in *E. coli* the error-prone DNA polymerase needs to form a stable complex with RecA molecules or filamers and the □-clamp to access the template strand and bypass the lesion, thus potentially generating a mutation (Goodman 2002; Pages and Fuchs 2002; Fuchs et al. 2004; Fujii and Fuchs 2004; Friedberg et al. 2005; Schlacher et al. 2005). Therefore RecA might be the crucial factor for targeting error-prone DNA polymerase, i.e. point mutations to a certain area during induced mutagensis by recognizing homology between the damaged DNA strand and a homologous strand. This assumption is further supported by the finding that error-prone DSBR mutagenesis is specific to interactions in the vicinity of sequences homologous to the double strand break surrounding area (Ponder et al. 2005). Targeting stress induced mutagenesis would allow preserving the integrity of most of the genome in any given cell while mutating small localized regions extensively.

Comparing adaptive mutagenesis via error-prone DSBR with the induced plastome mutagenesis in *Phaeodactylum tricornutum* reveals common features. Induced mutagenesis in the diatom as well as error-prone DSBR in bacteria and yeast lead to the formation of colonies despite growth inhibiting conditions. We also identified a nuclear encoded homologue of the error-prone PolIV (DinB) in *Phaeodactylum tricornutum*. The diatom DinB homologue is furthermore predicted to be imported into the chloroplast (data not shown). However, while mechanisms generating adaptive mutations in bacteria or yeast are stress-induced, it remains unclear whether mutagenesis in *Phaeodactylum tricornutum*'s chloroplast genome is triggered by environmental stress. Stressing the diatom cells with selective conditions alone was not sufficient to yield any resistant colonies, neither was the stress of the particle gun bombardment procedure using bare particles or empty plasmids precipitated to the particles. The most salient similarity between error-prone DSBR and induced mutagenesis in the diatom's plastome is the requirement for sequence homology. While a crucial prerequisite for error-prone DSBR is the homology between the double strand end and the DNA that acquires the mutation, also the induction of plastome mutagenesis depends on the introduction of additional target sequence copies into the organelle (or at least into the organism) via particle gun bombardment. Since introducing plasmids containing homologous sequences into *Phaeodactylum tricornutum* cells was sufficient to cause mutations in the respective genes, the importance of additional target sequence copies present in the cells, e.g. in

the form of episomes (Foster and Trimarchi 1995; Galitski and Roth 1995; Radicella et al. 1995; Foster 2004) or plasmids (Ponder et al. 2005), might be even underestimated by previous studies on error-prone DSBR. RecA which is required for adaptive mutations in E.coli and yeast and which might be essential for targeting the mutations is also encoded in the diatoms' nuclear genome and imported into the chloroplast after translation (see chapter II.4). Monitoring sequences which are rather distant to *psbA* or the 16S rRNA gene in the mutants did not reveal any differences to the wildtype sequence, thus suggesting that induced mutations in *Phaeodactylum tricornutum* 's plastome might be directed to a certain area with homology to the transformation vector. This finding correlates with the study of Ponder et al. (Ponder et al. 2005), who demonstrated that induced mutations during error-prone DSBR in *E. coli* are targeted to certain areas. However more extensive sequence analysis of plastome mutants is required to elucidate whether mutagenesis in the plastome of *Phaeodactylum tricornutum* is not only inducible but also directed. The advantage of potential directed mutagenesis might be that it avoids deleterious coincident mutations by restricting mutagenesis to small genome segments (Ponder et al. 2005).

Different inducible mutagenic mechanisms have been previously identified in prokaryotes and eukaryotes. A related mechanism inducing adaptive mutations in budding yeast (Heidenreich et al. 2003) involves proteins for non-homologous end-joining (NHEJ), thus applying a mechanism fundamentally different from DSBR. In addition to *E. coli* inducible error-prone DSBR also occurs in yeast (McGill et al. 1998) and Salmonella. Therefore, the fact that chloroplasts feature a related or independently evolved mutation inducing mechanism as well it is not unlikely. However, such a mechanism is certainly intriguing, especially in the context of plastome mutation rates being extraordinary low under normal conditions. Chloroplast genes are reported to generally experience less mutations ($1.1\text{-}2.9 \cdot 10^{-9}$ substitutions per synonymous site per year) than nuclear encoded genes ($5.8\text{-}31.5 \cdot 10^{-9}$ substitutions per synonymous site per year) (Wolfe et al. 1987; Muse 2000). Coherently chloroplasts were shown to feature high fidelity of DNA replication as well as highly efficient DNA repair (Maki 2002; GuhaMajumdar and Sears 2005). Therefore the implications of the organelle actively increasing mutation rates, and therewith rates of genome evolution, on our understanding of plastid evolution might not be conceivable to date.

Putative functions of an inducible and targeted mutagenic mechanism in a chloroplast are unclear. However controlled and targeted mutagenesis on the plastomic integrity certainly influences the effect of Muller's ratchet (Muller 1964) on a plastid genome. Muller was the first to realize the fact that an asexual population accumulates deleterious mutations in an irreversible manner (hence the word ratchet) while the genomes of sexual populations can easily reverse this process due to recombination. Chloroplasts are usually maternally inherited although biparentally or even paternally

inheritance exists as well (Birky 1995; Mogensen 1996), however even in those rare cases where inheritance is biparental no recombination between the two parental plastids could be observed (Medgyesy et al. 1985; Thanh and Medgyesy 1989). Therefore the lack of sexual recombination should have negative consequences in that deleterious mutations accumulate over time. Conventionally it was considered that the low mutation rates of organellar genomes counteract the speed of Muller's turning ratchet, while copy correction via biased gene conversion can even turn it back eventually when deleterious mutations appear in the inverted repeats of chloroplast genomes (Khakhlova and Bock 2006). A controlled and targeted mutagenesis system aligned with mechanisms such as biased gene conversion and effective DNA repair might faciliate organellar evolution in the absence of sexual reproduction. When required, mutations were induced and removed via gene conversion if deleterious. In this context the high ploidy of plastids - *Phaeodactylum tricornutum*'s chloroplasts contain ~100 plastome copies (see chapter II.4) – appear important for controlling new mutations.

The fact that the introduction of additional sequence copies into the cells induces mutagenesis while the naturally occurring already high amount of genome copies displays even extraordinary low mutation rates might also hint a system which sensors the gene copy numbers in the chloroplast. A similar mechanism was first discovered in the filamentous fungus Neurospora crassa which efficiently detects and mutates both copies of a sequence duplication in a process termed repeat-induced mutation (Selker et al. 1987; Cambareri et al. 1989; Selker 1990; Galagan et al. 2003).

Gene silencing – a new tool for combining genetics and physiology in diatoms

Materna, A.C.[1], Lavaud, J.[1], Ng Chin Yue, S.[1], and Kroth, P.G.[1]

1 Department of Plant Ecophysiology, University of Konstanz, Germany

II.3.1 Abstract

The purpose of this study was to establish a protocol for gene silencing in the diatom *Phaeodactylum tricornutum*. The diadinoxanthin de-epoxidase (*dde*) was chosen as target gene. Since the *dde* mediated conversion of diadinoxanthin (DD) to diatoxanthin (DT) under excess light conditions is inevitable for the photoprotective NPQ mechanism to develop, the silencing of this gene was expected to induce a clearly identifiable phenotype. RNA interference was induced by transformation of the cells with plasmids containing short (187bp) or long (509bp) antisense fragments under control of the fcpA promoter. A second strategy based on the transformation of *Phaeodactylum tricornutum* with a plasmid mediating the expression of a self- complementary hairpin like construct with a 5'-sense-overhang. The silencing approaches generated transformants with a phenotype clearly distinguishable from wildtype cells. PAM fluorescence measurements revealed that at least 80% of the transformant strains showed a decrease in NPQ by more than 10%. The majority of the examined transformants showed even between 30% to 47% reduction in NPQ compared to wildtype. Real-time PCR based quantification of *dde* transcripts showed differences in *dde* transcript levels between AS strains and wildtype cells but also between AS and RNAi strains, thus suggesting the presence of two different gene silencing mediating mechanisms in diatoms.

Key words: antisense, diadinoxanthin de-epoxidase, diatom, dsRNA, non-photochemical quenching, *Phaeodactylum tricornutum*, RNA interference, violaxanthin de-epoxidase

II.3.2 Introduction

Diatoms belong together with cyanobacteria to the most abundant phytoplanktonic organisms in the world's oceans and are therefore of great ecological relevance. *Phaeodactylum tricornutum* is a convenient model organism for studying marine diatoms. Several aspects of its physiology have been previously investigated including photosynthetic properties such as the xanthophyll cycle and photoprotection towards high light stress (Lohr and Wilhelm 1999; Goss et al. 2006; Lavaud and Kroth 2006) as well as aspects of its carbohydrate metabolism (Kroth et al. 2005; Michels et al. 2005). Other studies on *Phaeodactylum tricornutum* focused on the evolution of secondary chloroplasts and protein import into these complex plastids (Kroth and Strotmann 1999; Kroth 2002; Kilian and Kroth 2004, 2005). Also, its genome has been sequenced recently and annotation is to date nearly completed, which gave way to diatom comparative genomics (Montsant et al. 2005). A technique for efficient transformation of the genome of *Phaeodactylum tricornutum* via particle gun bombardment is available as well (Apt et al. 1996; Zaslavskaia et al. 2000). New molecular tools are now required in order to apply the genomic information to reverse genetics. Therefore we investigated the feasibility of posttranslational gene silencing in *Phaeodactylum tricornutum* by testing techniques known to induce RNA interference. Since `antisense-mediated silencing` has been discovered in 1995 (Guo and Kemphues 1995), double-stranded RNA (dsRNA) was proved to be an extremely potent activator of RNA interference (Fire et al. 1998; Meister and Tuschl 2004). DsRNA can originate from different sources. Endogenous sources include for instance, short forms of fold-back dsRNA (Bartel 2004) which are the precursor molecules of micro RNAs (miRNAs) (reviewed by (Ambros 2004). Another source of endogenous dsRNA is a class of enzymes called RNA-directed RNA polymerases (RDR). Abnormal or 'aberrant' transcripts are recognized by RDR and subsequently converted into endogenous dsRNA by primer-independent synthesis of complementary RNA (Schiebel et al. 1993; Dougherty and Parks 1995; Baulcombe 1996; Makeyev and Bamford 2002). However, dsRNA can also result from transgene transcription. Transgenic transcripts can be designed to either hybridize with complementary target transcripts or to contain complementary inverted repeats folding back on themselves to form dsRNA hairpins (Waterhouse and Helliwell 2003). Both, dsRNA and miRNA precursors are processed by special dicer to small interfering RNA (siRNA) or miRNA respectively (reviewed by (Meister and Tuschl 2004; Mello and Conte 2004). These small RNA products are subsequently rearranged into the RNA-induced silencing complex (RISC) (Hammond et al. 2000) or the miRNA containing effector complex (miRNP) (Mourelatos et al. 2002) which guide distinct protein complexes to the target RNAs. While the RISC mediates the target mRNA degradation, miRNP can also guide translational repression of

target mRNAs (Meister and Tuschl 2004). DsRNA triggers these RNAi processes, hence the allocation of dsRNA in cells is a convenient tool to perform gene silencing.

In order to establish a protocol for gene silencing in *Phaeodactylum tricornutum* we transformed the cells with different plasmids that allow the transcription of either self-complementary fragments of the target mRNA or antisense (AS) mRNA fragments that are reverse complementary to the target mRNA. We chose the diadinoxanthin de-epoxidase gene (*dde*) as target for the gene silencing approaches. Dde is involved in a photoprotective process termed non-photochemical Chl fluorescence quenching (NPQ). Photosynthetic growth in fluctuating light intensities requires a fast responding mechanism to protect the organism from potential damage by excess energy absorption at supersaturating light intensities. Plants and algae have evolved a number of protecting mechanisms including NPQ (Holt et al. 2004; Horton et al. 2005). NPQ mediates thermal dissipation of light energy absorbed in excess by the light-harvesting antenna complex (LHC) of PSII. This dissipation of the excess radiant energy is partly controlled by the interconversions between the carotenoids violaxanthin, antheraxanthin and zeaxanthin during the main xanthophyll cycle (Demmig-Adams and Adams 1996; Demmig-Adams et al. 1996; Eskling et al. 1997; Gilmore 1997). The xanthophyll cycle in diatoms differs from the xanthophyll cycle in plants. It is simpler and involves different pigments - antheraxanthin and zeaxanthin are replaced by diadinoxanthin (DD) and diatoxanthin (DT) (Stransky and Hager 1970). The accumulation of DT was shown to be crucial for NPQ (Arsalane et al. 1994; Olaizola et al. 1994; Olaizola and Yamamoto 1994; Lavaud et al. 2002).

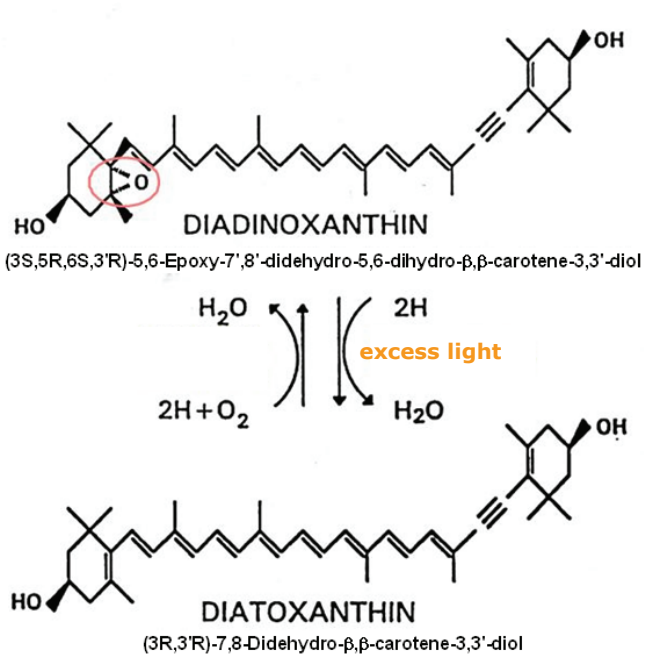


Figure 1: Conversion of Diadinoxanthin into Diatoxanthin under excess light conditions is mandatory for NPQ in diatoms.

The purpose of this study was to test the feasibility of gene silencing in the diatom *Phaeodactylum tricornutum*. Silencing of the *dde* gene is assumed to result in an impaired conversion of DD to DT, thus leading to a reduced NPQ. Since the phenotype NPQ can be measured via PAM imaging fluorescence measurements the *dde* appeared to be a convenient target gene for testing the suitability of gene silencing in diatoms.

II.3.3 Material and Methods

Strains and media

Phaeodactylum tricornutum Bohlin (University of Texas Culture Collection, strain 646) was grown at 22°C with continuous illumination at 35 µmol photons m^{-2} s^{-1} in Provasoli's enriched seawater (Starr and Zeikus 1993) using ''Tropic Marin'' artificial seawater at 50% concentration, compared to natural seawater. Solid media contained 1.2% Bacto Agar (Difco). Escherichia coli strain XL-1 Blue (Stratagene, Cedar Creek, TX, USA) was grown over night at 37°C in Luria Broth medium, using a shaker for liquid cultures (Sambrook et al. 1989). Solid media contained 1.5% Bacto Agar (Difco).

PCR and construction of plasmids

Standard cloning procedures were used (Sambrook et al. 1989). PCR was performed with a Master Cycler Gradient (Eppendorf, Hamburg, Germany) using recombinant Pfu polymerase (Fermentas, St.Leon-Rot, Germany) according to the manufacturer's instructions. Small letters in primer sequences indicate degenerated nucleotides (table 1). All AS fragments have been inserted into the *Phaeodactylum tricornutum* transformation vector pPha-T1 (Apt et al. 1996) which allows transformation of the diatom selecting for positive transformands using zeocin. The 187bp *dde* antisense fragment was amplified using the primers No. 1 and No. 2, the 509bp fragment was amplified using the primers No.3 and No.4. HindIII and BamHI restriction sites were added to both fragments by modifications in the primer sequences. After digesting the fragments and the pPha-T1 plasmid with HindIII and BamHI the inserts were ligated in antisense orientation into the plasmid downstream of the fcpA promoter, thus giving raise to pPha-T1 AS187 and pPha-T1 AS509 (Fig. 2). In order to construct the pPha-T1 RNAi plasmid several subsequent cloning steps were performed. First a 272bp fragment of the *dde* gene was amplified using the primers No.5 and No.4, again adding HindIII and BamHI restriction sites to the 5'- and 3'ends. Again the fragment was cloned in pPha-T1 in antisense orientation giving rise to pPha-T1 AS300. In the following steps a fragment containing the fcpA 5' untranslated region (UTR) downstream of the fcpA promoter followed by eGFP was amplified from a pPha-T1 vector - which contains eGFP - using the primers No. 6 and No.7. Further a 509bp long fragment of *dde* was amplified using the primers No. 8 and No. 9. Both fragments were ligated in the pF1-A directional cloning vector (Promega, Madison, WI, USA) according to the manufacturer's protocol. The resulting vectors were termed pF1-A DDEnc S and pF-1A pTV-MCS. pF1-A DDEnc S was digested with Ecl136II and FspI, pF-1A pTV-MCS was digested with PvuII and FspI. For both plasmids the half containing the respective insert was separated by horizontal gel-

electrophoresis and eluted from the agarose gel. Both halves were ligated, thus giving rise again to complete pF1-A vector which contains the 272bp *dde* fragment and the 5-UTR-eGFP construct, now ligated together. The resulting *dde* (sense)-5'-UTR-eGFP insert was amplified from the plasmid using the primers No.5 and No.7. The amplificon was subsequently digested with BamHI. After digesting the plasmid pPha-T1 AS300 with EcoRV and BamHI the digested amplificon was ligated into the plasmid, thus giving rise

No.	Primer	sequence 5' - 3'
1	DDEnc-187AS-HindIII-5'	GACAAgCttAATCCCTCGTACGGACTTCAGC
2	DDEnc-187AS-BamHI-3'	aaaaggatcccccccccACGGGAGAGTGACTGGTGTCG
3	DDEnc-HindIII-5'	AAAAAGcttCAACCAAAATCCCTCG
4	DDEnc-BamHI-3'	TTAATGgAtcCAACGTTGGCGAGGCAC
5	DDEnc-340-HindIII-5'	TCTCGGGCTCCAAgcTTGTCGCTCCGC
6	SgfI-NcoI-pTV-MCS-5'	aaaagcgatcgccaTGgCTGCAAGATCAGCTGGCCTAGC
7	pTV-MCS-BamHI-PmeI-3'	aaaacaaatttggaTCCCTGGTTGAGTTCGATAGCACG
8	SgfI-EcoRI-DDEnc-5'	aaaagcgatcgccgAAttcGACAACCAAAATCCCTCGTACGGAC
9	DDEnc-NcoI-PmeI-3'	aaaacaaatttgccATGgAAACAACGTTGGCGAGGCACTTGGG
RT-10	RT-DDE-629-fw	ACATCTCAGCCGGACAAAACA
RT-11	RT-DDE-729-rev	CCAATTCAGTTTGCCGAAGAAC
RT-12	RT-GapDH-775-fw	ACGGCCGATGTTTCTATGGT
RT-13	RT-GapDH-875-rev	ATCGGTCCTTCTGACGCCTT
RT-14	RT-Actin-24-fw	TGAGACCTTCAATGTCCCGG
RT-15	RT-Actin-124-rev	CATCGCCTGAGTCGAGAACAC

Table 2: Primers used for the construction of transformation vectors. The D1(A) primers were used to introduce point mutations into the vectors containing 0.8kb or 3.5kb of psbA sequence. Small letters in primer sequences indicate the conferred nucleotide substitutions.

to the final transformation vector pPha-T1 RNAi.

Biolistic transformation

Cells were bombarded using the Bio-Rad Biolistic PDS-1000/He Particle Delivery System (Bio-Rad Laboratories, Hercules, Canada) fitted with 1350 psi rupture discs. Tungsten particles (0.7µm median diameter) were coated with 5 µg of plasmid DNA in the presence of $CaCl_2$ and spermidine, as described by the manufacturer. One hour prior to bombardment approximately 10^8 cells were spread in the center of a plate containing 20 ml of solid culture medium. The plate was positioned at the second level within the Biolistic chamber for bombardment. Bombarded cells were allowed to recover for 24 h before being suspended in 1 ml of sterile 50% artificial seawater medium. 250 µl of this suspension were plated onto solid medium containing 50 µg/ml Zeocin. The plates were incubated at 20°C under constant illumination (40 µmol photons $\cdot$ m^{-2} $\cdot$ s^{-1}) for three weeks.

