University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Theses and Dissertations in Biochemistry | Biochemistry, Department of

Winter 12-5-2013

# MAKING *CHLAMYDOMONAS REINHARDTII* A BETTER MODEL ORGANISM: TACKLING THE INEFFICIENCY OF NUCLEAR TRANSGENE EXPRESSION AND IMPROVING METHODS FOR THE GENERATION AND CHARACTERIZATION OF INSERTIONAL MUTANT LIBRARIES

Thomas M. Plucinak
*University of Nebraska-Lincoln*, thomas.plucinak@huskers.unl.edu

Follow this and additional works at: https://digitalcommons.unl.edu/biochemdiss

Part of the Biochemistry Commons

Plucinak, Thomas M., "MAKING *CHLAMYDOMONAS REINHARDTII* A BETTER MODEL ORGANISM: TACKLING THE INEFFICIENCY OF NUCLEAR TRANSGENE EXPRESSION AND IMPROVING METHODS FOR THE GENERATION AND CHARACTERIZATION OF INSERTIONAL MUTANT LIBRARIES" (2013). *Theses and Dissertations in Biochemistry*. 15.
https://digitalcommons.unl.edu/biochemdiss/15

This Article is brought to you for free and open access by the Biochemistry, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Theses and Dissertations in Biochemistry by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# MAKING *CHLAMYDOMONAS REINHARDTII* A BETTER MODEL ORGANISM: TACKLING THE INEFFICIENCY OF NUCLEAR TRANSGENE EXPRESSION AND IMPROVING METHODS FOR THE GENERATION AND CHARACTERIZATION OF INSERTIONAL MUTANT LIBRARIES

by

Thomas M. Plucinak

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Biochemistry

Under the Supervision of Professor Donald P. Weeks

Lincoln, Nebraska

December, 2013

MAKING *CHLAMYDOMONAS REINHARDTII* A BETTER MODEL ORGANISM: TACKLING THE INEFFICIENCY OF NUCLEAR TRANSGENE EXPRESSION AND IMPROVING METHODS FOR THE GENERATION AND CHARACTERIZATION OF INSERTIONAL MUTANT LIBRARIES

Thomas M. Plucinak, Ph.D.

University of Nebraska, 2013

**Advisor: Donald P. Weeks**

The green algal species *Chlamydomonas reinhardtii* possesses many beneficial features that have made it a useful model organism for many decades. Many types of experimentation however are difficult to conduct with this organism due to the relative under-development of genetic tools available for use. Tasks such as transgene expression, overexpression of proteins of interest (POIs) or site specific genomic modification that are routine in other more facile microbial model organisms such as *Escherichia coli* and yeast are difficult to accomplish in *C. reinhardtii*. The second chapter of this thesis describes the development of a novel nuclear transgene expression system that alleviates many of the difficulties associated with transgene expression in *C. reinhardtii* and in some instances allows for overexpression of POIs. This system relies on viral 2A peptides to link the coding regions of selectable marker genes to those of other transgenes of interest – transgenes that do not confer a selectable phenotype to cells when expressed. Implementation of this system allows for the construction and utilization of polycistronic genes in *C. reinhardtii* that significantly outperform traditional vectors used for expressing transgenes. Overall, this new tool enhances the ability to

genetically engineer the nuclear genome of *C. reinhardtii* and expands the overall spectrum of experimentation that can be conducted with this organism.

The third chapter of this thesis describes the development of optimized methods for generating *C. reinhardtii* mutant libraries using insertional mutagenesis and for characterizing the resulting mutants. Because site specific modification of the genome of this organism is extremely inefficient, insertional mutagenesis has necessarily become the most favored methodology for generating mutants that are subsequently used to study various biological systems. Standard insertional mutagenesis procedures however suffer from random insertion of transgenes, resulting in the creation of large populations of mutants bearing insertions in sequences such as intergenic regions that do not contribute significantly to the function of any gene. The optimized insertional mutagenesis method reported here relies on the use of a selectable marker lacking a promoter to generate mutant populations consisting of a large majority of individuals that contain disruptions in transcribed regions. And finally, many of the most common methods that exist for identifying the exact genomic locations of foreign DNA insertion events can be inefficient. To address this deficiency, an optimized version of inverse PCR was developed that allows for the characterization of marker gene insertion events in a highly efficient, facile and cost-effective manner.

# TABLE OF CONTENTS

**TITLE** ..... i

**ABSTRACT** ..... ii

**TABLE OF CONTENTS** ..... iv

**ACKNOWLEDGEMENTS** ..... vii

**CHAPTER 1: Literature Review** ..... 1

*Chlamydomonas reinhardtii*: a useful model organism for many purposes ..... 2

- Photosynthetic properties of *C. reinhardtii* ..... 3
- Genetic transformation of *C. reinhardtii* ..... 5
- Utilization of the *C. reinhardtii* chloroplast as a protein production factory ..... 5

Overview of *C. reinhardtii* properties that restrict its utility as a model organism ..... 7

Utilization of polycistronic genes in eukaryotic organisms ..... 9

- Viral 2A peptides ..... 10

Mutagenesis and gene targeting ..... 13

- Insertional mutagenesis ..... 16

**CHAPTER 2: Development of an improved and versatile viral 2A platform for expression of polycistronic nuclear genes in *Chlamydomonas reinhardtii*** ..... 21

Abstract ..... 22

Introduction ..... 23

Results and Discussion ..... 26

- *Chlamydomonas* processes 2A peptides efficiently and an extended form of the FMDV 2A peptide is processed in a nearly quantitative manner ..... 26
- The 2A expression platform allows expression of GOIs at dependably high levels and significantly outperforms adjacent selection with respect to GOI expression frequency ..... 34
- The 2A expression platform facilitates subcellular localization studies in *Chlamydomonas* ..... 38
- Evaluation of the functionality and efficacy of various selection marker/2A peptide combinations ..... 41
- Utilization of the Heat Shock 70A gene promoter allows robust inducible expression of a 2A-dicistron when coupled with a low expression marker gene ..... 45
- Development of an inducible tricistronic expression cassette for use in *Chlamydomonas* ..... 49

Materials and Methods ..... 55

**CHAPTER 3: Enhanced methods for the generation and characterization of insertional mutant libraries in *Chlamydomonas reinhardtii* containing high percentages of genic insertions** ..... 62

Abstract ..... 63

Introduction ..... 64

Results ..... 69

- Utilization of a restriction enzyme cocktail to fragment the genome prior to inverse PCR generates a favorable restriction fragment size range for inverse PCR ..... 69
- Construction of a minimal marker gene and evaluation of an optimized inverse PCR method relying on the use of a restriction enzyme cocktail to fragment genomic DNA prior to inverse PCR ..... 78
- Analysis of insertion events resulting from transformation with truncated marker genes of various lengths and compositions ..... 83
- Generation of insertional mutants using AphVIII fragments that lack a promoter positively biases insertion events towards transcribed genomic loci ..... 90
- A commercial restriction enzyme buffer significantly improves transformation efficiency when included during electroporation of *Chlamydomonas* cells ... 94

Discussion ..... 97

Materials and Methods ..... 107

**REFERENCES:** ..... 112

## Acknowledgements

First and foremost, I would like to acknowledge the ways in which my advisor, Dr. Donald Weeks, has enriched my graduate school experience. Aside from being a very kind and amiable person, I would like to thank him for his ample support and helpful advice over the years. Furthermore, I would like to thank him for creating an enjoyable and fulfilling work environment where robust intellectual freedom was fostered and my co-workers and I felt like scientific colleagues rather than employees. In the absence of such an environment, life would have been far less enjoyable and I likely would not have been able to progress as a scientist to the extent that I did.

When I joined graduate school, I was a novice that knew next to nothing about the field I was entering. Dr. Kempton Horken was the individual assigned the daunting task of training me and showing me how to think like a scientist. One of the most fortunate aspects of my graduate career is that Kempton and I also shared an office. Throughout graduate school, and especially during the beginning years, I subjected Kempton to weekly and sometimes daily barrages of scientific and philosophical inquiries. Kempton's patience and willingness to help drastically accelerated my development as a scientist and a person. Furthermore, our discourse would often consist of exchanges of scientific ideas aimed at solving some experimental problem or developing new biotechnological tools. These discussions in particular were among the most enjoyable and fruitful of my graduate career and lead to many fascinating and intriguing scientific endeavors. I have long considered Kempton to be a second advisor and a friend. For all of Kempton's

kindness and graciousness in his mentorship as well as his companionship, I owe him an eternal debt of gratitude.

To my fellow graduate students from the lab of Donald Weeks, Amy Miller-Knobbe and Andrew Brueggeman, I have always valued your comradery and support and I thank you for creating a pleasant working environment.

I am thankful to the undergraduate student assistants, Jessica Fostvedt and Sanh Tan Nguyen, who assisted me with my research.

I would also like to thank Dr. Julie Stone and Dr. John Markwell for reading this dissertation and for their helpful comments and suggestions.

I am also thankful for the unconditional support, encouragement and love of my parents. Any success I have in life I owe to them for all they have done and continue to do for me.

Finally, I would also like to extend a sincere thank you to my wife, Yukiko Plucinak, for her continual patience, understanding, support and love through all of the difficulties of graduate school. She always encourages me to be my best and has made me a better person.

# CHAPTER 1: Literature Review

## ***Chlamydomonas reinhardtii*: a useful model organism for many purposes**

The green algal species *Chlamydomonas reinhardtii* is a unicellular, photosynthetic microbe that inhabits aquatic environments (1,2). The first wild type strain of this organism was collected in 1945 (Massachusetts, USA) for scientific purposes by Gilbert M. Smith (1). *C. reinhardtii* cells are encapsulated by a cell wall and contain two large flagella that are used to optimize exposure to light and nutrients (Fig. 1-1) (3,4). These flagella are also vital to the sexual mode of reproduction that occurs in this organism because they mediate the recognition and fusion processes that take place between cells of opposite mating types (2,5). This species is easily cultured in laboratory conditions and most strains have a doubling time of between 6 and 12 hours. Importantly, the *C. reinhardtii* nuclear genome exists in the haploid state during normal vegetative growth (1), rendering the creation and genetic analysis of interesting mutant phenotypes relatively simple compared to other organisms such as higher plants where diploidy and polyploidy complicate such endeavors (6,7).

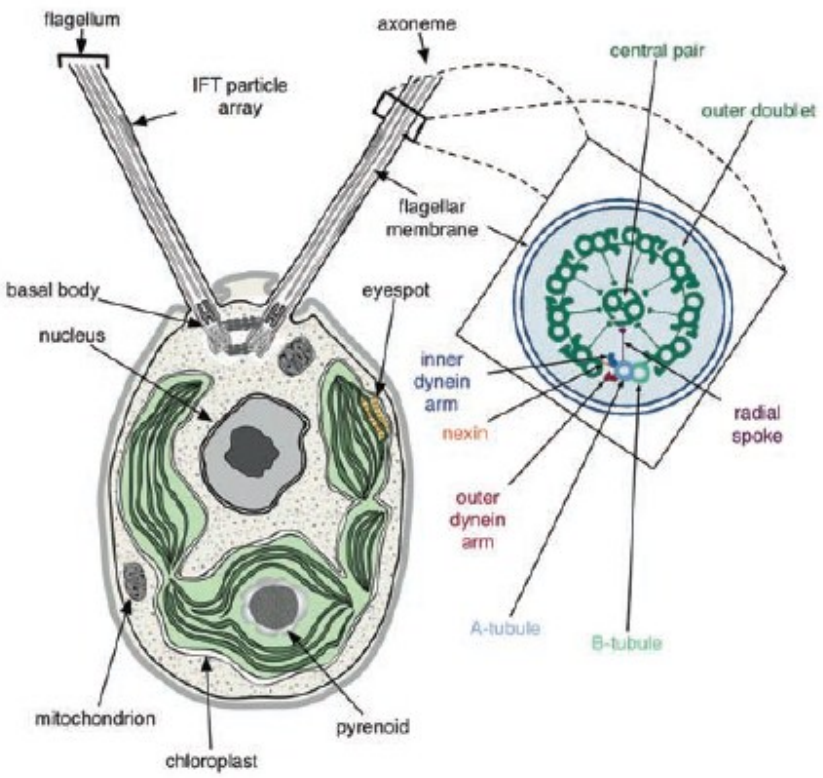



**Figure 1-1: Cross section of a *Chlamydomonas reinhardtii* cell.** Major structural features such as the individual cup-shaped chloroplast, pyrenoid, flagella and nucleus are indicated. A cross section of the flagella displays the organization of important structural proteins. Image adapted from (8).