Fluorescence measurements to determine NPQ

Since no Anti-DDE antibodies are available for *Phaeodactylum tricornutum* yet, the direct impact of the gene silencing approaches on the protein level could not be analysed. However, PAM (Pulse Amplitude Modulated) imaging fluorescence measurement was shown to be a convenient technique to determine the NPQ phenotype after silencing of the *dde* gene. Also, this technique allows the screening of numerous strains without the need of extensive cultivation or sample preparation. The only prerequisite for accurate and reproducible measurements is to form equal cell layers prior to fluorescence measurements in order to avoid signal variations. Therefore 50 µl of a fresh prepared solution containing 5 10^{-6} cells/ml in 50% seawater medium were spotted on solid medium plates. After allowing cells to divide for 3-4 days to form closed homogenous cell layers, the cells were dark adapted for at least three hours and subsequently analysed using a PAM fluorometer (PAM IMAG-K, Walz, Effeltrich, Germany).

The measurement is initiated by switching on the measuring light, thus giving a measure of the F_0 (minimal) level of fluorescence. Applying a saturating flash of light then allows the measurement of the maximal fluorescence F_m in the dark-adapted state. Subsequently actinic light is applied followed by further saturating flashes in appropriate intervals. From each of these, a value for F_m', the fluorescence maximum in the light, can be measured. With these different measured fluorescence levels NPQ can be evaluated according to the equation $NPQ = (F_m - F_m')/F_m'$.

In several subsequent rounds of measurements the measuring light intensity was increased from 0 µmol photons m^{-2} s^{-1} (over 1, 37, 67, 97, 172, 287, 457 and 742 µmol photons m^{-2} s^{-1}) to 1207 µmol Photons m^{-2} s^{-1}, thus measuring NPQ as a function of increasing light intensities. The settings of the used PAM fluorometer were adjusted as follows: "actinic light" = 20, "delay" = 60, "clock" = 30 and "duration" = 375.

Real-time PCR

Isolation of RNA and cDNA synthesis

Cells were harvested by a brief centrifugation for 1 min at room temperature. The pellet was immediately shock-frozen in liquid N_2. Cell pellets were homogenised by grinding the frozen pellets under liquid N_2. 1ml Trizol (Invitrogen, Carlsbad, CA, USA) was added to the deep frozen grinded powder, which was further homogenised in subsequent steps by vortexing and shaking the solution at room temperature. After adding 200µl of chloroform, rigorous mixing and centrifugation at 4°C, the aqueous phase was transferred into pre-cooled tubes. After adding 1 volume of ethanol (70%) the solution was transferred into RNeasy RNA purification columns (Qiagen, Hilden, Germany). Subsequent RNA purification steps were performed according to the manufacturer's instructions using the provided RW1 and RPE buffers. Although both Trizol and the

RNeasy purification are designed to remove genomic DNA from the RNA extract, we additionally treated aliquots of the extracted RNA with Turbo-DNase (Ambion, Woodward, TX, USA). The obtained genomic DNA-free RNA was reverse transcribed using the reverse transcriptase provided by the QuantiTect reverse transcription Kit (Qiagen, Hilden, Germany). Complete removal of genomic DNA from RNA samples was verified by PCR amplification of intergenic regions after cDNA synthesis. The resulting genomic DNA free cDNA was further used for Real-time PCR assays.

Real-time PCR assays

Real-time PCR was performed using the Real-Time PCR System 7500 (Applied Biosystems, Lincoln, CA, USA). The following program was utilized for all genes: 10 minutes of pre-incubation at 95°C followed by 40 cycles for 15 seconds at 95°C and one minute at 60°C. Individual real-time PCR reactions were carried out in 20 µl volumes in 96-well plates using Power SYBR® Green PCR Master Mix and optical covers by Applied Biosystems. All samples were analysed in six replicates per experiment and each experiment was repeated independently at least twice. At the end of each reaction, the Cycle threshold (Ct) was manually set at the level that reflected the best kinetic PCR parameters, and melting curves were acquired for analysis.

Data analysis using the $2^{-\Delta\Delta Ct}$ method and its validation

In this work, the $2^{-\Delta\Delta Ct}$ method for relative quantification (Livak & Schmittgen, 2001) was used to analyse the generated data and to calculate relative *dde* transcript quantities. This method allows the estimation of copy numbers in unknown samples. The method requires the existence of at least one calibrator for which the transcript quantity is assumed to be 1. Further a house-keeping gene (endogenous control) constantly expressed in all samples is required, to perform normalization of the quantitative data. In this work, wildtype cDNA was used as calibrator, while GapDH served as the endogenous control gene in all experiments. The ΔΔCt calculation for the relative quantification of target was used as follows $\Delta\Delta Ct = (Ct, \text{target gene}_{sample} - Ct, \text{endogenous control}_{sample}) - (Ct, \text{target gene}_{calibrator} - Ct, \text{endogenous control}_{calibrator})$. For amplification of the target (*dde*) fragment the primers No. RT-10 and No. RT-11 were used, the endogenous control fragment (GapDH) was amplified using the primers No. RT-12 and No. RT- 13. After validation of the method, the results for each sample were expressed in N-fold changes compared to the calibrator (wildtype) target copies. Therefore the results for each sample were normalized to the endogenous control GapDH relative to the copy number of the target copies in the calibrator sample, according to the following equation: amount of target = $2^{-\Delta\Delta Ct}$. The resulting normalised N-fold changes of target copy numbers compared to the calibrator were expressed in relative quantities (RQ).

All performed experiments were repeated using the actin gene as endogenous control (primers No. RT-14 and No. RT-15) in order to confirm the obtained results.

RNaseIII assays

In order to distinguish between the amounts of single-stranded and double-stranded *dde* transcript quantities determined by real-time PCR, an aliquot of each RNA extract was treated with ShortCut RNase III (NEB, Ipswitch, MA, USA) according to the manufacturer's instructions. RNase III converts long double-stranded RNA into a heterogeneous mix of short (18-25 bp) RNAs which cannot serve as template for real-template PCR. The digested RNA samples were reverse transcribed and the resulting cDNA analysed via real-time PCR. The obtained results were compared to relative *dde* transcript quantities in cDNA from untreated dsRNA containing (total) RNA.

II.3.4 Results

Fragments of the diadinoxanthin de-epoxidase (*dde*) (Fig. 2A) gene were ligated into pPha-T1 in reverse complementary (antisense) orientation downstream of the fcpA promoter (Fig. 2B). Hence the resulting plasmids pPha-T1 DDE AS 509 and pPha-T1 DDE AS187 were designed to mediate transcription of antisense fragments of 509bp and 187bp length. The antisense fragments were amplified from the N-terminal not conserved part of *dde* to avoid unspecific cross-reactions of the antisense transcripts with transcripts of other de-epoxidases (Fig. 2A). In addition the plasmid pPha-T1 DDE-RNAi

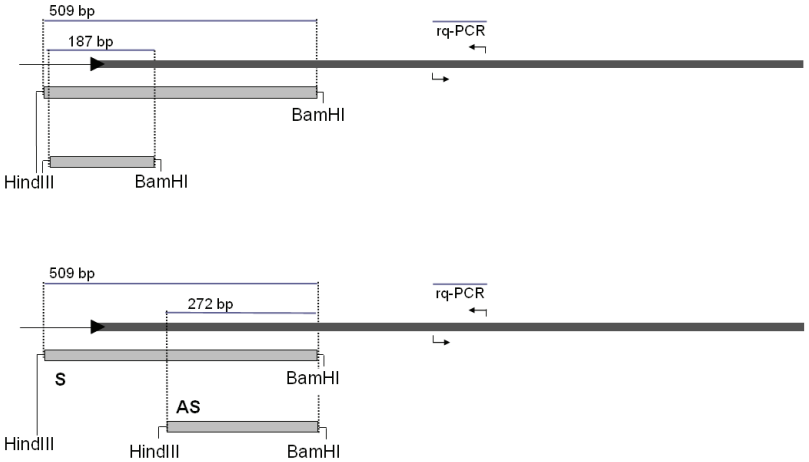


Figure 2: (A) Fragments of the non-conserved N-terminal part of the *dde* gene were cloned into pPha-T1 in antisense orientation. (B) fragments of the N-terminal part of dde were cloned into pPha-T1 in sense (S) and antisense (AS) orientation separated by a linker sequence. The amplified fragments are indicated by light grey bars, the dde gene by the dark greay bars. HindIII and BamHI restriction sites were added by PCR primers. "Rq-PCR" indicates the fragment amplified during relative quantification of *dde* transcripts.

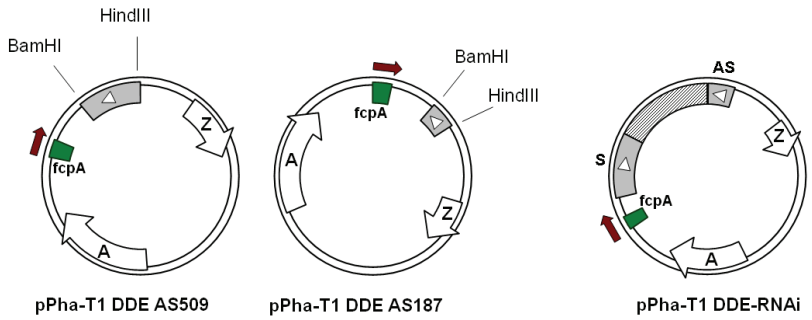


Figure 3: Resulting transformation vectors. The plasmids contan either 509bp (pPha-T1 DDE AS509)or 187bp (pPha-T1 DDE AS187) of N-terminal *dde* sequence in antisense orientation downstream of the fcpA promoter or a self complementary fragment of a 509bp dde sense fragment and a 272bp dde antisense fragment separated by fcpA-5'UTR and eGFP (pPha-T1 DDE-RNAi).

was designed to mediate the transcription of complementary inverted repeats which can fold back on themselves to form a dsRNA hairpin-like structure. Therefore 509bp of *dde* (sense orientation) were ligated to the fcpA 5'UTR followed by the eGFP gene.

Downstream of eGFP followed a 272bp *dde* fragment in antisense orientation. Instead of the commonly used intron linker between sense and antisense fragment we used eGFP as a linker featuring reporter gene function. Since it was still unclear whether *Phaeodactylum tricornutum* was suitable for RNA interference mediated gene silencing, the reporter linker was designed to potentially indicate transcript degradation vs. stability via fluorescing after its expression from the pPha-T1 DDE RNAi transcripts. Since *dde* contributes to NPQ by driving the conversion of diadinoxanthin to diatoxanthin in excess light, silencing of the *dde* gene was expected to result in an altered NPQ phenotype. NPQ is measurable via PAM imaging techniques, therefore examination of putative silenced transformants could easily be performed.

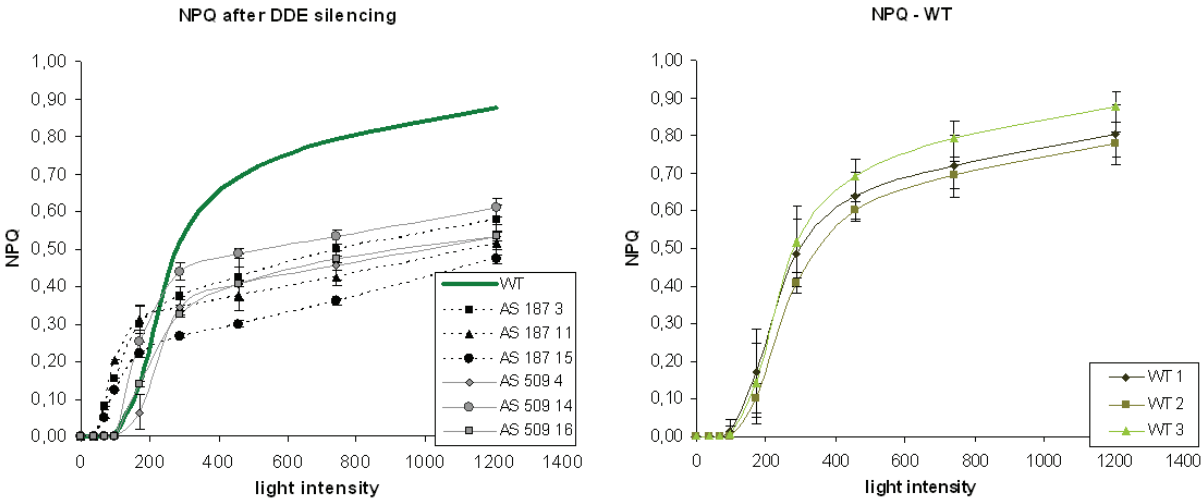


Figure 4: (A) NPQ as function of light intensity measured for six AS-strains. Transformation of cells with *dde* antisense vectors resulted in a significant decrease of NPQ compared to wildtype (corresponds to WT 3 in figure 4B. (B)Accuracy of PAM imaging fluorescence measurements. NPQ was measured as a function of the light intensity in three different wildtype (WT) cultures in independent experiments.

	PAR [µE]	WT	AS 187 3	AS 187 11	AS 187 15	AS 509 4	AS 509 14	AS 509 16
	97	0,022	0,039	0,051	0,031	0,000	0,000	0,000
□NPQ [%]			***79,72***	***135,02***	***42,86***	***-100,00***	***-100,00***	***-100,00***
	287	0,147	0,094	0,086	0,067	0,086	0,110	0,082
□NPQ [%]			***-36,23***	***-41,66***	***-54,55***	***-41,66***	***-25,37***	***-44,37***
	457	0,176	0,106	0,094	0,075	0,102	0,122	0,102
□NPQ [%]			***-39,81***	***-46,62***	***-57,41***	***-42,08***	***-30,72***	***-42,08***
	1207	0,221	0,145	0,129	0,118	0,133	0,153	0,133
□NPQ [%]			***-34,33***	***-41,58***	***-46,56***	***-39,76***	***-30,71***	***-39,76***

Table 2: NPQ in Antisense (AS) strains and wildtype at various light intensities. ΔNPQ [%] indicates increase or decrease in NPQ compared to wildtype.

All three vectors were transformed into *Phaeodactylum* cells using biolistic methods. The transformation yielded zeocin resistant colony forming units (cfu's) with rates of 10^{-7} cfu's/transformed cell. The phenotype NPQ of 50 obtained cfu's was analysed via Imaging PAM, measuring the NPQ parameters Fm and Fm´ as a function of light intensity. The fluorescence measurements revealed a majority of examined cfu's expressing phenotypes different from wild type. 64.3% of cfu's obtained from transformations with pPha-T1 DDE AS 509 (AS509 strains), 62.5% of cfu's from transformations with pPha-T1 DDE AS187 (AS187 strains) and 60% of the pPha-T1 DDE-RNAi transformant cfu's (RNAi strains) showed phenotypes with repressed NPQ. For more precise fluorescence measurements the strains identified by NPQ measurements were adjusted to 5 10^6 cells/ml in 50% seawater. Cell suspensions (50 µl) were allowed to settle on solid media plates thus forming thin and homogenous layers of *Phaeodactylum tricornutum* cells. After three to four days of incubation at constant illumination (35 µmol photons m^{-2} s^{-1}) the layers were dark adapted for three to six hours and analyzed via Imaging PAM measurements. NPQ parameters were measured at different light intensities and plotted in comparison to wild type cells. To confirm the reliability of both cell layer production and the fluorescence measurement protocol, the NPQ was measured for different wildtype cultures in three independent experiments performed on different days. The results - each obtained from five replicates - are shown in Figure 4A. The wildtype NPQ-light curves differ slightly between the independent experiments, but the standard deviations for the different curves (indicated by error bars) were below 0.024 at light intensities above 287µE (µmol photons m^{-2} s^{-1}), thus demonstrating the accuracy of the method.

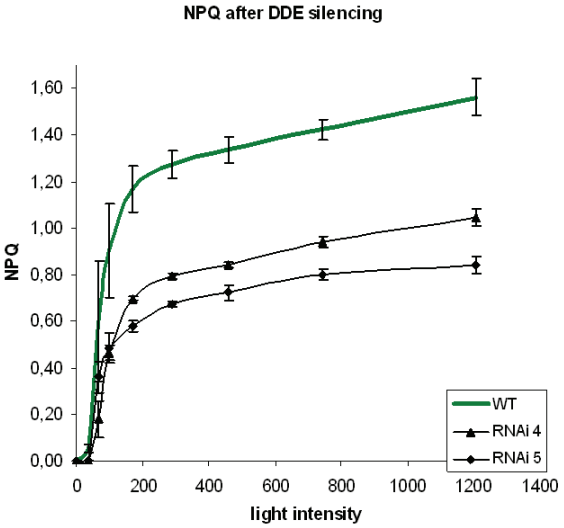


Figure 5: NPQ as function of light intensity measured for two RNAi-strains. Transformation of cells with *dde* RNAi vectors resulted in a significant decrease of NPQ compared to wildtype

	PAR [µE]	WT	RNAi 4	RNAi 5
	97	0,226	0,116	0,122
ΔNPQ [%]			***-48,67***	***-46,24***
	287	0,319	0,198	0,169
ΔNPQ [%]			***-37,85***	***-47,11***
	457	0,334	0,210	0,181
ΔNPQ [%]			***-37,13***	***-45,96***
	1207	0,391	0,261	0,210
ΔNPQ [%]			***-33,18***	***-46,24***

Table 3: NPQ in RNAi strains and wildtype at various light intensities. ☐NPQ [%] indicates increase or decrease in NPQ compared to wildtype.

In subsequent experiments wildtype cells and the NPQ repressed strains were always analysed in the same experiment. At least 80% of the AS or RNAi strains showed a decrease in NPQ by more than 10%. However, the majority of the examined transformants showed between 30% to 46.6% reduction in NPQ compared to wild type cells (Fig 4A, Fig 5). A selection of strains featuring prominent phenotypes is listed in Table 2 and Table 3. The maximal reduction in NPQ was observed for AS187-11 (46.6% NPQ reduction), AS509-4 and -16 (both: 39.8% NPQ reduction) and RNAi-5 (46.2% NPQ reduction). The strong NPQ phenotype suggested successful silencing of *dde*. In order to confirm these results, the transcript levels of the diadinoxanthin de-epoxidase gene were determined in selected AS and RNAi strains compared to wildtype via relative quantification using real-time PCR. Six different strains featuring the most prominent NPQ phenotypes were analysed regarding their *dde* transcript levels by applying the $2^{-\Delta\Delta Ct}$ method. RNA was extracted and reverse transcribed from cultures grown under normal light conditions (45µmol photons m^{-2} s^{-1}). The relative quantities (RQ values) (Table 4) were obtained from experiments using the GapDH gene as endogenous control. Remarkably, all analysed AS strains showed an increased *dde* transcript level compared to wildtype. Relative transcript quantities for the strains AS187-11, AS509-4 and AS509-16 were 6 to 10 fold higher. Only for AS187-15 the measured RQ was 3 fold higher than the wildtype transcript level. In contrast to these findings the detected *dde* RQ values in the RNAi strains and in the wildtype were the same. No significant increase or decrease in the *dde* transcript level could be observed. RNase III assays were performed in order to investigate whether the elevated transcript levels in the AS strains resulted from an accumulation of dsRNA in the cytoplasm. Total RNA from wildtype, AS- and RNAi strains was treated with RNase III, thus degrading all dsRNAs prior to reverse transcription and real-time PCR. The results shown in table 4 clearly demonstrate that the relative amounts of single-stranded (ss) *dde* transcript detected in the AS strains were approximately equal than in the wildtype. RNase III treatment led to a significant decrease of *dde* RQ values for all AS strains compared to wildtype (Fig. 6), thus indicating that the relative increase of *dde* transcripts in total RNA of AS strains resulted from accumulating dsRNAs. RNase III treatment had no significant effects on the relative amounts of *dde* transcript in the RNAi strains. RQ values detected for reverse transcribed total RNA were not significantly different from those observed in reverse transcribed total RNA of RNAi 4 and RNAi 5.

All monitored RNAi strains were also analysed via fluorescence microscopy for eGFP expression from the RNAi transcripts, but no fluorescence was observed for any of the analysed transformants.

	sample	RQ	RQ min	RQ max
total				
	AS 187-	*7,35*	*6,08*	*8,90*
	AS 187-	*3,04*	*2,10*	*4,51*
	AS 509-	*9,99*	*8,31*	*12,03*
	AS 509-	*6,20*	*4,81*	*8,00*
	RNAi 4	*1,49*	*1,16*	*1,94*
	RNAi 5	*0,76*	*0,55*	*1,07*
	WT	*1,00*	*0,80*	*1,26*
ss RNA				
	AS 187-	*1,96*	*1,45*	*2,64*
	AS 187-	*0,91*	*0,75*	*1,12*
	AS 509-	*0,66*	*0,35*	*1,30*
	AS 509-	*0,64*	*0,43*	*0,96*
	RNAi 4	*0,81*	*0,58*	*1,15*
	RNAi 5	*1,34*	*0,88*	*2,04*
	WT	*1,00*	*0,63*	*1,58*

Table 4: Relative quantification of dde transcripts via Real-time PCR. Relative transcript quantities (RQ) were determined in reverse transcribed total RNA obtained from wildtype (WT), AS- and RNAi strains. Transcript levels were also quantified in reverse transcribed RNase III treated total RNA (ssRNA). RQ min and max represent the deviation from the average values.