## Photosynthetic properties of *C. reinhardtii*

An important property of *C. reinhardtii* is that it can grow heterotrophically, using acetate as a sole carbon source, as well as phototrophically (9,10). Because of this, mutants unable to support photosynthesis can be generated and maintained in acetate containing growth media (11). Because viable photosynthetic mutants of land plants are generally not possible to propagate (12), this attribute has enabled significant photosynthetic research to be conducted which would have otherwise

been infeasible (13-15). An example of the power of the ability to generate and maintain photosynthetic mutants of *C. reinhardtii* can be found in studies of a specialized process unique to algal cells known as the Carbon Concentrating Mechanism (CCM) (16,17). Each *C. reinhardtii* cell contains a single large chloroplast, occupying approximately 40% of cellular volume, and under atmospheric $CO_2$ levels, a structure known as the pyrenoid forms inside each chloroplast (3,18). This spherical structure, generally surrounded by a starch sheath, takes up a relatively small portion of the chloroplast and contains its own subset of proteins as well. In *C. reinhardtii,* upwards of 90% of the population of the Rubisco holoenzyme localizes to this structure under atmospheric or limiting $CO_2$ levels (19). To increase the efficiency of the $CO_2$ fixation reaction, *C. reinhardtii* cells employ various transporters and other proteins to effectively “pump” $CO_2$/bicarbonate into the pyrenoid structure, effectively increasing the $CO_2$ concentration at the Rubisco active site (20,21). The presence of such a mechanism allows photosynthesis to occur efficiently in aquatic environments where $CO_2$ availability is restricted compared to terrestrial environments (22). The generation and characterization of numerous mutants possessing disruptions in various processes associated with the CCM has revealed much about this complex system (22). More complete elucidation of the workings of the *C. reinhardtii*l CCM may eventually prove useful in enhancing the photosynthetic efficiency of important land plants.

## Genetic transformation of *C. reinhardtii*

Knowing the complete genomic DNA sequence of any organism provides numerous advantages to those who study it. Thus it is important that the nuclear (8), mitochondrial (23) and chloroplast (24) genomes of *C. reinhardtii* have all been sequenced and annotated. Moreover, methods exist for genetically transforming each of these three genomes with exogenous DNA. The mitochondrial and chloroplast genomes can successfully be transformed by particle bombardment (25-28). With respect to the nuclear genome, particle bombardment (29), electroporation (30) and agitation in the presence of glass beads (31) have all been used to deliver exogenous DNA. However, electroporation produces the highest transformation rates among these three methods. There are also numerous selectable markers available for use in *C. reinhardtii* that enable one to select for transformation events. With regard to transformation of the nuclear genome, a plethora of foreign antibiotic resistance markers have been developed that enable high transformation rates. Prominent examples include: Ble (zeocin resistance) (32), AphVIII (paromomycin resistance) (33), AphVII (hygromycin resistance) (34) and AadA (spectinomycin resistance) (35). In addition, a smaller number of native nuclear genes that either complement auxotrophic mutants (*NIT1* (29) and *ARG7* (36)) or provide resistance to certain biocides when they contain strategic point mutations (*ALS* (37) and *CRY1* (38)*)* can be used as selection markers.

## Utilization of the *C. reinhardtii* chloroplast as a protein production factory

Another beneficial feature of *C. reinhardtii* stems from the fact that its chloroplast can be transformed with ease and made to express foreign proteins at relatively high levels. The beneficial implications of this feature are made even more valuable because this organism is photosynthetic and mutants can be generated, isolated and cultured quickly and easily. Because chloroplasts resemble prokaryotic organisms in many ways, they possess many beneficial features which facilitate transformation and protein production. Particle bombardment is primarily used to transform the *C. reinhardtii* chloroplast genome, where foreign DNA is primarily incorporated via homologous recombination (25). The C. *reinhardtii* chloroplast genome consists of approximately 80 redundant copies of an approximately 200 kb circular plasmid. (39) Stable transformation of all 80 copies with a transgene is achieved by prolonged selection at high antibiotic concentrations, which allows transgene copy number and thus expression level to be high. In addition, the transcriptional stability of transgenes incorporated into the chloroplast genome is high because silencing is not a concern (40). Consequently, expression of valuable proteins from the *C. reinhardtii* chloroplast commonly results in nearly 1%, and in rare cases up to 20%, of total cellular protein consisting of the transgenic product (39,41). To date, a copious number of reports have been published describing the expression of high value proteins with therapeutic or other biotechnological purposes from the *C. reinhardtii* chloroplast (39,42,43).

Because of these beneficial features, *C. reinhardtii* has for many decades been one of the best organisms available with which to study a number of important biological processes including flagella/cilia-based cell motility (4), flagellar assembly

(44,45), human diseases related to defects in cellular motility (46), circadian rhythm (47-49), photosynthesis (13) and the carbon concentrating mechanism (22). More recently, as greater focus has been placed on the development of renewable fuel sources, *C. reinhardtii* and other algal species have also been used to study the production of renewable biofuels in the form of hydrogen gas or oils (50-52). Photosynthetic microbes such as algae and cyanobacteria are promising renewable energy production vehicles for such compounds because they can utilize atmospheric $CO_2$ as their carbon source, grow in sea water or waste water, propagate quickly on otherwise non-arable land and use sunlight as an energy source (52).

## Overview of *C. reinhardtii* properties that restrict its utility as a model organism

Despite having a rich history of successful use to study a wide range of biological processes, there are significant drawbacks associated with the utilization of *C. reinhardtii* as a model organism. Two of the most beneficial features of more malleable model organisms such as *E. coli* and yeast are that they can be transformed with a number of self-replicating, high copy number plasmids, possessing diverse transgene expression control features, and their genomes can be easily manipulated using homology based targeting methods (53,54). These capabilities make it relatively simple to study genes of interest and to conduct tasks such as protein overexpression. To date, no DNA elements have been identified that allow for autonomous replication of plasmids in the *C. reinhardtii* nucleus (55). Rather, when DNA is introduced into the nucleus, the vast majority of

transformation events that occur are the result of direct integration of the transforming DNA into the nuclear genome in a random, non-homologous manner (30,31). Foreign DNA can be incorporated into the nuclear genome via homologous recombination but the ratio of non-homologous to homologous integration events approaches 1,000:1, at best (56,57). Furthermore, DNA introduced into *C. reinhardtii* cells is often fragmented, rearranged or extensively truncated before integrating into the nuclear genome (30,58). Conversely, when a plasmid containing an appropriate origin of replication is introduced into *E. coli* or yeast cells, in virtually all of the resulting transformants the structural integrity of the plasmid is not affected. Therefore in *C. reinhardtii*, extensive screening is often required following transformation to isolate suitable transformants that possess the intended transgenic DNA. In addition to these difficulties, transgenes that have been integrated into the nuclear genome are often prone to epigenetic silencing and position effects, rendering overexpression of proteins or long term transcriptional stability problematic in many cases (59,60). And finally, although many forms of experimentation do not require high transformation efficiency, it is highly preferred when it becomes necessary to screen large numbers of individual transformants or transgenes. The highest transformation rates that can be achieved in yeast and *E. coli* approach $10^8$ or $10^{10}$ CFU/ µg of DNA respectively (61,62). However, the maximum *C. reinhardtii* transformation rate that can be achieved using standard electroporation conditions is only ~$10^5$ CFU/ µg of DNA (30). Utilization of the glass bead or particle bombardment methods to transform *C. reinhardtii* with foreign DNA

yields considerably lower transformation efficiencies (29,31). Therefore, ambitious screening projects are difficult to conduct in *C. reinhardtii*.

## Utilization of polycistronic genes in eukaryotic organisms

In light of the information presented above, it is obvious that there is much potential for the development or optimization of methodologies which would expand the range of experimentation that can be conducted using *C. reinhardtii*. One of the most significant barriers is lack of efficient and reliable methods for transgene expression. One simple strategy that can be employed to overcome this limitation is fusion of the coding regions of a POI and a selectable marker to help ensure that the POI is expressed in all cells expressing the selectable marker protein. However, this technique is limited by the possibility that the activity of one or both proteins will be disrupted by fusion to another protein. In addition, both proteins must share the same subcellular localization patterns to achieve optimal functionality. Protease cleavage sites can also be included between POI and selectable marker proteins to enable post-translational separation. However, this requires that an appropriate protease also be expressed at the correct time and co-localize properly with its target. Internal Ribosomal Entry Sites (IRES) are natural genetic elements that also allow for the production of multiple proteins from one mRNA in eukaryotic organisms. These relatively large ~100-600 bp DNA elements are derived from various genes of a diverse range of viruses and eukaryotic organisms and, when transcribed into mRNA, facilitate initiation of translation independent of the canonical 5` Cap dependent eukaryotic translation initiation mechanism. Proteins translated from IRES sequences are free to localize independently and do not

contain extraneous amino acid residues at either end (63). IRES elements however can be inefficient and difficult to utilize due to their relatively large size (64). Of the dozens of IRES that have been identified, many perform well in some circumstances but still more are highly inefficient at initiating translation as compared to the efficiency with which translation initiates from canonical translation start sites (65). Many IRES are subject to host dependent down-regulation as well (63). A more ideal polycistronic mechanism would allow for the production of multiple, minimally altered proteins at roughly equivalent levels in eukaryotic organisms. Additionally, the resultant proteins must be able to localize independently once translated and the polycistronic mechanism should be mediated by relatively small elements that function efficiently in a broad range of cell types and organisms.

## Viral 2A peptides

The viral 2A peptide system is another natural mechanism that enables the generation of polycistronic genes that function efficiently in eukaryotic organisms. This mechanism relies on small ~19 amino acid domains called 2A peptides and is primarily employed by a broad range of viruses representing a wide range of host specificities (66). Through the study of the foot and mouth disease virus (FMDV), the 2A peptide mechanism was first observed and described. The genomes of many viruses, including FMDV, encode complex polyproteins with many functional

domains that are ultimately separated by various means. With respect to FMDV, it was discovered that 2A peptides were responsible for the separation of adjacent peptides into individual proteins (67). When these 2A peptides were taken out of their native context and placed between marker proteins, they exerted the same effect (Fig. 1-2) (68). Initially it was expected that they mediated a form of post-translational auto-proteolysis. However, subsequent investigations have ruled out this mode of action. Rather, when 2A peptides are translated, they catalyze what has been termed a 'ribosomal skip' event, resulting in the separation of polyprotein peptides co-translationally. When the translation complex reaches the end of a 2A domain, this peptide is able to mediate the release of the N-terminal nascent peptide chain from the ribosome while ensuring that translation of the remaining mRNA sequence proceeds normally until a stop codon is reached. This separation event, or rather failure to form a peptide bond in the ribosome, occurs between the last two amino acids of 2A peptides, which are always Glycine-Proline, that overlap with the highly conserved NPGP motif (69). Because of this, after translation, all proteins that reside upstream of a 2A peptide retain all but the last residue of the peptide as a C-terminal extension while downstream proteins necessarily contain an N-terminal proline residue. Furthermore, this mechanism is heavily dependent on eukaryotic translation-terminating release factors eRF1 and eRF3 (70). Altering the expression levels or activities of these factors leads to highly aberrant 2A processing patterns. Likely due to this mechanistic feature, 2A peptides do not mediate ribosomal skipping in prokaryotes (71).

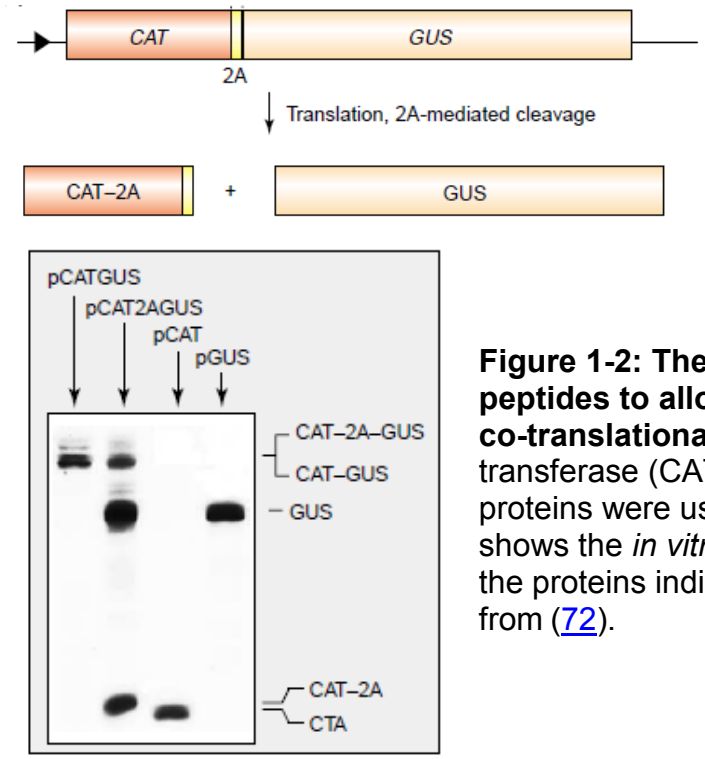



**Figure 1-2: The first example of the use of 2A peptides to allow for separation of marker proteins co-translationally**. The chloramphenicol acetyl-transferase (CAT) and b-glucuronidase (GUS) marker proteins were used in this experiment. Panel inset shows the *in vitro* translation profiles of genes encoding the proteins indicated above each lane. Figure adapted from (72).