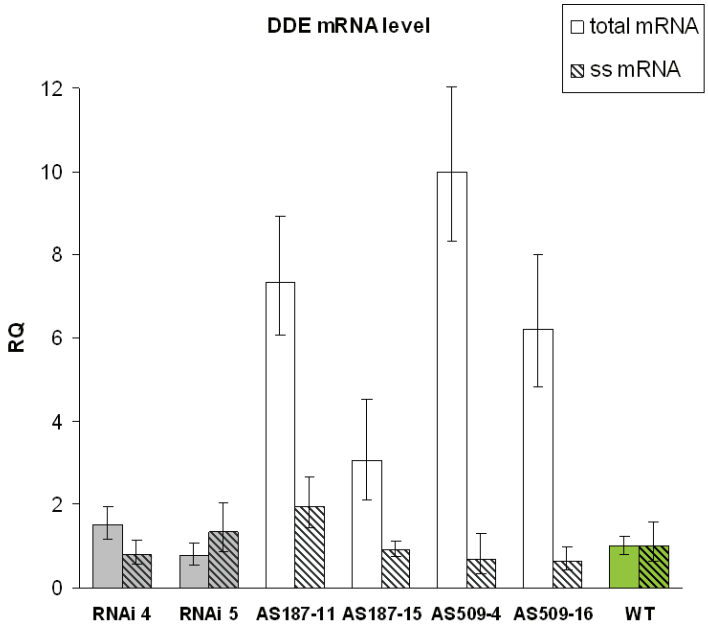


Figure 6: Relative quantification of *dde* transcripts in total RNA and ssRNA via Real-time PCR. Error bars indicate the deviations from the average RQ values.

II.3.5 Discussion

The purpose of this study was to establish a protocol for gene silencing in the diatom *Phaeodactylum tricornutum*. In order to induce RNA interference we transformed the cells with pPha-T1 AS plasmids designed to provide transcription of short (187bp) or long (509bp) fragments complementary (antisense) to the target gene. The second strategy implied the usage of the transformation vector named pPha-T1 RNAi which allows expression of a self- complementary hairpin-like construct with a 5'-sense overhang (Fig. 2). The two complementary flanking regions of the construct were separated by a linker consisting of the fcpA 5'-UTR followed by the eGFP gene. Both plasmid types AS and RNAi were designed to generate different dsRNA fragments homologous to target gene transcripts.

The diadinoxanthin de-epoxidase gene *dde* served as target for the gene silencing approaches. Since the *dde* mediated conversion of diadinoxanthin (DD) to diatoxanthin (DT) under excess light conditions is an inevitable reaction for NPQ, the silencing of this gene was likely to induce a clearly identifiable phenotype. Indeed the silencing approaches generated transformants with a phenotype clearly distinguishable from wildtype cells. Both approaches yielded similar rates of NPQ repressed transformants. Also the extent of NPQ reduction compared to wildtype was similar for AS and RNAi strains. Altering the lengths of the *dde* antisense fragments from 187bp to 509bp did not further affect the phenotype. In some cases NPQ was decreased by more than 40% compared to wildtype with a maximal observed decrease of 46.6%. Since NPQ is a complex process which is certainly influenced by various factors and genes, the observed impact of silencing the *dde* gene on NPQ is remarkable and demonstrates once more the importance of the de-epoxidation of DD for the thermal dissipation process. PAM imaging fluorescence measurements were shown to be accurate and suitable to identify algal strains impaired in NPQ (Fig 4A, 4B). Differences in absolute NPQ values may occur when analysing cells in independent experiments performed on different days. However the relative extents of NPQ decrease in silenced strains compared to wildtype remains the same. Nevertheless, dark adaptation and fluorescence measurements of wildtype and silenced strains should preferably be performed simultaneously and by applying the same experimental conditions.

Quantification of the *dde* transcript levels via Real-time PCR revealed significant differences between the wildtype and AS strains. The relative transcript quantities in AS strains were up to ten fold increased compared to wildtype. RNase III assays revealed that the increased transcript levels are mainly attributed to an accumulation of double-stranded target RNA. These findings indicate that *dde* mRNA effectively hybridises with the antisense fragments, thus forming dsRNA that which is not translated. Potential

attempts of the diatom to compensate negative effects of *dde* silencing might result in an upregulation of the endogenous *dde* transcription which would consequently lead to an increasing mRNA-antisenseRNA hybrid formation. The assumption that the formation of these dsRNAs is more efficient than the subsequent target RNA degradation explains the accumulation of dsRNA, as observed in this study. Another mechanism which is known to generate dsRNA molecules of the target RNA prior to its degradation involves cellular RNA-directed RNA polymerases (the abbreviation for plant RNA-directed RNA polymerases genes was recently changed from RdRP to RDR (Xie et al. 2004; Wassenegger and Krczal 2006)). RDR were shown to be involved in silencing pathways in plants (Dalmay et al. 2000; MacAlpine et al. 2000; Mourrain et al. 2000), fungi (Cogoni and Macino 1999), and *C. elegans* (Smardon et al. 2000; Sijen et al. 2001). This class of polymerases recognises abnormal or "aberrant" transcripts in cells and subsequently mediates primer-independent synthesis of the complementary RNA strand (Schiebel et al. 1993; Makeyev and Bamford 2002), thus generating double-stranded target RNA (Dougherty and Parks 1995; Baulcombe 1996). Neither this recognition mechanism nor are all the characteristics which classify aberrant RNA are completely understood, however the generated dsRNA further leads into downstream silencing pathways. Therefore it could be hypothesised that an RDR related mechanism also recognises the antisense transcripts as aberrant and consequentially initiates the synthesis of the complementary strand, thus increasing the amount of double-stranded target RNA in the cell.

Both approaches induced the same NPQ repressed phenotype in AS- and RNAi strains, the maximal extend of NPQ repression was in both cases ~45%. In contrast to the AS-strains the *dde* transcript levels in the RNAi-strains did not vary from wildtype. The RNase III assays did not reveal any accumulation of double-stranded *dde* transcripts. These findings suggest the presence of different underlying silencing mechanism in *Phaeodactylum tricornutum*. The RNAi strains were further analysed via fluorescence microscopy in order to investigate the pPha-T1 RNAi transcript stability. The linker in the self-complementary construct contained the eGFP gene equipped with an upstream fcpA 5'-UTR to guarantee the translation of eGFP. However, green fluorescence was not detected in any of the monitored RNAi strains. The absence of eGFP fluorescence might indicate an active RNA interference phenomenon which prevents not only the expression of the diadinoxanthin de-epoxidase but also of eGFP. The fact that NPQ reduction was observed while *dde* transcript levels in RNAi strains were not reduced compared to the wildtype hints to a silencing mechanism involving translational repression. The translational repression mechanism was already previously shown to reduce specifically protein synthesis without affecting mRNA levels (Bartel 2004).

In conclusion, the results presented in this study clearly demonstrate the feasibility of effective gene silencing in *Phaeodactylum tricornutum*. Our approaches to silence the *dde*

gene revealed strains displaying significantly reduced NPQ capacities. However, the data suggest the presence of different RNA interference mechanisms in diatoms. To date the underlying mechanisms leading to gene silencing are unclear – we emphasize that the mechanistic assumptions made in this report are speculative and require further investigations.

Future physiological and molecular examinations of the generated strains will provide further insights in the xanthophyll cycle in diatoms and help to increase our knowledge about photoprotection.

Quantification and visualisation of organellar genomes

Materna, A.C.[1], Gruber, A.[1] and Kroth, P.G.[1]

1 Department of Plant Ecophysiology, University of Konstanz, Germany

II.4.1 Abstract

Great progress has been made in sequencing genomes of a variety of different organisms. In contrast to sequence analysis which reveals the genetic information encoded in the nucleotide composition of DNA, this study presents two techniques focusing on the genomes themselves as structural entities, safely maintained and replicated within nuclei and organelles. We developed a protocol for the diatom *Phaeodactylum tricornutum*, to allow selective *in vivo* visualisation of chloroplast nucleoids. This labeling method was also applied to mitochondrial nucleoids. Further a high throughput capable method was designed to quantify organellar genomes thus determining ploidies with high accuracy. The method bases on a specially designed quantitative Real-time PCR protocol and was confirmed by fluorescence measurements. These techniques allow determination and monitoring of organellar ploidies and of nucleoid numbers per organelle. Further the subcellular localisation of nucleoids inside the organelles, as well as their genome contents can be studied.

Key words: *Phaeodactylum tricornutum*, chloroplast, plastid genome, mitochondria, mitochondrial genome, nucleoid, DNA content, genome copy number.

II.4.2 Introduction

Organelles, such as plastids and mitochondria are derived from prokaryotic ancestors, which have been engulfed by a eukaryotic host cell. In this endocytobiosis, both, genome and structure of the endosymbiont have been reduced during evolution of host and organelle likewise. Genes of the prokaryote have been deleted or predominantly transferred into the nucleus to be furthermore under control of their host. However, recent organelles still harbour reduced but functional circular genomes with sizes ranging from 120 to 160 kb for plastid genomes (plastomes) of higher plants (Sugiura 1992) and up to 190 kb for those of photosynthetic protists such as rhodophytes (Reith and Munholland 1995; Hagopian *et al.* 2004). In contrast the sizes of mitochondrial genomes (chondriomes) vary widely. While higher plants display the largest mitochondrial genomes (200-2,400 kb) those of animals (14-42 kb), protists (5.7-76 kb) and of fungi (18-176 kb) are comparatively small (Backert *et al.* 1997; Gray *et al.* 1998). Organellar genomes still contain a set of functional genes. Plastid genomes usually harbour approximately 120 genes while the chondriomes have a lower coding capacity and usually contain only approximately 60 genes. Organellar genes mainly encode for ribosomal RNA's, ribosomal proteins, and elongation factors, as well as for genome specific tRNA's to drive the formerly prokaryotic machinery for protein biosynthesis. Further genes encoding proteins involved in secretion, translocation, and in energy conservation (ATP Synthase subunits; NAD(P)H dehydrogenases) are also generally found in both mitochondrial and plastid genomes. Finally, organellar genomes contain a varying number of function specific genes involved in either photosynthesis in the case of plastomes (like the psa- and psb-genes, and the ycf ORF's) or in the oxidative electron transport for the mitochondrial genomes (e.g. cox-genes).

Plant and algal cells contain both organellar types, plastids and mitochondria. Therefore their entirety of inheritable traits is distributed among three different genomes localised in the compartments of the organelles and the nucleus. In contrast to the nucleus, the organelles are highly polyploid leading to genome copy numbers of up to 10,000 chloroplast genome copies in a single leaf cell(Kuroiwa 1982; Bendich 1987; Kuroiwa 1991; Coleman and Nerozzi 1999). The organellar genome copies usually are attached to each other thus forming several clusters per organelle, the so called nucleoids. Nucleoids usually bind to intraorganellar membranes employing special proteins with DNA binding and membrane binding domains(Jeong *et al.* 2003; Sato *et al.* 2003; Dai *et al.* 2005).

Visualising Nucleoids

The distribution of nucleoids in plastids is in the focus of several investigations since nucleoid localisation is believed to be subject to dynamic changes during plant cellular development(Sato *et al.* 2003). In accordance to these findings also nucleoids of

mitochondrial origin are involved in dynamic changes during organellar propagation(Dai *et al.* 2005). In the previous studies visualisation of organellar genomes has already been achieved in some rare cases. In order to study their function the thylakoid-associated nucleoid-binding proteins MFP1 and PEND have been fused to GFP (green fluorescent protein) resulting in fluorescing nucleoids inside the chloroplasts (Jeong *et al.* 2003; Terasawa and Sato 2005a, 2005b). These proteins and their homologues have been detected in various angiosperms, including Arabidopsis thaliana, Brassica napus, Medicago truncatula, Cucumis sativus (cucumber) and Prunus yedonensis (cherry). Monocot homologues have been also detected in barley and rice, but sequence conservation was low in monocots. However, related sequences have not been identified in non-flowering plants and any algae so far (Terasawa and Sato 2005a). Labelling of nucleoids in mitochondria has been achieved as well, however only in the budding yeast, Saccharomyces cerevisiae. The nucleoids have been labelled using the HMG-box DNA-binding protein Abf2 fused to GFP in order to monitor mitochondrial dynamics during meiosis and sporulation in the yeast(Gorsich and Shaw 2004).

Further fluorescence dyes such as DAPI have been widely used in the past for unspecific staining of nucleic acids, thus visualising organellar nucleoids (Williamson and Fennell 1975; Coleman 1979; Williamson and Fennell 1979; Stevens 1981). When using intact cells for exact localisation or even quantification experiments however, this technique is hardly the method of choice since the strongest source of fluorescence is the nucleus therefore outshining the nucleoids of nearby organelles. Furthermore fluorescence dyes like DAPI or SYBR Green are lethal to cells and therefore cannot provide access to in situ monitoring of living cells.

Quantification of organellar genomes

Not only the nucleoids' localisation, but also the overall copy numbers of organellar genomes are of interest, especially since many efforts have been made recently to monitor potentially occurring changes in plastid genome copy numbers (Oldenburg and Bendich 2004; Rowan *et al.* 2004; Li *et al.* 2006). Initially, transcription, replication and segregation of organellar genomes independent from the nuclear genome was shown in the sixties and seventies(Gibbs 1968; Gibbs and Poole 1973). The distribution of radioactively labelled DNA among progeny chloroplasts was therefore determined from autoradiographs via light microscopy(Gibbs 1968; Gibbs and Poole 1973). From the obtained data the amount of genomes (or segregating units) was calculated. Later techniques for determining genome DNA contents and ploidies based on reassociation kinetics of ^{125}I-labelled or UV spectrometricaly detected dsDNA(Gelb *et al.* 1971; Rawson and Boerma 1976; Ersland *et al.* 1981; Ersland and Cattolico 1981). Recent studies conventially rely on fluorophore staining of nucleic acids with subsequent fluorescence

quantification via fluorescence microscopy, thus allowing the determination of genome copy numbers(Misumi *et al.* 2001).

However, fluorophores such as DAPI provide only restricted capabilities to quantify genome copies via fluorescence due to the strong unspecific staining of different cellular structures as e.g. frustules of diatoms or sugars, such as extracellular polysaccharides, which are secreted by many microalgal species. Removal of these extracellular sugars is possible, however, by harsh treatments only which are lethal and therefore provide impact on the subcellular structure. In addition, the unspecific staining of all nucleic acids further contributes to the background fluorescence and therefore potentially produces errors during genome copy number determinations. When utilising isolated organelles for fluorescence quantification it is crucial to remove the chromosomal DNA during organelle isolation which might also lead to degradation of organellar DNA. Hence, fluorophore staining techniques are predominantly useful for monitoring changes in genome copy numbers, however they suffer from the incapability to reveal exact results. That applies equally to Southern blot hybridization of genome specific probes with total cellular DNA which has been recently demonstrated to be the most reliable method (Li *et al.* 2006)for monitoring changes in genome copy numbers.

In this study we are presenting techniques which allow the visualisation of organellar nucleoids selectively and *in vivo* in the diatom *Phaeodactylum tricornutum*. Furthermore a protocol was developed which allows for the first time to determine the level of ploidy, i.e. genome copy numbers of all three compartments - nucleus, plastid and mitochondrium utilising a SYBR Green I detection based quantitative real time PCR (RT-qPCR) technique taking into account PCR efficiencies. The planktonic brackish water alga *Phaeodactylum tricornutum* provides unique conditions to study organellar genome distribution since it harbours only one large chloroplast (up to 10 µm in length) and generally one mitochondrial network including two to three major protuberances. The diatom can be genetically modified via transformation(Apt *et al.* 1996) and is a convenient object for microscopy. Recently the nuclear- and organell-genomes of *P. tricornutum* have been sequenced (http://genome.jgi-psf.org/Phatr1/Phatr1.home.html, Bowler et al., in preparation) and the knowledge of the sequences will inspire and accelerate further physiological studies. Our new techniques to determine DNA contents, genome ploidies and nucleoid localisation allow physiological investigations on the structural aspects of the genomes.

II.4.3 Material and Methods

Strains and media

Phaeodactylum tricornutum Bohlin (University of Texas Culture Collection, strain 646) was grown at 22°C with continuous illumination at 35 µmol photons▪m^{-2}▪s^{-1} in Provasoli's enriched seawater (Starr and Zeikus 1993) using "Tropic Marin" artifcial seawater at 50% concentration, compared to natural seawater. Solid media contained 1.2% Bacto Agar (Difco). *Escherichia coli* strain XL-1 Blue (Stratagene, Cedar Creek, TX, USA) was grown over night at 37°C in Luria Broth medium, using a shaker for liquid cultures (Sambrook *et al.* 1989). Solid media contained 1.5% Bacto Agar (Difco).

Sequence analysis and prediction programs

To identify proteins putativly targeted to the nucleoids, we searched the current *Phaeodactylum tricornutum* genome sequencing project (http://genome.jgi-psf.org/Phatr1/Phatr1.home.html, Bowler et al., in preparation) for sequences with homology to recombinases using the BLAST algorithm (Altschul *et al.* 1997). Resulting hits were screened for the presence of signal peptides by help of the program SignalP (Bendtsen *et al.* 2004). For cleavage site predictions the results of SignalP's Hidden Markov Models (Nielsen and Krogh 1998) were used, for chloroplast or mitochondrial transit peptide prediction, the programs ChloroP (Emanuelsson *et al.* 1999) and TargetP (Emanuelsson *et al.* 2000) (Horton *et al.* 2006)were used. All these prediction servers can be found at http://www.cbs.dtu.dk/services/. For the construction of a mitochondrial targeted recombinase we used a known mitochondrial targeting sequence from the *Phaeodactylum tricornutum* phosphoglycerate kinase precursor (PGK2, GeneBank Accession AF108452), (Liaud *et al.* 2000).

PCR and construction of plasmids

Standard cloning procedures were used (Sambrook *et al.* 1989). PCR was performed in a Master Cycler Gradient (Eppendorf, Hamburg, Germany) using recombinant Pfu polymerase (Fermentas, Ontario, Canada) according to the manufacturer's instructions. Small letters in primer sequences below indicate degenerated nucleotides. All eGFP-fusion constructs have been ligated into the *Phaeodactylum tricornutum* transformation vector pPha-T1 (Apt *et al.* 1996) which allows transformation of the diatom selecting for positive transformants using zeocin. In a first step eGFP has been amplified using the primers Fw (CGGTCAGGCCTATGGTGAGCAAGGGCGAGGAG) and Rev (CGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATC) thus substituting the start codon ATG with the recognition site for StuI allowing in frame cutting. PPha-T1 was linearized using EcoRV and the modified eGFP fragment was ligated into the plasmid in the

orientation of the fcpA promoter resulting in the plasmid pPha-T1-eGFP. The plastid targeted recA without stop codon TAG was amplified using the primers Fw (5'-TGTATGATGCATCGAAAGATTGCG-3') and Rev (5'-GACGAGCAGATATTTTTGGAA-3'). After digesting pPha-T1-eGFP with StuI the recA amplicon was ligated into the plasmid upstream of and in frame with eGFP resulting in the vector RecA:GFP. The sequence encoding for the mature RecA protein without presequence was amplified using the primers Fw (5'-TTCGgaTATcGGACCCCTGTTAGCACGCAG-3') and Rev (5'-ATTCCAAAAATATCTGCTCGTCAACTTCTAG-3') thus adding an EcoRV restriction site to the 5'-end of the recA fragment which allows in-frame-cutting of the sequence. The modified amplicon was ligated in pPha-T1-eGFP after digesting the plasmid with StuI. Insertion upstream and in the orientation of eGFP resulted in the vector recA-eGFP. The presequence of recA was amplified using the primers Fw (5'-TGTATGATGCATCGAAAGATTGCG -3') and Rev (5'- CTTTCATGTTGTCCGCGTCACCGAG-3'). The amplicon was ligated in frame with and upstream of eGFP into pPha-T1-eGFP after digesting the plasmid with StuI resulting in Pt-eGFP. The presequence of the mitochondrial PGK was amplified using the primers Fw (5'-TTTGATATCATGTTCCGTATGTTGACTTC-3') and Rev (5'-CTTGGATCACTGGGGCAGCGTGAAAGC-3'). The obtained amplificat was ligated upstream of and frame with eGFP either in pPha-T1-eGFP after digesting the plasmid with StuI resulting in the plasmid PrePgk:GFP or in Ma-recA-eGFP after digesting the plasmid with EcoRV resulting in the plasmid PrePgk:RecA:GFP.