Utilization of the 2A peptide mechanism to produce functional polycistronic genes for use in eukaryotes has many advantages. First, because 2A peptides are small, the residual 2A peptide and proline motifs that necessarily remain attached to the C and N termini of proteins separated by a 2A domain are relatively unlikely to cause deleterious effects to fusion proteins. Second, because the 2A mediated separation event occurs during translation, proteins flanking a 2A peptide are free to localize independently of one another (73). Third, unlike with IRESs, *in vivo* translation of 2A containing polyproteins often leads to nearly equivalent production of all 2A-separated proteins. And finally, incorporation of numerous 2A peptides into a single polyprotein enables the production of complex, multicomponent

polycistrons that can greatly simplify tasks such as metabolic engineering in eukaryotic organisms. Accordingly, 2A peptides have been successfully used in a plethora of biotechnological applications since their initial discovery (72,73). For example, in one study three different 2A peptides were used to construct a functional tetracistronic gene that efficiently enabled the production and proper localization of 4 membrane associated proteins from a single mRNA (74). In another important study, a 2A peptide was used to separate the heavy and light chains of a therapeutic monoclonal antibody possessing anti-tumor properties. The resulting distronic gene was successful at promoting significant long term tumor suppression in mice (75). Because of the relative under development of transgenic tools available for use in *C. reinhardtii*, implementation of 2A peptide technology in this organism may alleviate significant difficulties associated with transgene expression, one of the most valuable and elementary biotechnological tools yet developed. Thus, taking advantage of the features of the 2A system should significantly expand the utility of *C. reinhardtii* as an organism that already possesses many beneficial traits which facilitate scientific exploration.

## Mutagenesis and gene targeting

One of the most valuable tools for studying gene function consists of generating genetic mutants followed by analysis of the resulting phenotypic alterations. This process has contributed immensely to our current understanding of many biological processes. Accordingly, many different methods have been devised for the

generation of mutants in a given organism. Perhaps the most facile and controlled means of generating mutants relies on homologous recombination to modify pre-determined genomic loci with engineered donor DNA fragments. However, as previously mentioned, site specific modification of the *C. reinhardtii* nuclear genome is highly inefficient (56). The same characteristic can be attributed to many other important eukaryotes such as plants (76), mammals (77) and fungi (78). In all of these organisms, foreign DNA that is introduced into the nucleus is integrated into the genome non-homologously in the vast majority of transformation events, regardless of whether or not the transforming DNA contains sequences homologous to endogenous genomic loci. Conversely, in yeast, the vast majority of transformation events that occur upon introduction of such fragments into cells are mediated by homologous recombination (Fig. 1-3) (53). This advantageous feature has allowed for the production of vast and complete gene knock out libraries of all non-essential genes in yeast (79). In addition, this capability greatly facilitates the study of more subtle genetic mutations such as domain replacements or site-directed point mutations.

Because the ability to precisely edit the genomic DNA of an organism is so valuable, much effort has been expended developing technologies aimed at increasing the efficiency of this process in organisms where it is naturally inefficient. It was previously discovered that the introduction of double stranded DNA breaks at a particular genomic locus greatly increases the chances that homologous recombination will be used to repair the damage (80). Indeed, this natural phenomenon is exploited by many organisms to generate site specific genetic

modifications for various biological purposes (81). Over the past decade, appreciation of this process has led to the development of various forms of custom nucleases possessing programmable DNA recognition domains which can be used to increase the efficiency of gene targeting in higher eukaryotes. The first successful custom nucleases developed were Zinc Finger Nucleases (ZFNs) that consist of customizable zinc finger DNA binding domains fused to the non-specific nuclease domain of the type IIs restriction enzyme FokI (82,83). This technology however has recently given way to Transcription Activator Like Effector Nucleases (TALENs) (84) and the Clustered Regularly Interspaced Palindromic Repeat (CRISPR)/Cas9 system (85) due to the greater ease with which these nucleases can be programmed to recognize and digest predetermined DNA sequences. However, all of these systems have remained difficult to implement in *C. reinhardtii* despite the expenditure of great effort by numerous groups experienced with this organism. In addition, utilization of these tools to generate each novel mutant requires the synthesis of novel targeting constructs and extensive screening to isolate valuable mutants from transformant populations containing high proportions of wild-type and otherwise unwanted individuals. Therefore, these tools are more suited for the generation of relatively small numbers of designer mutants rather than large, comprehensive mutant populations. Furthermore, in many situations, the genes responsible for a given phenotype are not yet known, necessitating the execution of a random mutagenesis approach followed by screening for disruption of the phenotype under study. None of the custom nuclease tools described above are suited for this purpose.

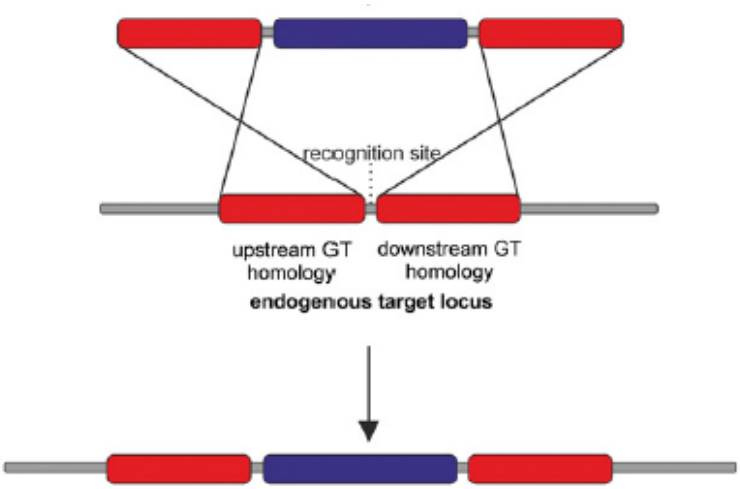


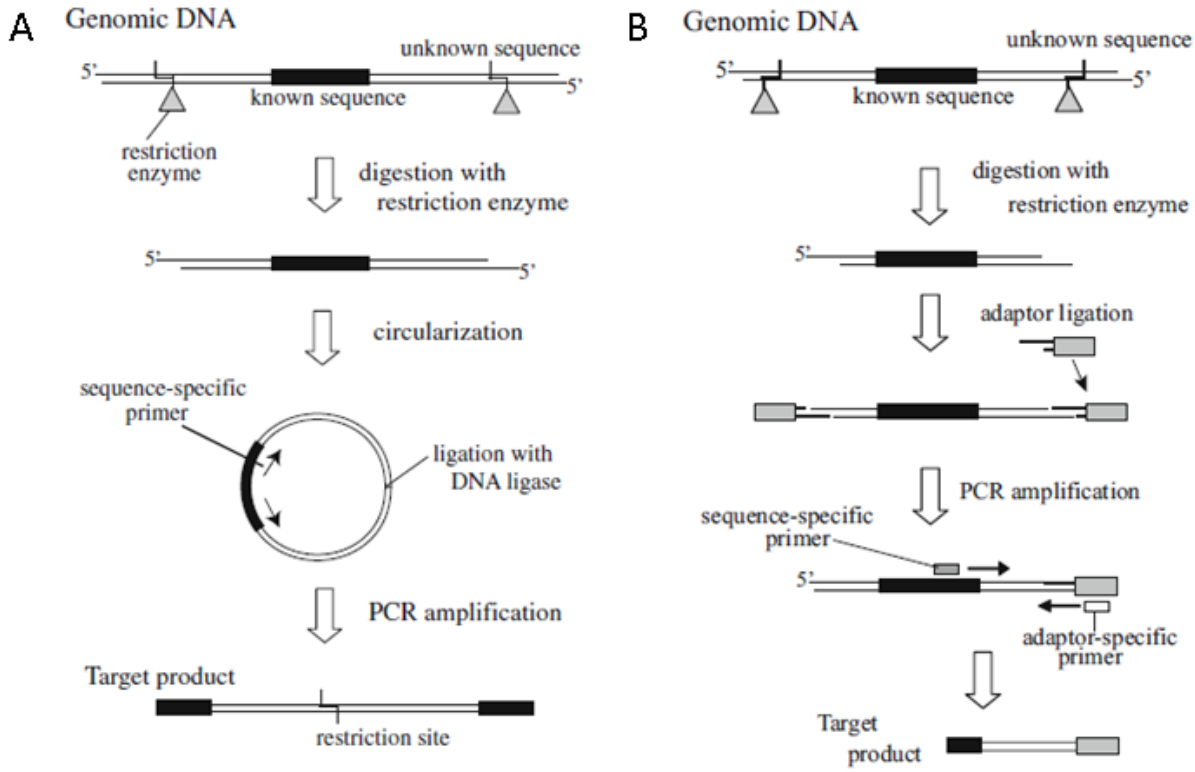

**Figure 1-3: Homologous recombination mediated genomic modification**. When double strand breaks are introduced into genomic DNA in the presence of a donor DNA (Top, blue rectangle flanked by red rectangles) bearing flanking regions (red) homologous to the targeted genomic locus (red), the frequency at which conversion of genomic DNA occurs is elevated. Adapted from (86).

## Insertional mutagenesis

Due to the practical difficulty of achieving gene targeting in *C. reinhardtii,* reliance on other forms of mutagenesis has become necessary. Although many forms of chemical or abiotic mutagenesis exist, most of them lead to the generation of multiple subtle mutations such as single base substitutions, insertions or deletions throughout the genome of each individual treated cell (87). This makes identification of the genetic elements responsible for a particular phenotypic disruption difficult to accomplish. Accordingly, another mutagenesis method has been developed that takes advantage of the randomness with which transgenes are integrated into the genomes of higher eukaryotes and prokaryotes when certain transformation procedures are used (88). This process is known as insertional mutagenesis and results in the production of populations of transformants bearing,

in ideal situations, only one copy of a marker gene randomly inserted in their genomes. Often these insertion events significantly affect the integrity and thus function of disrupted genetic elements, enabling their functional significance to be studied. Furthermore, because a marker DNA of known sequence is used to generate mutants, various methods that fall into the category of genome walking have been devised that utilize this sequence to ultimately identify the genomic regions flanking marker insertion sites (89). The three major methodologies used to accomplish this task are depicted in Figure 1-4. As a testament to its value, insertional mutagenesis has been extensively implemented for the generation of insertional mutant libraries in many important model organisms including plants (*Arabidopsis* (90), rice (91), corn (92)), yeast (93), fungi (94,95), bacteria (96), mammals (97), fish (98) and insects (99). Naturally, this technique has also been widely used for the study of *C. reinhardtii* (58,100). However, comprehensive insertional mutant libraries have not yet been made available in this organism. In addition, methods currently and traditionally used to generate such libraries in various organisms suffer from the unfortunate reality that numerous insertion events generated by transformation with complete marker genes containing promoters occur in regions, such as intergenic or putative promoter sequences, that are not expected to significantly contribute to the function of any genetic element or lead to significant phenotypic disruptions when mutated (101). In *Arabidopsis* for example, analysis of an extensive *Agrobacterium*-generated T-DNA insertion population consisting of ~88,000 individuals revealed that roughly 60% of all insertion events had occurred in such regions (90). This drawback increases the effort required to

generate and isolate valuable mutants. Consequently, numerous forms of gene trapping methodologies have been adopted for the generation of insertional mutants that bias insertion events towards transcribed or functionally relevant regions of a genome (101,102). As mentioned previously, numerous genome walking methods exist that for the most part depend on PCR based methods to amplify genomic sequences flanking known marker regions. Analyzing the sequence of these products enables marker insertion sites to be identified (89). Perhaps because no single method has yet proven to be highly efficient and broadly applicable, dozens of different genome walking methods have been developed, with new variants still being generated (103,104). Most of these methods however are only able to amplify genomic DNA sequence flanking one side of an inserted marker per reaction and often produce non-specific background products (89). Therefore, progress can still be made in the development of more facile techniques for characterizing insertional mutants.



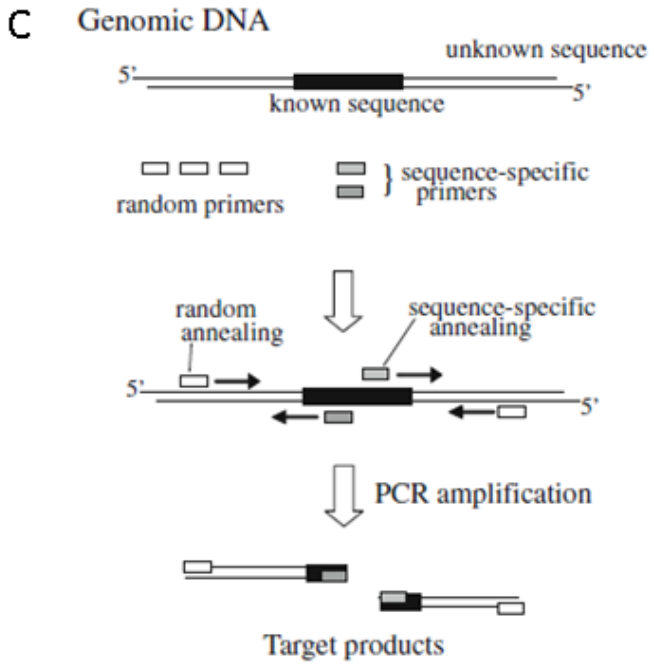




**Figure 1-4: The three major methodologies used to localize marker insertion sites in insertional mutants**. **A)** Inverse PCR. Genomic DNA is restricted and ligated to generate circular PCR templates which are amplified with two marker specific primers. **B)** Adapter-ligation mediated PCR. Genomic DNA is digested and ligated to a known adapter molecule. Flanking regions are amplified by marker and adapter specific primers. **C)** Thermal asymmetric interlaced PCR (TAIL-PCR). Semi-random primers are mixed with marker specific primers to amplify genomic regions flanking known sequences. Adapted from (105).