Biolistic transformation

Cells were bombarded using the Bio-Rad Biolistic PDS-1000/He Particle Delivery System (Bio-Rad Laboratories, Hercules, Canada) fitted with 1350 psi rupture discs. Tungsten particles (0.7 µm median diameter) were coated with 5 µg of plasmid DNA in the presence of $CaCl_2$ and spermidine, as described by the manufacturer. One hour prior to bombardment approximately 10^8 cells were spread in the center of a plate containing 20 ml of solid medium. The plate was positioned at the second level within the Biolistic chamber for bombardment. Bombarded cells were allowed to recover for 24 h before being suspended in 1 ml of sterile 50% artificial seawater medium. 250 µl of this suspension were plated on solid medium containing 50 µg/ml Zeocin. The plates were incubated at 20°C under constant illumination (35 µmol photons$\cdot m^{-2}\cdot s^{-1}$) for three weeks.

Microscopy

Transformants were screened for the expression of GFP using an Olympus BX51 epifluorescence microscope equipped with a Nikon DXM1200 digital camera system

(Olympus Europe, Hamburg, Germany). Nomarski's differential interference contrast illumination was used to view transmitted light images. Chlorophyll autofluorescence and green GFP fluorescence of the transformants have been dissected using the mirror unit U-MWSG2 (Olympus) and the filter set 41020 (Chroma Technology Corp, Rockingham, VT, USA), respectively. Six transformant lines expressing GFP in a detectable amount were selected. All lines displayed varying numbers of nucleoids in the plastids of individual cells which were analysed by counting fluorescing nucleoids during focusing through the chloroplast at 40x magnification (UplanFL objective, Olympus). A total of 313 cells randomly chosen from the transformant lines in equal amounts were examined.

To stain DNA in living *P. tricornutum* cells, Hoechst 33342 dye (Bisbenzimide H 33342 Fluorochrome, Behring Diagnostics, La Jolla, CA, USA) was directly added to the cells suspended in culture medium (10 µg/ml end concentration of the dye) prior to microscopical observation.

Analysis of GFP localisation in *P. tricornutum* transformants was performed with a confocal laser scanning microscope LSM 510 META (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) using a Plan-Apochromat 63x/1.4 Oil DIC objective (Carl Zeiss). GFP and chlorophyll fluorescence was excited at 488 nm, filtered with a beam splitter (HFT 488/543), and detected by two different photomultiplier tubes with a bandpass filter (BP 505–530) for GFP fluorescence and a low pass filter (LP 650) for chlorophyll (Chl) auto fluorescence. Transmitted light was detected at 488 nm excitation light. Hoechst 33342 (Hoe) was excited in a second track at 405 nm, filtered with a beam splitter (HFT 405/488) and detected by a photomultiplier tube with a bandpass filter (BP 420–480). Maximum intensity z-projections were calculated from slices of image stacks to ensure complete detection of Fluorophores within a cell.

Real-time PCR

Isolation of nucleic acids

Total nucleic acids were isolated from cells grown on solid media and from liquid cultures by a cetyltrimethylammoniumbromide (CTAB)-based method (Doyle and Doyle 1990). After collecting the cells by centrifugation and removal of the supernatant a small amount (of the final volume) of CTAB buffer containing 1% β-Mercaptoethanol was added first, followed by crushing the cells using a pistil fitting into the 1.5 ml reaction tubes. After subsequent precipitation the DNA pellet was washed with 80% Ethanol.

Oligonucleotide primer design

Real-time PCR oligonucleotide primers were designed using two different programs. Every primer set was initially computed using the Primer express software (Applied Biosystems, Lincoln,CA,USA). Subsequently the primers revealed by this program have been checked, where required adapted, or even discarded to assure maximal efficiency and sensitivity, according to the following parameters: avoidance of the formation of self and hetero-dimers, hairpins, and self-complementarity, primer length, and melting temperature. These properties were verified using DNASTAR primer select software (version 4.0). All primers were designed to generate an amplicon size of 100 (± 1) nucleotides at an average annealing temperature of 60°C. To verify the suitability for quantitative real-time PCR all oligonucleotide primers were tested by PCR with subsequent electrophoresis of the amplicons on 2% agarose TAE gels. Primers which led after 40 cycles of amplification to unspecific product bands, primer dimer bands or anything else than to a single distinct band were discarded. Further all primers were controlled by melting curve analysis; specific reactions should result in a single melting peak corresponding to the PCR product being amplified. In contrast, multiple melting peaks imply that that the reaction is either unspecific, generating more than one amplicon or that primer dimers are being formed. Only primer pairs leading to effective and specific amplification were used for RT-qPCR experiments.

Real-time PCR assays

Real-time PCR was performed using the Real-Time PCR System 7500, (Applied Biosystems, Lincoln,CA,USA). The default program was utilised for all genes: 10 minutes of pre-incubation at 95°C followed by 40 cycles for 15 seconds at 95°C and one minute at 60°C. Individual real-time PCR reactions were carried out in 20 µl volumes in 96-well plates using Power SYBR Green PCR Master Mix and optical covers by Applied Biosystems. At the end of each reaction, the Cycle threshold (Ct) was manually set at the level that reflected the best kinetic PCR parameters, further melting curves were acquired and analysed.

Data analysis and determination of real-time PCR efficiencies

For genome copy number determination sequences of the mitochondrial and plastid genome were amplified together with a sequence of the nuclear genome from the same template via real-time PCR. The amounts of individual organellar genome copies in the template were quantified relatively to the amount of nuclear genomes which serve as calibrator. Since the ploidy of the nuclear genome is known (diploid), the absolute number of organellar genomes is calculated by multiplying the relative quantities by two.

Instead of relying on the commonly used $2^{-\Delta\Delta Ct}$ method we decided to apply a method which incorporates the various PCR efficiencies of the amplifications of the genome specific sequences. Therefore a convenient mathematical model for RT-qPCR which includes efficiency correction for independent amplificats (Pfaffl, M.W. 2001), used by the software tool REST© (Pfaffl *et al.* 2002), was utilised. This model allows determination of gene expression in a target sample relative to a control sample. The equation of this mathematical model was adapted for relative quantification of genes located in the different genomes which constitute a total DNA extract [1]:

$$\text{copy number} = \frac{(E_{\text{t gen}})^{\Delta Ct\ \text{t gen(n gen sample - t gen sample)}}}{(E_{\text{endo}})^{\Delta Ct\ \text{endo(n gen endo - t gen endo)}}} \quad [1]$$

The target genome copy number is calculated basing on the efficiencies (E) and the Ct (Cycle threshold) difference (Δ) of an unknown sample, in which a sequence from the target genome is amplified versus the amplification of the calibrator, a nuclear genome encoded sequence $\Delta Ct_{(\text{n gen sample} - \text{t gen sample})}$. For computing relative expression levels in different templates the target gene expression is usually normalized by an endogenous control (E_{endo}), e.g. a housekeeping gene transcript. Since for the purpose of relative genome copy quantification the same template is used for the amplification of organellar genome sequences and of a nuclear genome sequence (calibrator), all samples contain the same amount and ratio of template molecules. Therefore the denominator of equation [1] becomes 1, which leads to the equation [2] for genome copy quantification via RT-qPCR:

$$\text{copy number} = (E_{\text{t gen}})^{\Delta Ct\ \text{t gen(n gen sample - t gen sample)}} \quad [2]$$

In order to determine PCR efficiencies for each amplified sequence *Phaeodactylum tricornutum* genomic DNA was diluted in serial 10-fold ranges and the Ct value at each dilution was measured. When ploting the Ct's depending on the dilutions (on a logarithmic scale) the slope of the resulting curve is used for efficiency calculated according to the equation [3] (Bustin 2000):

$$E = 10^{(-1/\text{slope})} \quad [3]$$

The theoretical maximal and optimal value for PCR efficiencies is 2, indicating a duplication of amplicons during every cycle. The coefficient of variation (CV) was calculated according to equation [4]:

CV = standard deviation / mean 100 [4]

Preparation of artificial templates

Artificial templates were prepared from three different plasmids, pCR Script DDE, pGEM-T FBA, and pCR Script PSBA which harbour sequences of the dde gene, thus representing the nuclear genome, and of the fba (a fructose bis-phosphate aldolase; class II) gene and the psbA gene representing the chondriome and the plastome respectively. After calculating the molecular weights of the vectors from their length and basepair composition, stock solutions of each vector were prepared containing the same amount of vector molecules (0.8 µM). The artificial templates were prepared from mixing the equimolar stock solutions of pCR Script DDE, pCR Script PSBA and pGEM-T FBA in the ratios 1:100:30, 30:1:20 and 1:1:1.

Fluorometric DNA detection

Phaeodactylum tricornutum was grown in liquid culture and the cell concentration was determined using a haemocytometer (Thoma chamber, 0.1 mm depth). To stain nucleid acids in the cells, 0.5 µl of the SYBR Green I stock solution (Cambrex Bio Science Rockland, Inc. Rockland, ME, USA) were added to 1.5 ml of the liquid culture. After 45 minutes incubation the stained cells were observed directly in the medium containing the staining agent using a Axiovert 200 M epifluorescence microscope (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) and a 63x oil immersion objective. (Plan-Apochromat 63x/1.40 Oil DIC M27, Zeiss). Autofluorescence of the cells was minimized by the combination of Zeiss Filter Set 44 (Excitation: 455-495 nm, Beam Splitter: 500 nm, Emission: 505-555) with the excitation system Lambda DG-4 (Sutter Instrument Company, Novato, CA, USA) containing an additional bandpass filter (485 to 502 nm), resulting in narrow banded excitation of the cells at 485-495 nm wavelength and emission detection at 505-555 nm wavelength. 12 bit grey level images were aquired with a AxioCamMR3 (Zeiss) controlled by the software AxioVision (Zeiss). Controll images of unstained cells were taken using the same setup and the same settings directly after aquisition of the measurement images to evaluate the background caused by autofluorescence of the cells and by scattered light.

Silhouettes of the cells (transmitted light chanel) and the nuclei (fluorescence chanel) were drawn into the images using AxioVision's outline spline function. The number of pixels (densitometric area, DensArea), the sum of grey values (densitometric sum, DensSum) and the mean grey value of the outlines (densitometric mean, DensMean) was then determined for the fluorescence chanel. To correct the data for background

fluorescence, DensMean of a circle with 10 µm radius nearby the cells was determined and the fluorescence (Fluor.) of a outline was calculated by substraction of the background fluorescence from the Dens Sum of the outline: $Fluor._{outline} = DensSum_{outline} - DensArea_{outline} * DensMean_{background}$. Cells were rejected if their DensMean did not exceed the DensMean of the background at least fourfold.

II.4.4 Results

Labelling of organellar nucleoids

In vivo labelling and visualisation of organellar nucleoids is only available for few species so far – mitochondrial labelling even for budding yeast only. A more general and species independent but especially organelle-independent application of this technique was not accessible so far. Our aim was to design a visualisation tool, potentially useful for a wider range of species which can be targeted to the organelle of interest. Therefore the RecA protein was deployed. RecA is evolutionary conserved with functional and structural analogs in all kingdoms of life(Edelmann and Kucherlapati 1996; Seitz and Kowalczykowski 2000; Wyman and Kanaar 2004). During homologous recombination RecA forms closely packed nucleoprotein filaments leading to a so called presynaptic complex of dsDNA and ssDNA, and therefore features salient DNA binding capabilities(Wyman and Kanaar 2004).

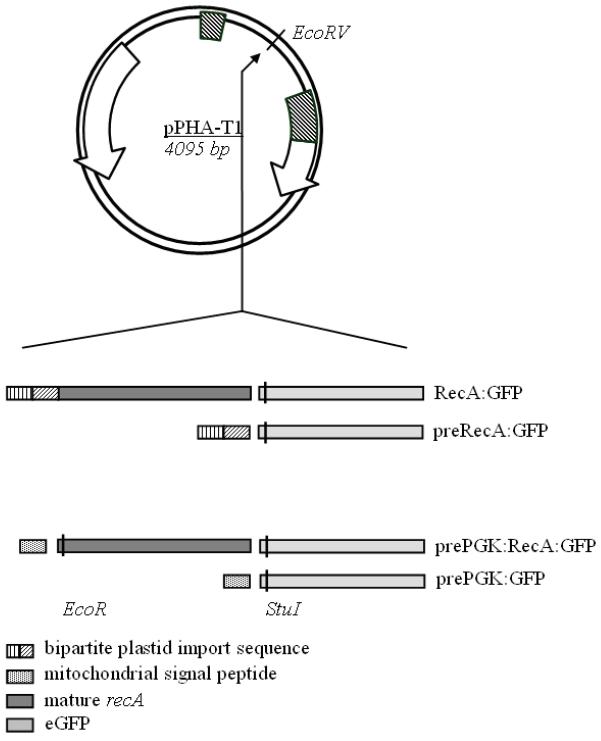


Figure 1: Transformation vectors for thelabelling of nucleoids. RecA:GFP fusions were ligated into the pPha-T1 transformation vector.

We identified a RecA homologue in *Phaeodactylum tricornutum* which posesses a putative N-terminal plastid targeting sequence and prepared a set of vectors suitable to label organellar nucleoids (figure 1). To verify plastidic localisation of the gene product, a fragment coding for the presequence only was fused to eGFP (PreRecA:GFP, figure 1). *Phaeodactylum tricornutum* transformants expressing this fusion protein showed GFP fluorescence colocalised with the chlorophyll autofluorescence (Figure 2 b), the RecA protein of *P. tricornutum* is therefore indeed targeted to the chloroplast.

In order to label plastic nucleoids in *Phaeodactylum tricornutum* the chloroplast targeted recA gene was fused to eGFP (RecA:GFP, figure 1). *Phaeodactylum tricornutum* transformants expressing this fusion protein showed GFP fluorescence associated with the chlorophyll autofluorescence, but restricted to small spots at the periphery of the chloroplasts (figure 2 c). Within these spots GFP fluorescence colocalised with the

a GFP fusion protein constructs

preRecA:GFP

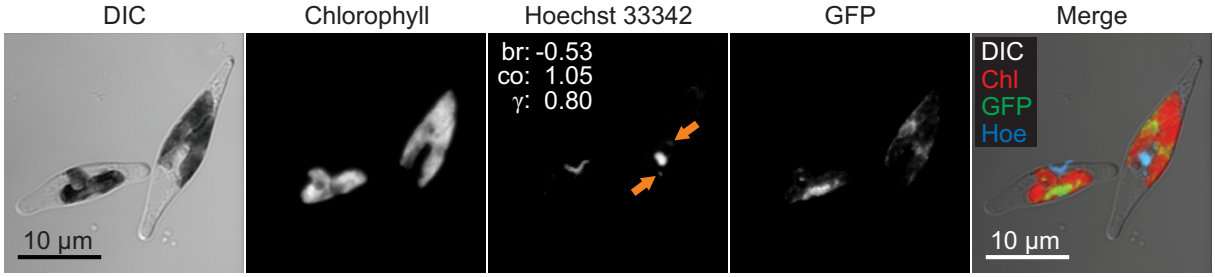
01 05 10 15 20 25 30 35 40 45 50 55 60
MMHRKIALVAGLWFCLLGSSCHGFGLLGGAAHSRTVRWPQQWAVSSPIGPLL*ARSKKSSN*
*DSNDDNELENERPTMDPAKRAALDGVLNQIERNYGRGSIVKLGDADNMK*p:eGFP
61 65 70 75 80 85 90 95 100 105 109 349

RecA:GFP

01 05 10 15 20 25 30 35 40 45 50 55 60
MMHRKIALVAGLWFCLLGSSCHGFGLLGGAAHSRTVRWPQQWAVSSPIGPLL*ARSKKSSN*
DSNDDNELENERPTM - 355 aa - *E*p:eGFP
61 65 70 75 431 671

BOLD: signal peptide predicted by SignalP's hidden Markov models,
UNDERLINED: estimated transit peptide domain,
ITALIC: mature protein, lower case: artificial proline,
GREY: conserved phenylalanine at signal peptide cleavage site

b Localisation of preRecA:GFP



c Localisation of RecA:GFP

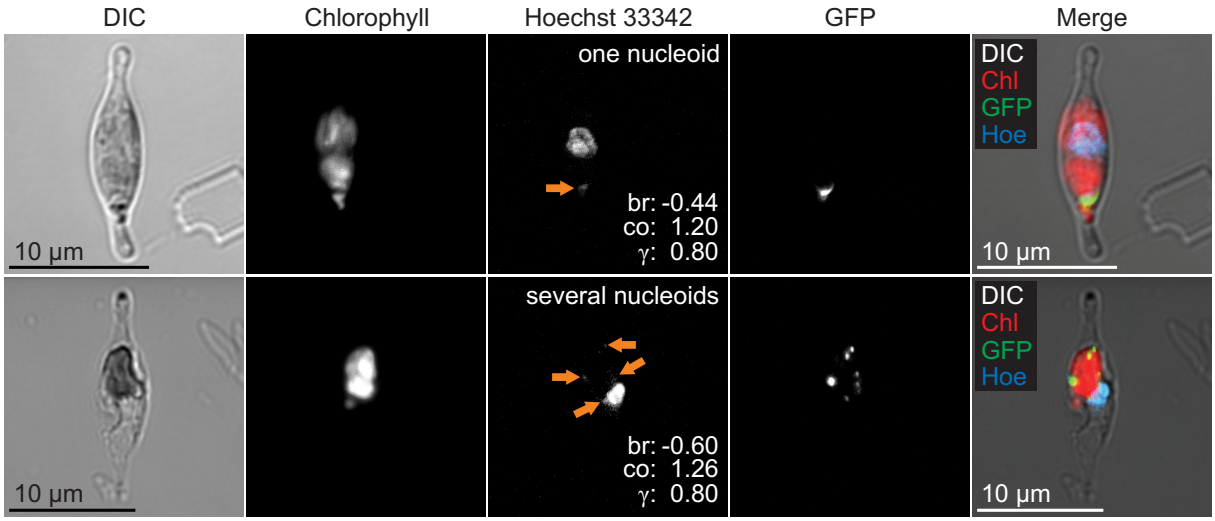


Figure 2: Selective labeling of plastidic nucleoids. a: GFP fusion protein constructs, presequence (preRecA:GFP) and full length (RecA:GFP) Phaeodactylum tricornutum RecA fused to GFP, presequence structure as indicated. b: Localisation of preRecA:GFP, Nomarski's differential interference contrast (DIC), Chlorophyll autofluorescennce (Chl), DNA stained with Hoechst 33342 dye (Hoe), GFP fluorescence and a merged image showing the respective channels in the indicated colours are shown from left to right, adjustment of the display settings brightness (br), contrast (co) and gamma correction (y) reveals plastidic nucleoids (orange arrows) in the Hoe chanel, maximum intensity z-projections from 12 slices of a stack with 2,9 µm depth, GFP is imported into the plastid and equally distributed within the stroma. c: Localisation of RecA:GFP, plastidic nucleoids can be recognised in the Hoe chanel and two phenotypes can be distinguished in the upper and the lower row (one nucleoid vs several nucleoids, orange arrows), maximum intensity z-projections from 18 slices of a stack with 4,4 µm depth (upper row) and from 12 slices of a stack with 2,9 µm depth (lower row), GFP selectively labels these plastidic nucleoids.

fluorescence of DNA stained with Hoechst 33342, but nuclear DNA which showed the brightest fluorescence in the Hoechst 33342 stain was not labeled with GFP (figure 2 c). Remarkably two different phenotypes occurred, displaying either one big nucleoid (figure 2 c, upper row) or several smaller nucleoids randomly distributed in proximity to the chloroplast's inner membrane (figure 2 c, lower row).

For labelling of the mitochondrial nucleoids a fusion protein of mature RecA and eGFP was N-terminally fused with the presequence of the mitochondrial phosphoglycerate kinase (Pgk) to provide the import of the fusion protein into the mitochondrial compartment (prePgk:RecA:GFP, figure 1). The PGK-presequence only was fused to eGFP for verification of correct mitochondrial import of the constructs (prePgk:GFP, figure 1). The labelling of mitochondrial nucleoids revealed only one spot of GFP fluorescence in the mitochondrial network of the diatom for all analysed transformants. Until now we could not show the presence of DNA in these spots, probably because the amount of mitochondrial DNA is only 0.25 times the amount of chloroplast DNA (Table 4). The correct localisation of the PrePgk:RecA:GFP fusion protein was verified via fluorescence microscopy of the PrePgk:GFP transformants which clearly displayed a green fluorescing mitochondrial network wrapped around the chloroplast (data not shown).

Quantification of plastidic nucleoids

To determine the average number of nucleoids, the chloroplast RecA:GFP transformants were analysed via fluorescence microscopy. While focusing through the chloroplast the fluorescing nucleoids have been counted for 313 individual diatoms. The ratio of cells containing only one big nucleoid (157 cells) to cells with several small nucleoids (156

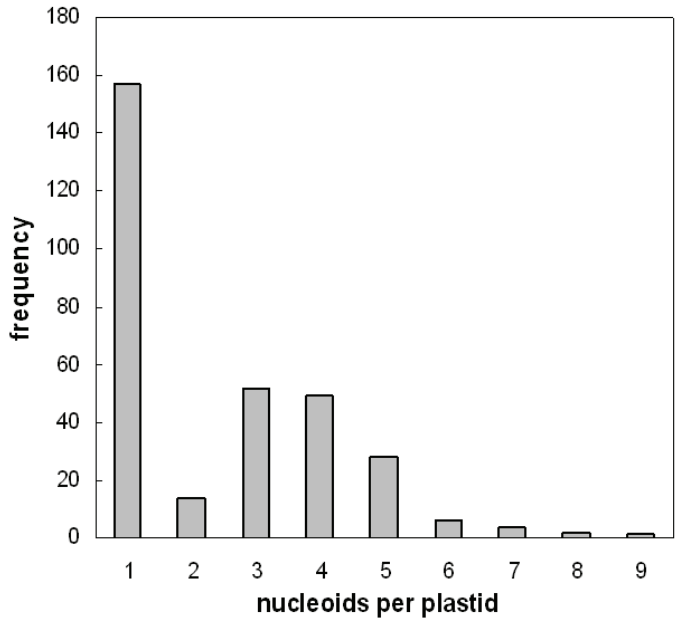


Figure 3: Number of nucleoids per plastid. Two major phenotypes could be distinguished, containing either one large or several small nucleoids per plastid. The frequencies of the different nucleoid counts observed in the second phenotype followed a Poisson-distribution with a maximum at 3.9.

cells) was 1.01. On average plastids contain 3.90 nucleoids, the standard deviation is 1.274. The maximum number of nucleoids per chloroplast was 9. However in 82.2% of all observations the chloroplasts contained 3 to 5 nucleoids as shown in figure 3. An average of 3.9 nucleoids in the cells with several plastidic nucleoid would mean an average of ~26 genome copies per plastidic nucleoid.

Real-time PCR quantification of organellar genomes

The heterocontophyte *Phaeodactylum tricornutum* harbours three different genome types, the diploid nuclear genome and the polyploid organellar genomes located in the chloroplast and mitochondrial network. In order to determine the genome copy numbers via relative quantification, short sequences of genes specific for the three individual genome types were amplified during RT-qPCR. Different sets of primers have been derived and tested for best suitability in RT-qPCR experiments as described in Materials and Methods. The primer pairs chosen for the experiments amplify the nuclear encoded gene dde, the chloroplast encoded psbA gene and the chondriome gene cox1. In order to cover possible variations generated by either the PCR reactions themselves or by differences in DNA extractions, all reactions have been performed in replicates using templates from three independent DNA extractions originating from three independently grown cultures. The amplification efficiencies of the three genome specific primers were determined previous to the relative quantification of the different genomes in a first RT-PCR experiment. Therefore the DNA extractions have been pooled as recommended (Bubner *et al.* 2004) and diluted in 5 serial 10 fold dilution steps. Efficiencies were calculated from the slopes of Ct/log dilution plots according to equation [3]. The three amplicons amplify with similar efficiencies of 2.29 (dde genes), 2.27 (psbA) and 2.04 (cox1) as shown in figure 4.

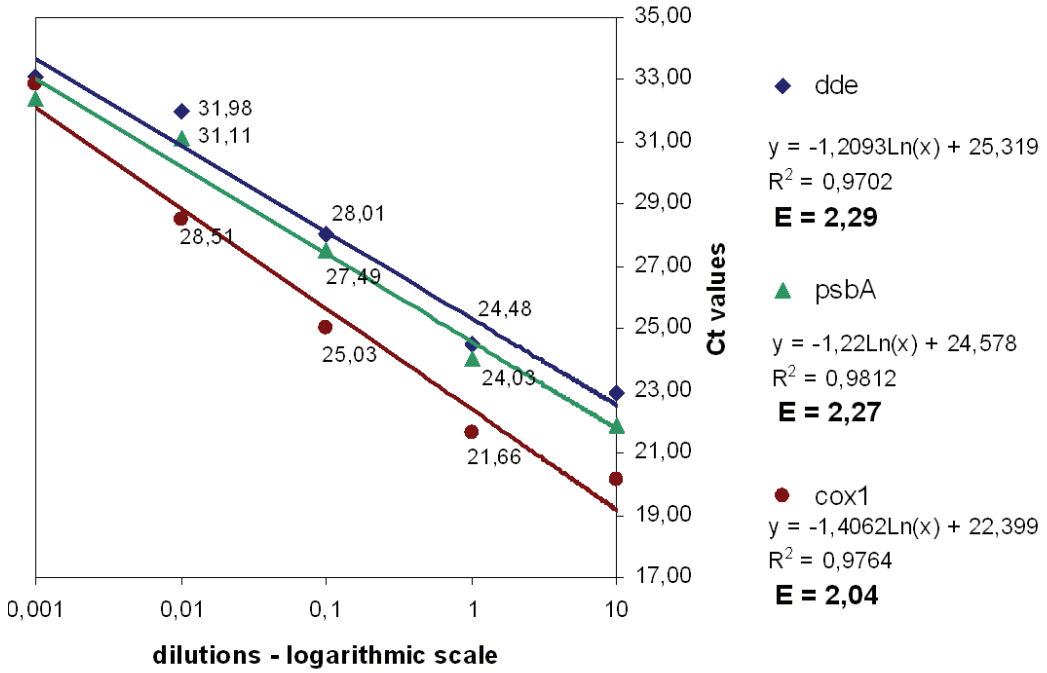


Figure 4: Calculation of different PCR efficiencies for the RT-qPCR primer pairs

For determining genome ploidies RT-qPCR was performed, the amount of psbA or cox1 copies have been quantified relatively to the nuclear calibrator dde. Each PCR reaction was performed in six replicates per template and per run. All runs have been repeated independently twice. The cycle threshold (Ct) values obtained from RT-qPCR vary only slightly within the six replicates, standard deviations were usually below 0.1 cycles, in all cases below 0.17 cycles. (table 1). Differences in average Ct values obtained by two independent RT-qPCR experiments are always below 0.3 cycles as suggested by Bubner & Baldwin(Bubner and Baldwin 2004) thus confirming the reproducibility of results. The relative quantities of the three genomes were calculated from the respective Ct values, taking into account the PCR efficiencies of the genome specific primers according to equation [2] (see Materials). Ct values and relative quantities are shown in table 1. In order to obtain absolute genome copy numbers the relative quantities obtained for the

Template - Run		Nucleus	Plastid	Mito
I-1	n	*6*	*6*	*6*
	Ct mean	*24,041*	*19,076*	*20,241*
	standard deviation	*0,066*	*0,040*	*0,048*
	CV [%]	*0,275*	*0,210*	*0,237*
	RQ	*1,000*	*58,569*	*15,017*
I-2	n	*6,000*	*6,000*	*6,000*
	Ct mean	*24,270*	*19,209*	*20,418*
	standard deviation	*0,056*	*0,024*	*0,066*
	CV [%]	*0,231*	*0,125*	*0,323*
	RQ	*1,000*	*63,365*	*15,585*
I mean	RQ	*1,000*	*60,967*	*15,301*
II-1	n	*6*	*6*	*6*
	Ct mean	*29,743*	*24,968*	*26,087*
	standard deviation	*0,014*	*0,007*	*0,136*
	CV [%]	*0,047*	*0,028*	*0,521*
	RQ	*1,000*	*50,121*	*13,552*
II-2	n	*6,000*	*6,000*	*6,000*
	Ct mean	*30,010*	*25,201*	*26,304*
	standard deviation	*0,058*	*0,055*	*0,166*
	CV [%]	*0,193*	*0,218*	*0,631*
	RQ	*1,000*	*51,538*	*14,044*
II mean	RQ	*1,000*	*50,830*	*13,798*
III-1	n	*6*	*6*	*6*
	Ct mean	*27,802*	*23,343*	*24,018*
	standard deviation	*0,087*	*0,018*	*0,017*
	CV [%]	*0,313*	*0,077*	*0,071*
	RQ	*1,000*	*38,683*	*14,847*
III-2	n	*6,000*	*6,000*	*6,000*
	Ct mean	*28,052*	*23,560*	*24,214*
	standard deviation	*0,058*	*0,026*	*0,013*
	CV [%]	*0,207*	*0,110*	*0,054*
	RQ	*1,000*	*39,744*	*15,430*
III mean	RQ	*1,000*	*39,213*	*15,138*

Table 1: RT-qPCR results. Quantification of relative DNA amounts in total DNA extractions from three independent cultures. Listed are n: number of replicates per run, Ct: Cycle threashold and standard deviation, CV: coefficient of variance, and the RQ: relative quantity.

psbA and cox1 genes have to be multiplied with the number of dde copies in the nucleus. The dde primers amplify the nuclear encoded diadinoxanthin de-epoxidase gene which serves as calibrator. Since the nuclear genome is diploid it contains altogether 2 copies of the calibrator gene. Hence the detected relative quantities of plastome specific psbA and of the chondriome specific cox1 gene are therefore multiplied by 2. Analysis of all data obtained from different RT-qPCR runs using three independent templates reveal an average plastome copy number of 100.7, the average chondriome copy number is 29.5 (table 2, figure 5).

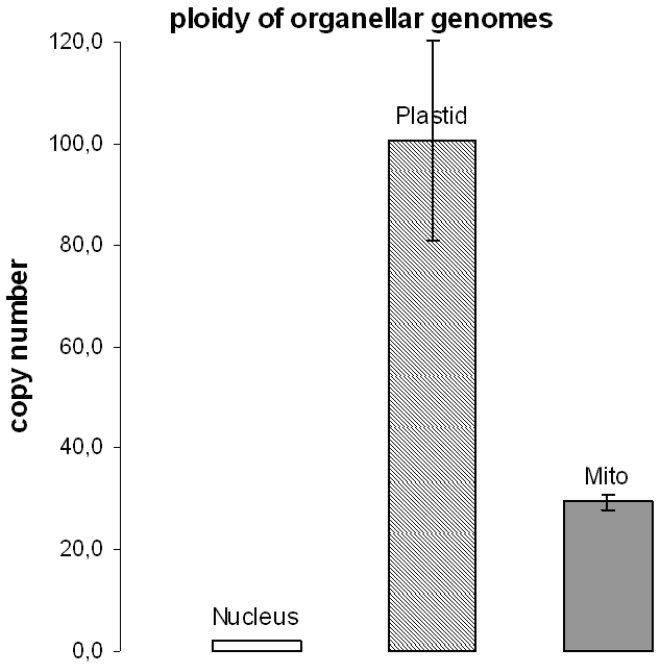


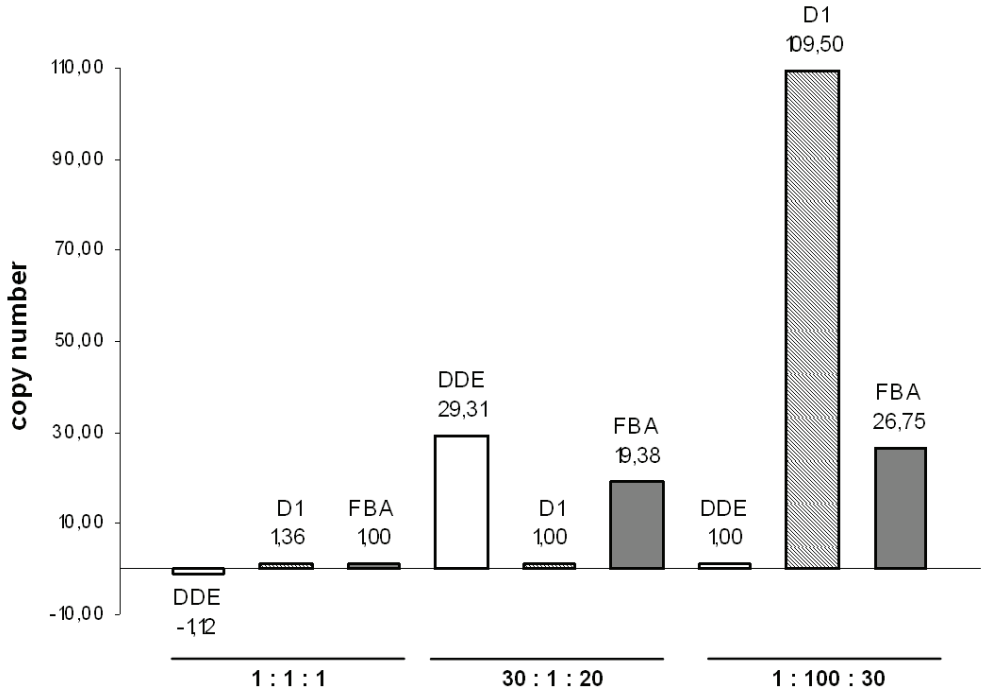
Table 2 and Figure 5: Ploidies of organellar genomes obtained by RT-qPCR based quantification relative to the diploid nuclear genome. The copy numbers represent the average value calculated from the analysis of three independent cultures. It is remarkable that plastid ploidies vary within different cultures while chondriome ploidies remain constant

Genomes per	**Nucleus**	**Plastid**	**Mito.**
average Copy No.:	*2,0*	*100,7*	*29,5*
standard deviation	*0,0*	*19,7*	*1,6*
CV [%]	*0,0*	*19,6*	*5,4*

While the chondriome copy numbers obtained from different runs and different templates vary only slightly (coefficient of variation = 5.4%) the plastome copy numbers vary by 19.6%. PsbA replicates containing the same template produced very similar results, hence relative quantities (RQ) obtained from two independent RT-qPCR runs differed by only 1.4 or 1 copies for template II and III respectively. Only for template I the relative psbA quantities obtained from two runs differed from each other by ~4.8 copies (table 2). Since plastome copy numbers remained stable during independent quantification experiments using the same template, the variance in the plastome copy numbers must be attributed to differences emanating from the three independent templates. This variance seems to be rather a culture and chloroplast specific feature than due to unsteadiness in preparation of different templates. The constancy in chondriome copy numbers, even in different templates, provides evidence for this assumption (figure 5).

Artificial templates

To verify both, the obtained ploidy levels of organellar genomes in *Phaeodactylum tricornutum* and the accuracy of the presented method, a setup was developed in which artificially prepared templates have been analysed. These artificial templates consist of plasmid mixtures simulating the different genomes within total DNA extractions. The composition of the different genomes was virtually represented by the relative amounts of the respective plasmids in the mixture. Three different plasmids have been used containing either the dde sequence representing the nuclear genome, the psbA sequence representing the plastome, or an fba gene representing a chondriome encoded sequence. The plasmids have been mixed in the same ratio also found for the genomes in *Phaeodactylum tricornutum* (1:100:30). Further two more artificial templates have been prepared containing the dde-plasmid, the psbA-plasmid and the fba-plasmid in the ratios 1:1:1 and 30:1:20. The artificial templates were analysed via RT-qPCR, each reaction was performed in six replicates. Again relative quantities were calculated using equation [2], for the 1:1:1 experiment fba was used as calibrator, while for the 30:1:20 experiment psbA served as calibrator. For the experiment which simulates the conditions found in Phaeodactylum, again dde was used as calibrator. The results of all three experiments (see table 3, figure 6) reflect the plasmid compositions of the respective artificial templates which supports the functionality of the relative genome quantification method. Minor aberrations from expected ratios are most likely due to inaccuracies in manual adjustment of the artificial template components. The replicate standard deviations are below 0.02 cycles, demonstrating the very high reproducibility of results and the accuracy of the RT-qPCR setup.



dde* : *psbA* : *fba		***dde***	***psbA***	***fba***
1 : 1 : 1	n	*6*	*6*	*6*
	mean	*12,577*	*17,130*	*13,234*
	standard deviation	*0,015*	*0,002*	*0,019*
	CV [%]	*0,119*	*0,012*	*0,144*
	RQ	***-1,119***	***1,359***	***1,000***
30 : 1 : 20	n	*6*	*6*	*6*
	mean	*12,577*	*17,130*	*13,234*
	standard deviation	*0,015*	*0,002*	*0,019*
	CV [%]	*0,119*	*0,012*	*0,144*
	RQ	***29,313***	***1,000***	***19,377***
1 : 100 : 30	n	*6*	*6*	*6*
	mean	*18,213*	*11,801*	*13,893*
	standard deviation	*0,018*	*0,006*	*0,060*
	CV [%]	*0,099*	*0,051*	*0,432*
	RQ	***1,000***	***109,502***	***26,754***

Table 3 and Figure 6: Virtual ploidies of artificial templates obtained by RT-qPCR based quantification. The obtined copy numbers confirm the reliability of the method.

Fluorometric DNA detection

To evaluate the findings of the copy number determination we calculated the distribution of DNA within a *Phaeodactylum tricornutum* cell. Genome sizes of the organelles were multiplied with the respective copy numbers determined in this study and via the average molecular weight of double stranded DNA (660 g/mol). The average DNA amount per cell was calculated to be 0.073 pg/cell (Table 1). The ratio of nuclear DNA was calculated to be 78.4 % on average with a range from 75.5 % to 90.2 % nuclear DNA (Table 4). For

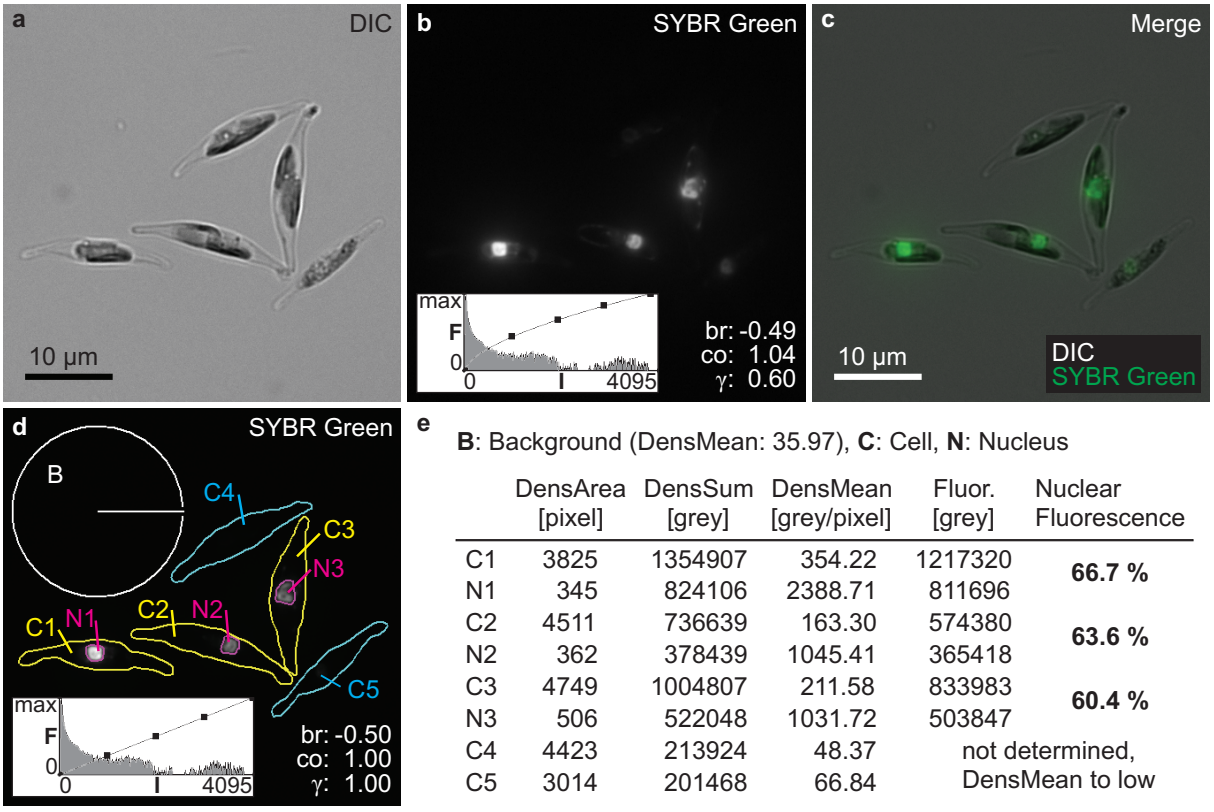
		average ratio			maximal ratio			minimal ratio		
	Genome size [bp]	ploidy	size [bp/cell]	weight [pg/cell]	ploidy	size [bp/cell]	weight [pg/cell]	ploidy	size [bp/cell]	weight [pg/cell]
Nucleus	26000000	2	52000000	0.057	4	104000000	0.1144	2	52000000	0.057
Chloroplast	117369	101	11854269	0.013	77	9037413	0.0099	122	14319018	0.016
Mitochondrium	83793	30	2513790	0.003	27	2262411	0.0025	31	2597583	0.003
Sum	26201162		66368059	0.073		115299824	0.1268		68916601	0.076
% of nuclear DNA			**78.4**			**90.2**			**75.5**	

Table 4: Distribution of DNA within the Phaeodactylum tricornutum cell. Genome sizes of the organelles were multiplied with the respective copy numbers determined in this study and the DNA ammount per cell was calculated via the average molecular weight of double stranded DNA (660 g/mol). For the average ratio a diploid set of nuclear DNA was assumed and the average copy numbers of the plastome and chondriome were used, the maximal ratio was calculated using the highest theoretical nuclear DNA ammount (4 haploid sets in a nucleus shortly before division) and the minimal determined copy numbers of the plastome and chondriome, the minimal ratio was calculated assuming a diploid set of nuclear DNA (pennates do not form free living haploid cells) and the maximal determined copy numbers of the plastome and chondriome were used.

the average ratio of nuclear to total DNA of the cell, a diploid set of nuclear DNA was assumed and the average copy numbers of the plastid genome and the mitochondrial genome were used, the maximal ratio was calculated using the highest theoretical nuclear DNA amount (4 haploid sets in a nucleus shortly before division) and the minimal determined copy numbers of the plastid genome and the mitochondrial genome, the minimal ratio was calculated assuming a diploid set of nuclear DNA (pennates do not form free living haploid cells) and the maximal determined copy numbers of the plastid genome and the mitochondrial genome.