Although insertional mutagenesis is an important and sometimes essential tool that has been used for decades to study biological processes, there are still ways in which it can be improved in some organisms. One limitation stems from the nature of transformation events that occur in most eukaryotic organisms leading to random transgene integration throughout the genome. Utilization of procedures or techniques to bias insertion events towards transcribed genomic loci might significantly improve the value of insertional mutant libraries in *C. reinhardtii* especially. Furthermore, many of the techniques that have been devised to characterize insertional mutants can be inefficient and costly. Therefore, there is still a need to develop more efficient methods for characterizing genetic lesions in insertional mutant populations.

# CHAPTER 2: Development of an improved and versatile viral 2A platform for expression of polycistronic nuclear genes in *Chlamydomonas reinhardtii*

## Abstract

A greatly expanded demonstration of the efficiency and utility of the viral 2A system for expression of dicistronic and polycistronic genes in *Chlamydomonas reinhardtii* is presented. In particular, data are provided demonstrating high efficiency in producing two separate proteins from dicistronic genes in which the two genes are separated by viral 2A peptide coding regions, including an especially efficient “extended” FMDV 2A sequence. Confocal fluorescence microscopy is used to show that remnants of the 2A peptide sequences at the N-terminus or C-terminus of proteins produced from 2A-containing dicistronic genes does not interfere with proper subcellular localization of native and foreign proteins. Importantly, data are also presented that demonstrate the utility of the 2A system in allowing reliable, high-level expression of foreign genes in *Chlamydomonas*. Furthermore, to expand the versatility of the 2A expression system, many commonly used selectable marker proteins were assessed for their compatibility with the extended FMDV 2A peptide. Additional experiments highlight the ability to use the 2A system to couple low expression marker proteins with a strong conditional *Chlamydomonas* promoter to allow for facile delivery of genes of interest whose expression need to be regulated to optimize cellular economy or mitigate toxic effects that constitutive high level expression may cause. Finally, a demonstration of the 2A system to allow for the production of tricistonic genes in *Chlamydomonas* is presented. Together, the studies presented here demonstrate the power and utility of the 2A system to allow dependable, high-level expression of transgenes in *C. reinhardtii*, a capacity heretofore lacking in this important algal model system.

## Introduction

The unicellular green alga *Chlamydomonas reinhardtii* (*Chlamydomonas* hereafter) has long served as a model organism for studies of photosynthesis (106), cellular motility (45) and, more recently, biofuel production (51). Among its useful features are its publicly available nuclear genome sequence and the ability to transform all three of its genomes (nuclear, mitochondria and chloroplast) with foreign DNA (8,107). Although transformation of the nuclear genome is possible with relatively high efficiency (30,31), there are a number of difficulties associated with the integration and expression of foreign DNA. First, transgenes are integrated randomly. This leads to a high degree of expression variability due to position effects (31). Second, transformation efficiency is low compared to other organisms such as *Saccharomyces cerevisiae* and *E. coli* (30,31). Using electroporation, approximately 1 in 25,000 cells can be transformed in a stable manner leading to a maximum CFU/µg of DNA of ~$10^5$ (30). Third, transgenes are often fragmented or rearranged prior to or during integration. This often necessitates extensive screening to identify desired transformants for further study (30,58). Fourth, transgenes are commonly subject to silencing (108,109). This results in unreliable long-term expression of genes of interest (GOI). One method that has been utilized to address some of these difficulties is to fuse the coding regions of a GOI to a selectable marker gene such that one promoter is used to drive expression of one fusion protein (110). This method increases the chances that primary transformants will express the desired GOI, provides more reliable gene expression over time and reduces the amount of screening required to identify suitable transformants.

However, there is substantial risk that fusion of a GOI to a marker gene will hinder the function of the GOI, the marker or both.

Viral 2A peptides represent another mechanism by which two or more genes can be translated from a single di- or poly-cistronic mRNA. The 2A system is utilized by several different viral species whose hosts include a broad variety of eukaryotic organisms (66,69,70). Most minimal 2A peptide domains are approximately 19 amino acids in length and are used to separate functional protein domains in viral polyproteins (66). 2A peptides catalyze an event termed a 'ribosomal skip' during translation that results in the production of two independent peptides (70). Following this event, the N-terminal protein partner retains at its C-terminus all of the 2A amino acids except the C-terminal residue, which is always a proline. This proline remains covalently attached to the N-terminus of the downstream protein (66,69). 2A peptides have been shown to function in all eukaryotes tested to date (72); however, they do not function in prokaryotic organisms (71). This is because eukaryotic-specific translation termination cofactors appear to be required for the 2A mechanism (69,70). Various 2A peptides catalyze the ribosomal skip process with different efficiencies (66). The most active 2A peptides catalyze this reaction with at least 90% efficiency, resulting in the production of near stoichiometric levels of each protein product (66,72). In addition, because the 'skip' mechanism occurs during translation, the proteins separated by a 2A peptide are free to localize independently. However, successful targeting to specific subcellular locations is not always achieved (72). In regard to *Chlamydomonas*, utilization of the 2A system offers several particular advantages for simultaneous expression of multiple GOIs:

1 – The 2A system allows fusion of a GOI to a selectable marker which may drastically reduce screening and make gene expression more reliable; 2 - Depending on the GOI and marker combination, relatively significant levels of GOI overexpression can be obtained; 3 - Marker and GOI proteins have the potential to localize independently; and 4 – Because polycistronic expression cassettes containing 2, 3 or 4 GOIs have been shown to function well in other organisms (72,74) utilization of similarly complex polycistronic genes might be feasible in *Chlamydomonas*. Currently, the logistical difficulty of simultaneously expressing in a reliable manner multiple GOIs that do not confer a selectable phenotype in *Chlamydomonas* precludes most attempts to achieve such a goal.

We and others (111,112) have provided preliminary reports of the utility of the 2A system for gaining expression of dicistronic messenger RNAs in *Chlamydomonas*. Here we provide a much expanded characterization of the 2A system in *Chlamydomonas* including quantitative analyses and comparison of two 2A peptides, a comparison of the efficiency with which the 2A system and traditional adjacent selection achieve expression of a GOI as well as an examination of the compatibility of the 2A system with a number of selectable marker genes. In addition, we demonstrate the utility of inducible promoters in conjunction with low expression selectable markers for achieving low constitutive expression and allowing for transient over-expression of various dicistronic 2A constructs. As a final addition to the current *Chlamydomonas* 2A toolbox, we present evidence that with proper design considerations, tricistronic 2A constructs

can be produced that are capable of simultaneously expressing three separate and functional proteins in an inducible fashion.