To confirm this ratio, we analysed fluorescence images of a *P. tricornutum* liquid culture ($1.39 \cdot 10^7$ cells$\cdot$ml^{-1}) stained with SYBR Green I (0.5 µl SYBR Green I stock solution into 1.5 ml liquid culture). Bright fluorescent nuclei could be observed in most of the cells and with adjustments of the display settings brightness (br), contrast (co) and gamma correction (γ) also extranuclear DNA could be visualised, mostly associated with the plastids. (Figure 7). In controll images of unstained *P. tricornutum* cells, equal DensMean values were measured for outlines of cells and backgrounds of the respective images (data not shown), autofluorescence of the cells was therefore efficiently excluded from the measurements. To calculate the ratio of nuclear to total fluorescence of individual cells ($Fluor._{nucleus}/Fluor._{cell}$) outlines were drawn manualy into the image. Fluor. was

calculated by substraction of the background from the DensSum of the outline if the DensMean exceeded the DensMean of the Background at least fourfold (Figure 7). The originally acquired images were used for the measurements without further adjustment of display settings. 45 cells showed nuclear fluorescences from 53.5 % to 75.5 % with an average of 66.3 % (standard deviation 4.6 %) (Figure 8).



e

B: Background (DensMean: 35.97), **C**: Cell, **N**: Nucleus

	DensArea [pixel]	DensSum [grey]	DensMean [grey/pixel]	Fluor. [grey]	Nuclear Fluorescence
C1	3825	1354907	354.22	1217320	**66.7 %**
N1	345	824106	2388.71	811696	
C2	4511	736639	163.30	574380	**63.6 %**
N2	362	378439	1045.41	365418	
C3	4749	1004807	211.58	833983	**60.4 %**
N3	506	522048	1031.72	503847	
C4	4423	213924	48.37	not determined, DensMean to low	
C5	3014	201468	66.84		

Figure 7: Relative fluorometric DNA quantification. **a**: Transmitted light view of *Phaeodactylum tricornutum* cells stained with SYBR Green I. **b**: SYBR Green I fluorescence of the same specimen with bright fluorescent nuclei, brightness (br), contrast (co) and gamma correction (γ) adjustment reveals also fluorescence of non nuclear DNA, the distribution of grey values (histogramm inserted at lower left) shows no overexposure of the image, a prerequisite for quantitative analyses. **c:** Merged image of a and b showing that most non nuclear DNA fluorescence is associated with the plastids. **d**: Outlines drawn manualy into the image to calculate the ratio of nuclear to total fluorescence of individual cells, measurements are done on the originally acquired image (see histogramm inserted at lower left and brightness, contrast and gamma correction settings). **e**: Densitometric values measured from d, fluorescence (Fluor.) was calculated by substraction of the background from the densitometric sum (DensSum) of the outline if the densitometric mean (DensMean) exceeded the DensMean of the Background at least fourfold, see text for details.

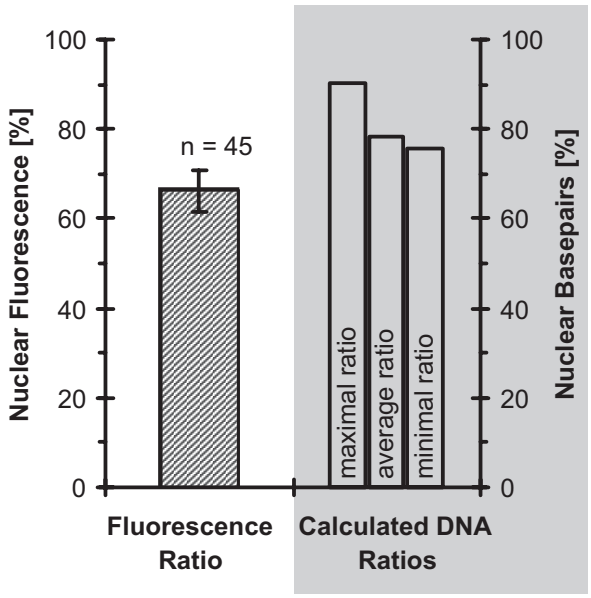


Figure 8: Ratio of nuclear fluorescence to total flourescence of Phaeodactylum tricornutum cells stained with SYBR Green I. The ratio of nuclear fluorescence is compared to ratios of nuclear basepairs calculated from the genome sizes and our determination of genome copy numbers. Nuclear fluorescence ratios were calculated for 45 cells, error bar, standard deviation.

II.4.5 Discussion

In vivo labelling of organellar nucleoids

In this study we presented a protocol for labelling specifically and *in vivo* organellar nucleoids in the diatom *Phaeodactylum tricornutum*. The visualisation of nucleoids in microalgae as well as the targeted and independent labelling of both, plastid and mitochondrial nucleoids in the same organism is unique so far. In contrast to previous studies we designed a fusion protein featuring organelle-independent DNA binding capabilities. Organelle-specific targeting was achieved via N-terminal fusion of the DNA binding protein with an appropriate presequence. The RecA protein was utilised to mediate DNA binding. RecA effects homologous recombination which is a crucial underlying mechanism for DNA repair and proper cellular replication, furthermore RecA is believed to be involved in genome evolution via mediating events such as gene conversion.

Labelling of organellar nucleoids was achieved via targeting RecA-eGFP fusion proteins to the compartment of interest. Therefore the fusion protein was provided with appropriate presequences. The specific import of the fusion proteins into either the chloroplast or the mitochondrial network was confirmed by transforming the cells with preRecA:GFP or PrePgk:GFP which contain the respective presequences fused to eGFP. The observed eGFP-fluorescence in the RecA:GFP and PrePgk:RecA:GFP transformants could be colocalised with fluorescing spots which resulted from DNA staining using the fluorophore Hoechst 33342, thus confirming the eGFP-labelled spots within the organelles to be rich of DNA which supports their identification as nucleoids. Taken together these results strongly suggest the capability of targeted RecA-eGFP fusion proteins to selectively label plastid or mitochondrial nucleoids.

Older techniques often deployed fluorophores such as DAPI, SYBR Green or ethidium bromide which have been widely used to stain genomes and even to study organellar nucleoids via fluorescence microscopy (Coleman 1979; Williamson and Fennell 1979; Stevens 1981). However, besides suffering from generated background fluorescence, these methods cannot be directed specifically to a genome of interest. Fluorophores rather stain all nucleic acids in the cell, thus generating errors when quantifying nucleoids. In more recent studies nucleoids have been already successfully labelled using DNA binding proteins such as PEND (Sato *et al.* 1993; Sato and Ohta 2001), or Abf2 (Gorsich and Shaw 2004). However, these proteins are species and organelle-specific and not available in most plants, neither in microalgae (Terasawa and Sato 2005a). The system presented here intends to overcome these restrictions to provide access to both organelles via utilising a general DNA binding protein which is found ubiquitously in the

most species. Therefore the evolutionary conserved RecA protein (Edelmann and Kucherlapati 1996; Seitz and Kowalczykowski 2000; Wyman and Kanaar 2004)was used to mediate DNA binding. We suggest that RecA-eGFP fusion proteins might generally contribute to genome sciences by representing a convenient alternative to label organellar nucleoids in a wide range of species.

Labelling of *Phaeodactylum tricornutum*'s chloroplast nucleoids revealed the presence of two different equally distributed phenotypes. The chloroplasts contained either one large nucleoid or several small nucleoids. In accordance with previous findings the presence of two different phenotypes clearly demonstrates significant changes in morphology and in the dynamics of plastid nucleoids. The cause of these changes remains elusive, however it appears plausible that these changes are linked to cell cycle and organellar propagation (Sato *et al.* 2003; Dai *et al.* 2005). In addition to these findings we further experienced changes in nucleoid arrangements as a reaction to certain stress conditions (data not shown).

RecA:GFP fuson protein targetet to the mitochondria leads to one fluorescence spot within the mitochondrial network of the cells. Assuming the same genome copy number in mitochondrial nucleoids than found in plastids, it appears possible that the complete mitochondrial DNA content is concentrated within a single nucleoid. This finding is surprising insofar as that the whole mitochondrial network of *Phaeodactylum tricornutum* usually consists of two to three major protuberances and this would include that some parts of the mitochondrial network are temporarily DNA depleted.

Relative quantification of organellar genomes

Initially ploidies and amounts of segregating units within organelles were determined via methods which based on autoradiography (Gibbs 1968, 1979), reassociation kinetics of ^{125}I-labelled or UV spectrometrically detected dsDNA (Gelb *et al.* 1971; Rawson and Boerma 1976; Ersland *et al.* 1981; Ersland and Cattolico 1981), or fluorophore staining of nucleic acids with subsequent fluorescence quantification (Misumi *et al.* 2001). These techniques involve complex experimental setups which are time consuming, thus lacking high-throughput capabilities.

The protocol developed here was designed to meet the needs for accuracy as well as for the capability to analyse a large number of samples at the same time. High throughput characteristics allow monitoring of different cultures or even species in order to study influences of genome contents on organellar development and propagation, or environmental influences on the genome contents. Our protocol was designed to determine copy numbers of organellar genomes via RT-qPCR. The absolute ploidies were quantified relatively to the diploid nuclear genome. Since quantification by real-time PCR

is an amplification-based procedure small fluctuations in the starting conditions of a PCR reaction lead to large fluctuation of the product amount which is expressed in Ct values. Our protocol relies on three major steps to keep amplification-based variations low: (i) All RT-qPCR primer pairs were designed to provide equal conditions regarding annealing temperature, amplicon length and amplicon T_M, further hetero-dimer or self-dimer formation was avoided. The predicted conditions have been verified by two independent software tools. In the next step of control the amplicons generated by the primer pairs were analysed via gel-electrophoresis and by melting curve analysis under RT-qPCR conditions. Primers pairs producing detectable amounts of either dimers or unspecific amplicons have been discarded. (ii) Different primer pairs and templates lead to different amplification efficiencies thus altering Ct values. For determining exact copy numbers via relative quantification an efficiency corrected mathematical model was applied. (iii) In our approach the three different genomes in one distinct DNA extraction are quantified relatively. Therefore all reactions and replicates of one experiment contain exactly the same template, neither template concentrations need to be equally adjusted, nor an endogenous control is required to normalize different template concentrations. Since the Ct total variation is the sum of the variation from sample replicates and from the endogenous control replicates (Bubner and Baldwin 2004) the possibility to neglect the latter decreases the standard deviations of the genome quantities. The methodology was shown to be sensitive, indicated by very low Ct standard deviations which were in most cases below 0.1 cycles, but always below 0.17 cycles. The reliability of the genome quantification protocol was further demonstrated by the simulation experiments in which artificial templates were analysed. Given that manual errors cannot be avoided during preparation of the artificial templates, the ratios of virtual genomes in these templates (1:1:1; 30:1:20; 1:100:30) could be determined surprisingly accurate (-1.12:1.36:1; 29.31:1:19.38; 1:109.50:26.75). It is likely that the small deviations from the given ratios can be attributed rather to the manual adjustments of the artificial templates than to inaccuracies of the quantification protocol. Finally the results for *Phaeodactylum tricornutum* generated by this protocol provide further support for the reliability of the method: For the three different cultures analysed in this study the plastome content per chloroplast was found to vary among different cultures by 19.6% due to yet unknown circumstances. Supporting our findings, variations in plastome contents during chloroplast development have been already detected for green algae (Chelm *et al.* 1977) and higher plants (Miyamura *et al.* 1986; Baumgartner *et al.* 1989). Plastome content variations have also been proposed to occur due to cell proliferation and aging in Chlamydomonas (Misumi *et al.* 2001) or prior to leaf senescence in higher plants (Sodmergen *et al.* 1989; Sodmergen *et al.* 1991; Inada *et al.* 1998). However, strongest support for both, the detected plastome content variations and the reliability of the genome quantification protocol is provided by the stable chondriome copy numbers

detected for the *P. tricornutum* mitochondrial network. In all analysed cultures the copy numbers of the chondriomes varied by only 5,4% which is the equivalent to 1.89 copies. This coefficient of variance is mainly due to the template of culture II (see table X) displaying the highest Ct value standard deviations (SD = 0.136, respectively 0.166) which even might be reduced by increasing the applied template concentration and therefore the obtained Ct values (Bubner and Baldwin 2004).

The knowledge of the genome sizes and the respective copy numbers enables us to calculate the total DNA amount in a *P. tricornutum* cell (0.073 pg/cell). Previously determined total cellular DNA contents ranged from 0.0133pg/cell (Scala *et al.* 2002) and and 0.12 pg/cell (Darley 1968) to even 0.249 pg/cell (Veldhuis *et al.* 1997). Since this significant variance in published results can be hardly satisfying, we wanted to confirm our finding by applying a different method. Comparison of the observed ratio of nuclear fluorescence to the calculated ratio of nuclear DNA revealed similar results for both methods. The observed nuclear fluorescences were slightly lower than the calculated ratio of nuclear DNA. The systematic underestimation of nuclear DNA ratios by measuring nuclear fluorescence can be explained by different structures and densities of the different genomes, resulting in a weaker staining of the densely packed nuclear DNA. Moreover, the probability of self-absorption of excitation light and quenching of the fluorescence is higher in the intensely fluorescing nuclei. This finding confirms the copy numbers obtained with the RT-qPCR approach.

Conclusion

Combining labelling and quantification techniques provides potential support to genome sciences, especially to studies focusing on organelles. Both techniques together allow to determine organellar ploidies, the amount of nucleoids per organelle and to estimate the genome content per nucleoid. Furthermore estimation of genome sizes can be optimized by including copy numbers and DNA contents of the organelles. Being able to monitor these data for a set of cultures might increase our knowledge about organellar propagation and the effects of genome contents on cellular development. Furthermore little is known yet, about environmental effects on nucleoid localisation, and on the genome content of both, organelles and nuclei. Thus, the methods presented here might contribute to recent studies in photosynthetic and non-photosynthetic protists and higher plants (MacAlpine *et al.* 2000; Misumi *et al.* 2001; Gorsich and Shaw 2004; Oldenburg and Bendich 2004; Rowan *et al.* 2004; Terasawa and Sato 2005a, 2005b; Li *et al.* 2006).

III Appendices

References cited

Ajlani G, Meyer I, Vernotte C, Astier C (1989) Mutation in phenol-type herbicide resistance maps within the psbA gene in Synechocystis 6714. *FEBS Lett*. **246**: 207-210.

Alekseeva SA, Shevchenko NM, Kusaikin MI, Ponomorenko LP, Isakov VV, Zviagintseva TN, Likhoshvai EV (2005) [Polysaccharides of diatoms occurring in Lake Baikal]. *Prikl Biokhim Mikrobiol*. **41**: 213-219.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. **25**: 3389-3402.

Ambros V (2004) The functions of animal microRNAs. *Nature*. **431**: 350-355.

Apt KE, Kroth-Pancic PG, Grossman AR (1996) Stable nuclear transformation of the diatom Phaeodactylum tricornutum. *Mol Gen Genet*. **252**: 572-579.

Arsalane W, Rousseau B, Duval JC (1994) Influence of the pool size of the xanthophyll cycle on the effects of light stress in a diatom: competition between photoprotection and photoinhibition. *Photochemistry and Photobiology*. **60**: 237-243.

Avni A, Edelman M (1991) Direct selection for paternal inheritance of chloroplasts in sexual progeny of Nicotiana. *Mol Gen Genet*. **225**: 273-277.

Backert S, Nielsen BL, Börner T (1997) The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. *Trends Plant Sci*. **2**: 477-483.

Baena-Gonzalez E, Gray JC, Tyystjarvi E, Aro EM, Maenpaa P (2001) Abnormal regulation of photosynthetic electron transport in a chloroplast ycf9 inactivation mutant. *J Biol Chem*. **276**: 20795-20802.

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. **116**: 281-297.

Bateman JM, Purton S (2000) Tools for chloroplast transformation in Chlamydomonas: expression vectors and a new dominant selectable marker. *Mol Gen Genet*. **263**: 404-410.

Baulcombe DC (1996) RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants. *Plant Mol Biol*. **32**: 79-88.

Baumgartner BJ, Rapp JC, Mullet JE (1989) Plastid Transcription Activity and DNA Copy Number Increase Early in Barley Chloroplast Development. *Plant Physiol*. **89**: 1011-1018.

Bendich AJ (1987) Why do chloroplasts and mitochondria contain so many copies of their genome? *Bioessays*. **6**: 279-282.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol*. **340**: 783-795.

Bhattacharya D, Medlin LK (2004) Dating algal origin using molecular clock methods. *Protist*. **155**: 9-10.

Birky CW, Jr. (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci U S A*. **92**: 11331-11338.

Birky CW, Jr., Walsh JB (1992) Biased gene conversion, copy number, and apparent mutation rate differences within chloroplast and bacterial genomes. *Genetics*. **130**: 677-683.

Bock R (2001) Transgenic plastids in basic research and plant biotechnology. *J Mol Biol*. **312**: 425-438.

Bock R, Kossel H, Maliga P (1994) Introduction of a heterologous editing site into the tobacco plastid genome: the lack of RNA editing leads to a mutant phenotype. *Embo J*. **13**: 4623-4628.

Boiteux S, Gellon L, Guibourt N (2002) Repair of 8-oxoguanine in Saccharomyces cerevisiae: interplay of DNA repair and replication mechanisms. *Free Radic Biol Med*. **32**: 1244-1253.

Borowitzka MA, Volcani BE (1978) The polymorphic diatom Phaeodactylum tricornutum: ultrastructure of its morphotypes. *Journal of Phycology*. 10–21.

Borowitzka MA, Chiappino ML, Volcani BE (1977) Ultrastructure of a chainforming diatom Phaeodactylum tricornutum. *Journal of Phycology*. 162–170.

Bouck GB (1965) Fine structure and organelle association in brown algae. *J Cell Biol*. **26**: 523-537.

Boynton JE, Gillham NW, Harris EH, Hosler JP, Johnson AM, Jones AR, Randolph-Anderson BL, Robertson D, Klein TM, Shark KB, et al. (1988) Chloroplast transformation in Chlamydomonas with high velocity microprojectiles. *Science*. **240**: 1534-1538.

Bridges BA (2001) Hypermutation in bacteria and other cellular systems. *Philos Trans R Soc Lond B Biol Sci*. **356**: 29-39.

Bubner B, Baldwin IT (2004) Use of real-time PCR for determining copy number and zygosity in transgenic plants. *Plant Cell Rep*. **23**: 263-271.

Bubner B, Gase K, Baldwin IT (2004) Two-fold differences are the detection limit for determining transgene copy numbers in plants by real-time PCR. *BMC Biotechnol*. **4**: 14.

Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol*. **25**: 169-193.

Cairns J, Foster PL (1991) Adaptive reversion of a frameshift mutation in Escherichia coli. *Genetics*. **128**: 695-701.

Cambareri EB, Jensen BC, Schabtach E, Selker EU (1989) Repeat-induced G-C to A-T mutations in Neurospora. *Science*. **244**: 1571-1575.

Carrer H, Hockenberry TN, Svab Z, Maliga P (1993) Kanamycin resistance as a selectable marker for plastid transformation in tobacco. *Mol Gen Genet*. **241**: 49-56.

Cavalier-Smith T (1998) A revised six-kingdom system of life. *Biol Rev Camb Philos Soc*. **73**: 203-266.

Cavalier-Smith T (2003) Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). *Philos Trans R Soc Lond B Biol Sci*. **358**: 109-133; discussion 133-104.

Cerutti H, Osman M, Grandoni P, Jagendorf AT (1992) A homolog of Escherichia coli RecA protein in plastids of higher plants. *Proc Natl Acad Sci U S A*. **89**: 8068-8072.

Chelm BK, Hoben PJ, Hallick RB (1977) Cellular content of chloroplast DNA and chloroplast ribosomal RNA genes in Euglena gracilis during chloroplast development. *Biochemistry*. **16**: 782-786.

Chiovitti A, Molino P, Crawford SA, Teng R, Spurck T, Wetherbee R (2004) The glucans extracted with warm water from diatoms are mainly derived from intracellular chrysolaminaran and not extracellular polysaccharides. *European Journal of Phycology*. **39**: 117 - 128.

Cogoni C, Macino G (1999) Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. *Nature*. **399**: 166-169.

Coleman AW (1979) Use of the fluorochrome 4'6-diamidino-2-phenylindole in genetic and developmental studies of chloroplast DNA. *J Cell Biol*. **82**: 299-305.

Coleman AW, Nerozzi AM (1999) Temporal and spatial coordination of cells with their plastid component. *Int Rev Cytol*. **193**: 125-164.

Cox MM (2003) The bacterial RecA protein as a motor protein. *Annu Rev Microbiol*. **57**: 551-577.

Dai H, Lo YS, Litvinchuk A, Wang YT, Jane WN, Hsiao LJ, Chiang KS (2005) Structural and functional characterizations of mung bean mitochondrial nucleoids. *Nucleic Acids Res*. **33**: 4725-4739.

Dalla Chiesa M, Friso G, Deak Z, Vass I, Barber J, Nixon PJ (1997) Reduced turnover of the D1 polypeptide and photoactivation of electron transfer in novel herbicide resistant mutants of Synechocystis sp. PCC 6803. *Eur J Biochem*. **248**: 731-740.

Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC (2000) An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell*. **101**: 543-553.

Dany AL, Tissier A (2001) A functional OGG1 homologue from Arabidopsis thaliana. *Mol Genet Genomics*. **265**: 293-301.

Darley WM (1968) Deoxyribonucleic acid content of the three cell types of Phaeodactylum tricornutum Bohlin. *Journal of Phycology*. 219-220.

Delwiche CF (1999) Tracing the Thread of Plastid Diversity through the Tapestry of Life. *Am Nat*. **154**: S164-S177.

Delwiche CF, Palmer JD (1997) The origin of plastids and their spread via secondary endosymbiosis. *Plant Syst Evol*. **11**: 53-86.

Demmig-Adams B, Adams WW (1996) The role of xanthophyll cycle carotenoids in the protection of photosynthesis. *Trends in Plant Science*. **1**: 21.

Demmig-Adams B, Gilmore AM, Adams WW (1996) In vivo functions of carotenoids in higher plants. *FASEB Journal*. **10**: 403–412.

Doetsch NA, Favreau MR, Kuscuoglu N, Thompson MD, Hallick RB (2001) Chloroplast transformation in Euglena gracilis: splicing of a group III twintron transcribed from a transgenic psbK operon. *Curr Genet*. **39**: 49-60.

Dougherty WG, Parks TD (1995) Transgenes and gene suppression: telling us something new? *Curr Opin Cell Biol*. **7**: 399-405.

Doyle JJ, Doyle JL (1990) A rapid total DNA preparation procedure for fresh plant tissue. *Focus*. 13-15.

Drum RW, Gordon R (2003) Star Trek replicators and diatom nanotechnology. *Trends Biotechnol*. **21**: 325-328.

Dufourmantel N, Pelissier B, Garcon F, Peltier G, Ferullo JM, Tissot G (2004) Generation of fertile transplastomic soybean. *Plant Mol Biol*. **55**: 479-489.

Dunahay TG, Jarvis EE, Roessler PG (1995) Genetic transformation of the diatoms Cyclotella cryptica and Navicula saprophila. *Journal of Phycology*. 1004–1012.

Edelmann W, Kucherlapati R (1996) Role of recombination enzymes in mammalian cell survival. *Proc Natl Acad Sci U S A*. **93**: 6225-6227.

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci*. **8**: 978-984.

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol*. **300**: 1005-1016.

Ersland DR, Cattolico RA (1981) Nuclear deoxyribonucleic acid characterization of the marine chromophyte Olisthodiscus luteus. *Biochemistry*. **20**: 6886-6893.

Ersland DR, Aldrich J, Cattolico RA (1981) Kinetic Complexity, Homogeneity, and Copy Number of Chloroplast DNA from the Marine Alga Olisthodiscus luteus. *Plant Physiol*. **68**: 1468-1473.

Eskling M, Arvidsson P-O, Åkerlund H-E (1997) The xanthophyll cycle, its regulation and components. *Physiologia Plantarum*. **100**: 806-816.

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

Falciatore A, Casotti R, Leblanc C, Abrescia C, Bowler C (1999) Transformation of Nonselectable Reporter Genes in Marine Diatoms. *Mar Biotechnol (NY)*. **1**: 239-251.

Falkowski PG, Barber RT, Smetacek VV (1998) Biogeochemical Controls and Feedbacks on Ocean Primary Production. *Science*. **281**: 200-207.

Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJ (2004) The evolution of modern eukaryotic phytoplankton. *Science*. **305**: 354-360.

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*. **281**: 237-240.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*. **391**: 806-811.

Fischer H, Robl I, Sumper M, Kröger N (1999) Targeting and covalent modification of cell wall and membrane proteins heterologously expressed in the diatom Cylindrotheca fusiformis (Bacillariophyceae). *Journal of Phycology*. 113–120.

Foster PL (1999) Mechanisms of stationary phase mutation: a decade of adaptive mutation. *Annu Rev Genet*. **33**: 57-88.

Foster PL (2004) Adaptive mutation in Escherichia coli. *J Bacteriol*. **186**: 4846-4852.

Foster PL (2005) Stress responses and genetic variation in bacteria. *Mutat Res*. **569**: 3-11.

Foster PL, Trimarchi JM (1995) Adaptive reversion of an episomal frameshift mutation in Escherichia coli requires conjugal functions but not actual conjugation. *Proc Natl Acad Sci U S A*. **92**: 5487-5490.

Foster PL, Trimarchi JM, Maurer RA (1996) Two enzymes, both of which process recombination intermediates, have opposite effects on adaptive mutation in Escherichia coli. *Genetics*. **142**: 25-37.

Fowler RG, White SJ, Koyama C, Moore SC, Dunn RL, Schaaper RM (2003) Interactions among the Escherichia coli mutT, mutM, and mutY damage prevention pathways. *DNA Repair (Amst)*. **2**: 159-173.

Franklin S, Ngo B, Efuet E, Mayfield SP (2002) Development of a GFP reporter gene for Chlamydomonas reinhardtii chloroplast. *Plant J*. **30**: 733-744.

Friedberg EC (2002) The intersection between the birth of molecular biology and the discovery of DNA repair. *DNA Repair (Amst)*. **1**: 855-867.

Friedberg EC, Walker GC, Siede W (1995) DNA Repair and Mutagenesis. 698 pp.

Friedberg EC, Wagner R, Radman M (2002) Specialized DNA polymerases, cellular survival, and the genesis of mutations. *Science*. **296**: 1627-1630.

Friedberg EC, Lehmann AR, Fuchs RP (2005) Trading places: how do DNA polymerases switch during translesion DNA synthesis? *Mol Cell*. **18**: 499-505.

Fuchs RP, Fujii S, Wagner J (2004) Properties and functions of Escherichia coli: Pol IV and Pol V. *Adv Protein Chem*. **69**: 229-264.

Fujii S, Fuchs RP (2004) Defining the position of the switches between replicative and bypass DNA polymerases. *Embo J*. **23**: 4342-4352.

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CP, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap J, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B (2003) The genome sequence of the filamentous fungus Neurospora crassa. *Nature*. **422**: 859-868.

Galitski T, Roth JR (1995) Evidence that F plasmid transfer replication underlies apparent adaptive mutation. *Science*. **268**: 421-423.

Gelb LD, Kohne DE, Martin MA (1971) Quantitation of Simian virus 40 sequences in African green monkey, mouse and virus-transformed cell genomes. *J Mol Biol*. **57**: 129-145.

Gibbs SP (1968) Autoradiographic evidence for the *in situ* synthesis of chloroplast and mitochondrial RNA. *J Cell Sci*. **3**: 327-340.

Gibbs SP (1979) The route of entry of cytoplasmically synthesized proteins into chloroplasts of algae possessing chloroplast ER. *J Cell Sci*. **35**: 253-266.

Gibbs SP (1981) The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae. *Ann N Y Acad Sci*. **361**: 193-208.

Gibbs SP, Poole RJ (1973) Autoradiographic evidence for many segregating DNA molecules in the chloroplast of Ochromonas danica. *J Cell Biol*. **59**: 318-328.

Gilmore AM (1997) Mechanistic aspects of xanthophyll cycle-dependent photoprotection in higher plant chloroplasts and leaves. *Physiologia Plantarum*. **99**: 197-209.

Golds T, Maliga P, Koop H-U (1993) Stable Plastid Transformation in PEG-treated Protoplasts of Nicotiana tabacum. *Nat Biotech*. **11**: 95.

Goldschmidt-Clermont M (1991) Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker of site-directed transformation of chlamydomonas. *Nucleic Acids Res*. **19**: 4083-4089.

Goodman MF (2002) Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem*. **71**: 17-50.

Gorsich SW, Shaw JM (2004) Importance of mitochondrial dynamics during meiosis and sporulation. *Mol Biol Cell*. **15**: 4369-4381.

Goss R, Ann Pinto E, Wilhelm C, Richter M (2006) The importance of a highly active and DeltapH-regulated diatoxanthin epoxidase for the regulation of the PS II antenna function in diadinoxanthin cycle containing algae. *J Plant Physiol*. **163**: 1008-1021.

Gray MW, Lang BF, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Brossard N, Delage E, Littlejohn TG, Plante I, Rioux P, Saint-Louis D, Zhu Y, Burger G (1998) Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Res*. **26**: 865-878.

Grossman AR (2000) Chlamydomonas reinhardtii and photosynthesis: genetics to genomics. *Curr Opin Plant Biol*. **3**: 132-137.

GuhaMajumdar M, Sears BB (2005) Chloroplast DNA base substitutions: an experimental assessment. *Mol Genet Genomics*. **273**: 177-183.

Guo S, Kemphues KJ (1995) par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell*. **81**: 611-620.

Hager M, Biehler K, Illerhaus J, Ruf S, Bock R (1999) Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome b(6)f complex. *Embo J*. **18**: 5834-5842.

Hagopian JC, Reis M, Kitajima JP, Bhattacharya D, de Oliveira MC (2004) Comparative analysis of the complete plastid genome sequence of the red alga Gracilaria tenuistipitata var. liui provides insights into the evolution of rhodoplasts and their relationship to other plastids. *J Mol Evol*. **59**: 464-477.

Hammond SM, Bernstein E, Beach D, Hannon GJ (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature*. **404**: 293-296.

Harris EH (2001) Chlamydomonas As A Model Organism. *Annu Rev Plant Physiol Plant Mol Biol*. **52**: 363-406.

Harris RS, Longerich S, Rosenberg SM (1994) Recombination in adaptive mutation. *Science*. **264**: 258-260.

Harris RS, Ross KJ, Rosenberg SM (1996) Opposing roles of the holliday junction processing systems of Escherichia coli in recombination-dependent adaptive mutation. *Genetics*. **142**: 681-691.

He AS, Rohatgi PR, Hersh MN, Rosenberg SM (2006) Roles of E. coli double-strand-break-repair proteins in stress-induced mutation. *DNA Repair (Amst)*. **5**: 258-273.

Heidenreich E, Novotny R, Kneidinger B, Holzmann V, Wintersberger U (2003) Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells. *Embo J*. **22**: 2274-2283.

Hengge-Aronis R (2002) Recent insights into the general stress response regulatory network in Escherichia coli. *J Mol Microbiol Biotechnol*. **4**: 341-346.

Hersh MN, Ponder RG, Hastings PJ, Rosenberg SM (2004) Adaptive mutation and amplification in Escherichia coli: two pathways of genome adaptation under stress. *Res Microbiol*. **155**: 352-359.

Holt NE, Fleming GR, Niyogi KK (2004) Toward an understanding of the mechanism of nonphotochemical quenching in green plants. *Biochemistry*. **43**: 8281-8289.

Horton P, Wentworth M, Ruban A (2005) Control of the light harvesting function of chloroplast membranes: the LHCII-aggregation model for non-photochemical quenching. *FEBS Lett*. **579**: 4201-4206.

Horton P, Park K-J, Obayashi T, Nakai K. Protein Subcellular Localization Prediction with WoLF PSORT. Proceedings of the 4th Annual Asia Pacific Bioinformatics Conference APBC06; 2006; Taipei, Taiwan. pp. 39-48.

Huang FC, Klaus SM, Herz S, Zou Z, Koop HU, Golds TJ (2002) Efficient plastid transformation in tobacco using the aphA-6 gene and kanamycin selection. *Mol Genet Genomics*. **268**: 19-27.

Inada N, Sakai A, Kuroiwa H, Kuroiwa T (1998) Three-dimensional analysis of the senescence program in rice (Oryza sativa L.) coleoptiles - Investigations by fluorescence microscopy and electron microscopy. *Planta*. **4**: 585-597.

Ishida K-i, Cavalier-Smith T, Green BR (2000) ENDOMEMBRANE STRUCTURE AND THE CHLOROPLAST PROTEIN TARGETING PATHWAY IN HETEROSIGMA AKASHIWO (RAPHIDOPHYCEAE, CHROMISTA). *Journal of Phycology*. **36**: 1135-1144.

Jeong SY, Rose A, Meier I (2003) MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. *Nucleic Acids Res*. **31**: 5175-5185.

Kavanagh TA, Thanh ND, Lao NT, McGrath N, Peter SO, Horvath EM, Dix PJ, Medgyesy P (1999) Homeologous plastid DNA transformation in tobacco is mediated by multiple recombination events. *Genetics*. **152**: 1111-1122.

Khakhlova O, Bock R (2006) Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J*. **46**: 85-94.

Kilian O, Kroth PG (2004) Presequence acquisition during secondary endocytobiosis and the possible role of introns. *J Mol Evol*. **58**: 712-721.

Kilian O, Kroth PG (2005) Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. *Plant J*. **41**: 175-183.

Kindle KL, Richards KL, Stern DB (1991) Engineering the chloroplast genome: techniques and capabilities for chloroplast transformation in Chlamydomonas reinhardtii. *Proc Natl Acad Sci U S A*. **88**: 1721-1725.

Klein RM, Wolf ED, Wu R, Sanford JC (1992) High-velocity microprojectiles for delivering nucleic acids into living cells. 1987. *Biotechnology*. **24**: 384-386.

Kooistra WH, De Stefano M, Mann DG, Medlin LK (2003) The phylogeny of the diatoms. *Prog Mol Subcell Biol*. **33**: 59-97.

Koop HU, Steinmuller K, Wagner H, Rossler C, Eibl C, Sacher L (1996) Integration of foreign sequences into the tobacco plastome via polyethylene glycol-mediated protoplast transformation. *Planta*. **199**: 193-201.

Kroth PG (2002) Protein transport into secondary plastids and the evolution of primary and secondary plastids. *Int Rev Cytol*. **221**: 191-255.

Kroth PG, Strotmann H (1999) Diatom plastids: secondary endocytobiosis, plastid genome and protein import. *Physiologia Plantarum*. **107**: 136-141.

Kroth PG, Schroers Y, Kilian O (2005) The peculiar distribution of class I and class II aldolases in diatoms and in red algae. *Curr Genet*. **48**: 389-400.

Kumar S, Dhingra A, Daniell H (2004) Stable transformation of the cotton plastid genome and maternal inheritance of transgenes. *Plant Mol Biol*. **56**: 203-216.

Kunkel TA, Pavlov YI, Bebenek K (2003) Functions of human DNA polymerases eta, kappa and iota suggested by their properties, including fidelity with undamaged DNA templates. *DNA Repair (Amst)*. **2**: 135-149.

Kuroiwa T (1982) Mitochondrial nuclei. *Int Rev Cytol*. **75**: 1-59.

Kuroiwa T (1991) The replication, differentiation, and inheritance of plastids with emphasis on the concept of organelle nuclei. *Int Rev Cytol*. **128**: 1–62.

Lapidot M, Raveh D, Sivan A, Arad SM, Shapira M (2002) Stable chloroplast transformation of the unicellular red alga Porphyridium species. *Plant Physiol*. **129**: 7-12.

Lavaud J, Kroth PG (2006) In diatoms, the transthylakoid proton gradient regulates the photoprotective non-photochemical fluorescence quenching beyond its control on the xanthophyll cycle. *Plant Cell Physiol*. **47**: 1010-1016.

Lavaud J, Rousseau B, van Gorkom HJ, Etienne AL (2002) Influence of the diadinoxanthin pool size on photoprotection in the marine planktonic diatom Phaeodactylum tricornutum. *Plant Physiol*. **129**: 1398-1406.

Lebeau T, Robert JM (2003a) Diatom cultivation and biotechnologically relevant products. Part II: current and putative products. *Appl Microbiol Biotechnol*. **60**: 624-632.

Lebeau T, Robert JM (2003b) Diatom cultivation and biotechnologically relevant products. Part I: cultivation at various scales. *Appl Microbiol Biotechnol*. **60**: 612-623.

Lederberg J, Lederberg EM (1952) Replica plating and indirect selection of bacterial mutants. *J Bacteriol*. **63**: 399-406.

Lelivelt CL, McCabe MS, Newell CA, Desnoo CB, van Dun KM, Birch-Machin I, Gray JC, Mills KH, Nugent JM (2005) Stable plastid transformation in lettuce (Lactuca sativa L.). *Plant Mol Biol*. **58**: 763-774.

Lewin JC, Lewin RA, Philpott DE (1958) Observations on Phaeodactylum tricornutum. *J Gen Microbiol*. **18**: 418-426.

Li W, Ruf S, Bock R (2006) Constancy of organellar genome copy numbers during leaf development and senescence in higher plants. *Mol Genet Genomics*. **275**: 185-192.

Liaud MF, Lichtle C, Apt K, Martin W, Cerff R (2000) Compartment-specific isoforms of TPI and GAPDH are imported into diatom mitochondria as a fusion protein: evidence in favor of a mitochondrial origin of the eukaryotic glycolytic pathway. *Mol Biol Evol*. **17**: 213-223.

Lohr M, Wilhelm C (1999) Algae displaying the diadinoxanthin cycle also possess the violaxanthin cycle. *Proc Natl Acad Sci U S A*. **96**: 8784-8789.

Lopez PJ, Descles J, Allen AE, Bowler C (2005) Prospects in diatom research. *Curr Opin Biotechnol*. **16**: 180-186.

Luria SE, Delbruck M (1943) MUTATIONS OF BACTERIA FROM VIRUS SENSITIVITY TO VIRUS RESISTANCE. *Genetics*. **28**: 491-511.

MacAlpine DM, Perlman PS, Butow RA (2000) The numbers of individual mitochondrial DNA molecules and mitochondrial DNA nucleoids in yeast are co-regulated by the general amino acid control pathway. *Embo J*. **19**: 767-775.

Maenpaa P, Gonzalez EB, Chen L, Khan MS, Gray JC, Aro EM (2000) The ycf 9 (orf 62) gene in the plant chloroplast genome encodes a hydrophobic protein of stromal thylakoid membranes. *J Exp Bot*. **51 Spec No**: 375-382.

Maier RM, Neckermann K, Igloi GL, Kossel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J Mol Biol*. **251**: 614-628.

Makeyev EV, Bamford DH (2002) Cellular RNA-dependent RNA polymerase involved in posttranscriptional gene silencing has two distinct activity modes. *Mol Cell*. **10**: 1417-1427.

Maki H (2002) Origins of spontaneous mutations: specificity and directionality of base-substitution, frameshift, and sequence-substitution mutageneses. *Annu Rev Genet*. **36**: 279-303.

Maliga P (2003) Progress towards commercialization of plastid transformation technology. *Trends Biotechnol*. **21**: 20-28.

Maliga P (2004) Plastid transformation in higher plants. *Annu Rev Plant Biol*. **55**: 289-313.

Maliszewska-Tkaczyk M, Jonczyk P, Bialoskorska M, Schaaper RM, Fijalkowska IJ (2000) SOS mutator activity: unequal mutagenesis on leading and lagging strands. *Proc Natl Acad Sci U S A*. **97**: 12678-12683.

Mann DG (1993) Patterns of sexual reproduction in diatoms. *Hydrobiologia*. **269-270**: 11.

Mann DG (1999) The species concept in diatoms. *Phycologia*. **38**: 437-495.

Matic I, Taddei F, Radman M (2004) Survival versus maintenance of genetic stability: a conflict of priorities during stress. *Res Microbiol*. **155**: 337-341.

McGill CB, Holbeck SL, Strathern JN (1998) The chromosome bias of misincorporations during double-strand break repair is not altered in mismatch repair-defective strains of Saccharomyces cerevisiae. *Genetics*. **148**: 1525-1533.

McKenzie GJ, Harris RS, Lee PL, Rosenberg SM (2000) The SOS response regulates adaptive mutation. *Proc Natl Acad Sci U S A*. **97**: 6646-6651.

McKenzie GJ, Lee PL, Lombardo MJ, Hastings PJ, Rosenberg SM (2001) SOS mutator DNA polymerase IV functions in adaptive mutation and not adaptive amplification. *Mol Cell*. **7**: 571-579.

Medgyesy P, Fejes E, Maliga P (1985) Interspecific chloroplast recombination in a Nicotiana somatic hybrid. *Proc Natl Acad Sci U S A*. **82**: 6960-6964.

Meister G, Tuschl T (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature*. **431**: 343-349.

Melancon P, Lemieux C, Brakier-Gingras L (1988) A mutation in the 530 loop of Escherichia coli 16S ribosomal RNA causes resistance to streptomycin. *Nucleic Acids Res*. **16**: 9631-9639.

Mello CC, Conte D, Jr. (2004) Revealing the world of RNA interference. *Nature*. **431**: 338-342.

Michels AK, Wedel N, Kroth PG (2005) Diatom plastids possess a phosphoribulokinase with an altered regulation and no oxidative pentose phosphate pathway. *Plant Physiol*. **137**: 911-920.

Miller JH (1996) Spontaneous mutators in bacteria: insights into pathways of mutagenesis and repair. *Annu Rev Microbiol*. **50**: 625-643.

Misumi O, Nishimura Y, Kuroiwa T (2001) Effects if chloroplast DNA content on the cell proliferantion and aging in *Chlamydomonas reinhardtii*. *J Plant Res*. **114**: 125-131.

Miyamura S, Nagata T, Kuroiwa T (1986) Quantitative fluorescence microscopy on dynamic changes of plastid nucleoids during wheat development. *Protoplasma*. **133**: 66.

Mo JY, Maki H, Sekiguchi M (1991) Mutational specificity of the dnaE173 mutator associated with a defect in the catalytic subunit of DNA polymerase III of Escherichia coli. *J Mol Biol*. **222**: 925-936.

Mogensen HL (1996) The Hows and Whys of Cytoplasmic Inheritance in Seed Plants. *American Journal of Botany*. **83**: 383-404.

Montandon PE, Nicolas P, Schurmann P, Stutz E (1985) Streptomycin-resistance of Euglena gracilis chloroplasts: identification of a point mutation in the 16S rRNA gene in an invariant position. *Nucleic Acids Res*. **13**: 4299-4310.

Montsant A, Jabbari K, Maheswari U, Bowler C (2005) Comparative genomics of the pennate diatom Phaeodactylum tricornutum. *Plant Physiol*. **137**: 500-513.

Morton BR, Clegg MT (1993) A chloroplast DNA mutational hotspot and gene conversion in a noncoding region near rbcL in the grass family (Poaceae). *Curr Genet*. **24**: 357-365.

Mourelatos Z, Dostie J, Paushkin S, Sharma A, Charroux B, Abel L, Rappsilber J, Mann M, Dreyfuss G (2002) miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev*. **16**: 720-728.

Mourrain P, Beclin C, Elmayan T, Feuerbach F, Godon C, Morel JB, Jouette D, Lacombe AM, Nikic S, Picault N, Remoue K, Sanial M, Vo TA, Vaucheret H (2000) Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell*. **101**: 533-542.

Muller HJ (1927) Artificial transmutation of the gene. *Science*. **66**: 69-92.

Muller HJ (1964) The Relation Of Recombination To Mutational Advance. *Mutat Res*. **106**: 2-9.

Muse SV (2000) Examining rates and patterns of nucleotide substitution in plants. *Plant Mol Biol*. **42**: 25-43.

Newcombe HB (1949) Origin of bacterial variants. *Nature*. 150-151.

Newman SM, Boynton JE, Gillham NW, Randolph-Anderson BL, Johnson AM, Harris EH (1990) Transformation of chloroplast ribosomal RNA genes in Chlamydomonas: molecular and genetic characterization of integration events. *Genetics*. **126**: 875-888.

Nielsen H, Krogh A (1998) Prediction of signal peptides and signal anchors by a hidden Markov model. *Proc Int Conf Intell Syst Mol Biol*. **6**: 122-130.

Norton TA, Andersen RA, Melkonian M (1996) Algal Biodiversity. *Phycologia*. **35**: 308-326.

O'Neill C, Horvath GV, Horvath E, Dix PJ, Medgyesy P (1993) Chloroplast transformation in plants: polyethylene glycol (PEG) treatment of protoplasts is an alternative to biolistic delivery systems. *Plant J*. **3**: 729-738.

Ohad N, Hirschberg J (1992) Mutations in the D1 subunit of photosystem II distinguish between quinone and herbicide binding sites. *Plant Cell*. **4**: 273-282.

Okumura S, Sawada M, Park YW, Hayashi T, Shimamura M, Takase H, Tomizawa K (2006) Transformation of poplar (Populus alba) plastids and expression of foreign proteins in tree chloroplasts. *Transgenic Res*. **15**: 637-646.

Olaizola M, Yamamoto HY (1994) Short-term response of the diadinoxanthin cycle and fluorescence yield to high irradiance in Chaetoceros muelleri (Bacillariophyceae). *Journal of Phycology*. **30**: 606–612.

Olaizola M, Roche J, Kolber Z, Falkowski PG (1994) Non-photochemical fluorescence quenching and the diadinoxanthin cycle in a marine diatom. *Photosynthesis Research*. **41**: 357.

Oldenburg DJ, Bendich AJ (2004) Most chloroplast DNA of maize seedlings in linear molecules with defined ends and branched forms. *J Mol Biol*. **335**: 953-970.

Owens TG (1986) Light-Harvesting Function in the Diatom Phaeodactylum tricornutum: II. Distribution of Excitation Energy between the Photosystems. *Plant Physiol*. **80**: 739-746.

Pages V, Fuchs RP (2002) How DNA lesions are turned into mutations within cells? *Oncogene*. **21**: 8957-8966.

Palmer JD (1985) Comparative organization of chloroplast genomes. *Annu Rev Genet*. **19**: 325-354.

Parkinson J, Gordon R (1999) Beyond micromachining: the potential of diatoms. *Trends Biotechnol*. **17**: 190-196.

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res*. **30**: e36.

Ponder RG, Fonville NC, Rosenberg SM (2005) A switch from high-fidelity to error-prone DNA double-strand break repair underlies stress-induced mutation. *Mol Cell*. **19**: 791-804.

Radicella JP, Park PU, Fox MS (1995) Adaptive mutation in Escherichia coli: a role for conjugation. *Science*. **268**: 418-420.

Radicella JP, Dherin C, Desmaze C, Fox MS, Boiteux S (1997) Cloning and characterization of hOGG1, a human homolog of the OGG1 gene of Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A*. **94**: 8010-8015.

Rawson JR, Boerma C (1976) Influence of growth conditions upon the number of chloroplast DNA molecules in Euglena gracilis. *Proc Natl Acad Sci U S A*. **73**: 2401-2404.

Reenan RA, Kolodner RD (1992) Isolation and characterization of two Saccharomyces cerevisiae genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins. *Genetics*. **132**: 963-973.

Reinfelder JR, Kraepiel AM, Morel FM (2000) Unicellular C4 photosynthesis in a marine diatom. *Nature*. **407**: 996-999.

Reith M, Munholland J (1995) Complete nucleotide sequence of the Porphyra purpurea chloroplast genome. *Plant Mol Biol Rep*. **13**: 333-335.

Rosenberg SM (2001) Evolving responsively: adaptive mutation. *Nat Rev Genet*. **2**: 504-515.

Round FE, Crawford RM, Mann DG (1990) The Diatoms: Biology and Morphology of a Genera.

Rowan BA, Oldenburg DJ, Bendich AJ (2004) The demise of chloroplast DNA in Arabidopsis. *Curr Genet*. **46**: 176-181.

Ruf S, Kossel H, Bock R (1997) Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. *J Cell Biol*. **139**: 95-102.

Ruf S, Biehler K, Bock R (2000) A small chloroplast-encoded protein as a novel architectural component of the light-harvesting antenna. *J Cell Biol*. **149**: 369-378.

Ruf S, Hermann M, Berger IJ, Carrer H, Bock R (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. *Nat Biotechnol*. **19**: 870-875.

Sakumi K, Furuichi M, Tsuzuki T, Kakuma T, Kawabata S, Maki H, Sekiguchi M (1993) Cloning and expression of cDNA for a human enzyme that hydrolyzes 8-oxo-dGTP, a mutagenic substrate for DNA synthesis. *J Biol Chem*. **268**: 23524-23530.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sato N, Ohta N (2001) DNA-binding specificity and dimerization of the DNA-binding domain of the PEND protein in the chloroplast envelope membrane. *Nucleic Acids Res*. **29**: 2244-2250.

Sato N, Terasawa K, Miyajima K, Kabeya Y (2003) Organization, developmental dynamics, and evolution of plastid nucleoids. *Int Rev Cytol*. **232**: 217-262.

Sato N, Albrieux C, Joyard J, Douce R, Kuroiwa T (1993) Detection and characterization of a plastid envelope DNA-binding protein which may anchor plastid nucleoids. *Embo J*. **12**: 555-561.

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

Schiebel W, Haas B, Marinkovic S, Klanner A, Sanger HL (1993) RNA-directed RNA polymerase from tomato leaves. II. Catalytic in vitro properties. *J Biol Chem*. **268**: 11858-11867.

Schlacher K, Leslie K, Wyman C, Woodgate R, Cox MM, Goodman MF (2005) DNA polymerase V and RecA protein, a minimal mutasome. *Mol Cell*. **17**: 561-572.

Seitz EM, Kowalczykowski SC (2000) The DNA binding and pairing preferences of the archaeal RadA protein demonstrate a universal characteristic of DNA strand exchange proteins. *Mol Microbiol*. **37**: 555-560.

Selker EU (1990) Premeiotic instability of repeated sequences in Neurospora crassa. *Annu Rev Genet*. **24**: 579-613.

Selker EU, Cambareri EB, Jensen BC, Haack KR (1987) Rearrangement of duplicated DNA in specialized cells of Neurospora. *Cell*. **51**: 741-752.

Sidorov VA, Kasten D, Pang SZ, Hajdukiewicz PT, Staub JM, Nehra NS (1999) Technical Advance: Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. *Plant J*. **19**: 209-216.

Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, Plasterk RH, Fire A (2001) On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell*. **107**: 465-476.

Slupska MM, Baikalov C, Luther WM, Chiang JH, Wei YF, Miller JH (1996) Cloning and sequencing a human homolog (hMYH) of the Escherichia coli mutY gene whose function is required for the repair of oxidative DNA damage. *J Bacteriol*. **178**: 3885-3892.

Smardon A, Spoerke JM, Stacey SC, Klein ME, Mackin N, Maine EM (2000) EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. *Curr Biol*. **10**: 169-178.

Smeda RJ, Hasegawa PM, Goldsbrough PB, Singh NK, Weller SC (1993) A serine-to-threonine substitution in the triazine herbicide-binding protein in potato cells results in atrazine resistance without impairing productivity. *Plant Physiol*. **103**: 911-917.

Smith NA, Singh SP, Wang MB, Stoutjesdijk PA, Green AG, Waterhouse PM (2000) Total silencing by intron-spliced hairpin RNAs. *Nature*. **407**: 319-320.

Sodmergen, Kawano S, Tano S, Kuroiwa T (1989) Preferential digestion of chloroplast nuclei (nucleoids) during senescence of the coleoptile ofOryza sativa. *Protoplasma*. **152**: 65.

Sodmergen, Kawano S, Tano S, Kuroiwa T (1991) Degradation of chloroplast DNA in second leaves of rice (Oryza sativa) before leaf yellowing. *Protoplasma*. **160**: 89.

Springer B, Kidan YG, Prammananan T, Ellrott K, Bottger EC, Sander P (2001) Mechanisms of streptomycin resistance: selection of mutations in the 16S rRNA gene conferring resistance. *Antimicrob Agents Chemother*. **45**: 2877-2884.

Starr RC, Zeikus JA (1993) UTEX: the culture collection of algae at the University of Texas at Austin, 1993 list of cultures. *J Phycol*. **29**: 1–106.

Staub JM, Maliga P (1992) Long regions of homologous DNA are incorporated into the tobacco plastid genome by transformation. *Plant Cell*. **4**: 39-45.

Staub JM, Maliga P (1994) Extrachromosomal elements in tobacco plastids. *Proc Natl Acad Sci U S A*. **91**: 7468-7472.

Staub JM, Maliga P (1995) Marker rescue from the Nicotiana tabacum plastid genome using a plastid/Escherichia coli shuttle vector. *Mol Gen Genet*. **249**: 37-42.

Stevens B (1981) Mitochondrial structure. In: Strathern JN, Jones EW, Broach JR, editors. The Molecular Biology of the Yeast Saccharomyces. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. pp. 471–504.

Storseth TR, Hansen K, Reitan KI, Skjermo J (2005) Structural characterization of beta-D-(1->3)-glucans from different growth phases of the marine diatoms Chaetoceros mulleri and Thalassiosira weissflogii. *Carbohydr Res*. **340**: 1159-1164.

Stransky H, Hager A (1970) [The carotenoid pattern and the occurrence of the light-induced xanthophyll cycle in various classes of algae. VI. Chemosystematic study]. *Arch Mikrobiol*. **73**: 315-323.

Sugiura M (1992) The chloroplast genome. *Plant Mol Biol*. **19**: 149-168.

Svab Z, Hajdukiewicz P, Maliga P (1990) Stable transformation of plastids in higher plants. *Proc Natl Acad Sci U S A*. **87**: 8526-8530.

Swiatek M, Kuras R, Sokolenko A, Higgs D, Olive J, Cinque G, Muller B, Eichacker LA, Stern DB, Bassi R, Herrmann RG, Wollman FA (2001) The chloroplast gene ycf9 encodes a photosystem II (PSII) core subunit, PsbZ, that participates in PSII supramolecular architecture. *Plant Cell*. **13**: 1347-1367.

Tang M, Shen X, Frank EG, O'Donnell M, Woodgate R, Goodman MF (1999) UmuD'(2)C is an error-prone DNA polymerase, Escherichia coli pol V. *Proc Natl Acad Sci U S A*. **96**: 8919-8924.

Tenaillon O, Denamur E, Matic I (2004) Evolutionary significance of stress-induced mutagenesis in bacteria. *Trends Microbiol*. **12**: 264-270.

Terasawa K, Sato N (2005a) Occurrence and characterization of PEND proteins in angiosperms. *J Plant Res*. **118**: 111-119.

Terasawa K, Sato N (2005b) Visualization of plastid nucleoids in situ using the PEND-GFP fusion protein. *Plant Cell Physiol*. **46**: 649-660.

Thanh ND, Medgyesy P (1989) Limited chloroplast gene transfer via recombination overcomes plastomegenome incompatibility between Nicotiana tabacum and Solanum tuberosum. *Plant Molecular Biology*. **12**: 87-93.

Tonon T, Harvey D, Larson TR, Graham IA (2002) Long chain polyunsaturated fatty acid production and partitioning to triacylglycerols in four microalgae. *Phytochemistry*. **61**: 15-24.

van Bel AJ, Hibberd J, Prufer D, Knoblauch M (2001) Novel approach in plastid transformation. *Curr Opin Biotechnol*. **12**: 144-149.

van den Hoek C, Mann DG, Jahns HM (1997) Algae. An introduction to Phycology.

van der Kemp PA, Thomas D, Barbey R, de Oliveira R, Boiteux S (1996) Cloning and expression in Escherichia coli of the OGG1 gene of Saccharomyces cerevisiae, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine. *Proc Natl Acad Sci U S A*. **93**: 5197-5202.

Van Oijen T, Veldhuis MJW, Gorbunov MY, Nishioka J, Van Leeuwe MA, De Baar HJW (2005) Enhanced carbohydrate production by Southern Ocean phytoplankton in response to in situ iron fertilization. *Marine Chemistry*. **93**: 33.

Varum KM, Myklestad S (1984) Effects of light, salinity and nutrient limitation on the production of [beta]-1,3--glucan and exo--glucanase activity in Skeletonema costatum (Grev.) Cleve. *Journal of Experimental Marine Biology and Ecology*. **83**: 13.

Veldhuis MJW, Cucci TL, Sieracki ME (1997) Cellular DNA content of marine phytoplankton using two new fluorochromes: taxonomic and ecological implications. *Journal of Phycology*. 527-541.

Wakasugi T, Tsudzuki T, Sugiura M (2001) The genomics of land plant chloroplasts: Gene content and alteration of genomic information by RNA editing. *Photosynth Res*. **70**: 107-118.

Wang T, Li Y, Shi Y, Reboud X, Darmency H, Gressel J (2004) Low frequency transmission of a plastid-encoded trait in Setaria italica. *Theor Appl Genet*. **108**: 315-320.

Wassenegger M, Krczal G (2006) Nomenclature and functions of RNA-directed RNA polymerases. *Trends Plant Sci*. **11**: 142-151.

Waterhouse PM, Helliwell CA (2003) Exploring plant genomes by RNA-induced gene silencing. *Nat Rev Genet*. **4**: 29-38.

Williamson DH, Fennell DJ (1975) The use of fluorescent DNA-binding agent for detecting and separating yeast mitochondrial DNA. *Methods Cell Biol*. **12**: 335-351.

Williamson DH, Fennell DJ (1979) Visualization of yeast mitochondrial DNA with the fluorescent stain "DAPI". *Methods Enzymol*. **56**: 728-733.

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci U S A*. **84**: 9054-9058.

Wyman C, Kanaar R (2004) Homologous recombination: down to the wire. *Curr Biol*. **14**: R629-631.

Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol*. **2**: E104.

Xiong L, Sayre RT (2004) Engineering the chloroplast encoded proteins of chlamydomonas. *Photosynth Res*. **80**: 411-419.

Yeiser B, Pepper ED, Goodman MF, Finkel SE (2002) SOS-induced DNA polymerases enhance long-term survival and evolutionary fitness. *Proc Natl Acad Sci U S A*. **99**: 8737-8741.

Yoshiyama K, Higuchi K, Matsumura H, Maki H (2001) Directionality of DNA replication fork movement strongly affects the generation of spontaneous mutations in Escherichia coli. *J Mol Biol*. **307**: 1195-1206.

Zamore PD, Tuschl T, Sharp PA, Bartel DP (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*. **101**: 25-33.

Zaouk R, Park BY, Madou MJ (2006) Introduction to microfabrication techniques. *Methods Mol Biol*. **321**: 5-15.

Zaslavskaia LA, Lippmeier JC, Kroth PG, Grossman AR, Apt KE (2000) Transformation of the diatom Phaeodactylum tricornutum (Bacillariophyceae) with a variety of selectable marker and reporter genes. *Journal of Phycology*. **36**: 379-386.

Zaslavskaia LA, Lippmeier JC, Shih C, Ehrhardt D, Grossman AR, Apt KE (2001) Trophic conversion of an obligate photoautotrophic organism through metabolic engineering. *Science*. **292**: 2073-2075.

Zurzolo C, Bowler C (2001) Exploring bioinorganic pattern formation in diatoms. A story of polarized trafficking. *Plant Physiol*. **127**: 1339-1345.

Contributions

Conceptualization and implementation of the experimental designs, data gathering and analysis, and the writing of the manuscripts were initiated and performed by A.C. Materna.

Co-Authors contributed in parts on experimental works. Ansgar Gruber, the second author of chapter II.4, planned and implemented the visual analysis of the transformed strains via fluorescence microscopy and confocal laser scanning microscopy. He further contributed to the manuscript editing (chapter II.4).

Prof. Peter G. Kroth was responsible for the supervision of this dissertation.