## Results and Discussion

### *Chlamydomonas* processes 2A peptides efficiently and an extended form of the FMDV 2A peptide is processed in a nearly quantitative manner

To determine if various 2A peptides are properly processed in *Chlamydomonas* and, if so, how efficiently they are processed, a series of marker genes was constructed consisting of the bleomycin resistance gene ($Ble^r$) and the fluorescent mCherry protein separated by a 2A peptide coding region (Fig. 2-1A, Fig. 2-2). 2A peptides are highly varied in their amino acid sequences and the efficiency with which they are processed in various eukaryotic organisms (66). The most commonly used 2A peptides are usually about 19 AAs in length. It has been shown, however, that the inclusion of a longer native N terminal region can improve 2A performance (66). The 2A peptide used in previous *Chlamydomonas* studies (111,112) was based on the minimal FMDV 2A peptide which has been shown to be as much as 45-50% less efficient than larger FMDV 2A peptides (66). Thus, an extended version of the FMDV 2A peptide was synthesized and evaluated (Fig. 2-1B). Moreover, we reasoned that when designing complex 2A-based polycistronic genes, it may be beneficial to use multiple dissimilar 2A peptides to minimize the chances of recombination occurring between their respective DNA coding regions during or after transformation (74,113,114). Therefore, another commonly used 2A peptide derived from the Thosea asigna Virus (TaV 2A) (72), which has been

reported to promote efficient 'ribosolmal skip' activity (66) was also chosen for evaluation and for comparison to the extended FMDV 2A.

The Ble$^r$ and mCherry markers were chosen to assess the efficacy of the 2A mechanism in *Chlamydomonas* for a number of reasons. First, they both produce easily detectable phenotypes. The Ble$^r$ protein allows for selection of transformants resistant to the antibiotic zeocin (115) while the mCherry protein is easily detectable via fluorescence microscopy. In addition, imaging of fluorescent proteins that emit maximally in the red range leads to less background from chlorophyll autofluorescence in photosynthetic organism compared to the more typically employed GFP (116). Second, both of these proteins have been shown to tolerate various fusion partners (110,116). Third, their native localization patterns in eukaryotic cells are different. The Ble$^r$ protein and Ble$^r$/GFP chimeras have been shown to localize predominantly to the nucleus of *Chlamydomonas* (110) while RFP-like proteins such as mCherry are found scattered throughout both the cytoplasm and nucleus when they lack a localization signal (117). And finally, the Ble$^r$ protein is required in relatively high levels because it must form 1:1 complexes with the antibiotics bleomycin or zeocin to neutralize their ability to create DNA double strand breaks. Accordingly, fusion of a marker protein such as mCherry to Ble$^r$ will ensure that its expression level is adequate for detection. If these two markers are separated by a properly functioning 2A peptide, mCherry will be detected in the cytoplasm and nucleus. Conversely, if a particular 2A peptide is not processed properly, the vast majority of the mCherry signal will be detected in the nucleus.

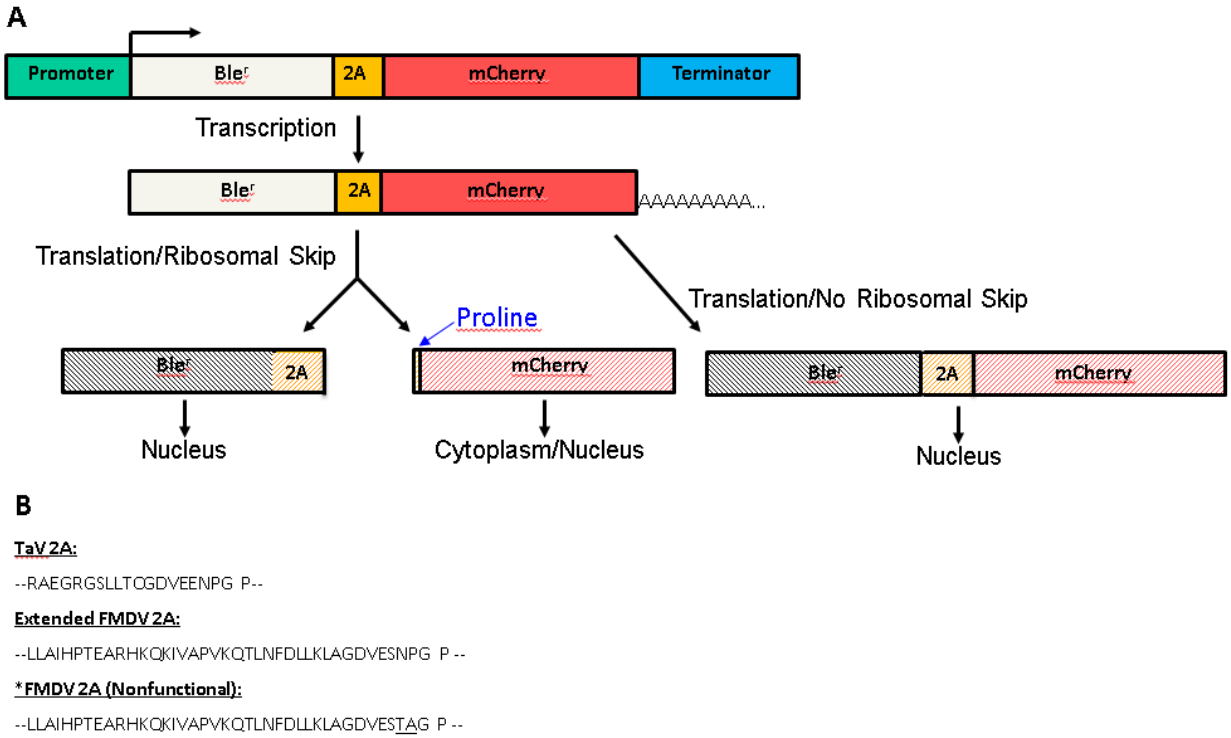



**Figure 2-1: Strategy for verification of 2A function in *Chlamydomonas*. A)** Dicistronic, 2A-containing construct used to assess 2A function *in vivo*. If the mRNA trasnscribed from the Ble$^r$-2A-mCherry discistronic gene is translated into a polypeptide that is cleaved when it reaches the 2A region, separate Ble$^r$ and mCHerry proteins will be produced with the Ble$^r$ protein migrating solely to the nucleus and the mCherry protein occupying both the cytoplasm and the nucleus. Failure of cleavage at the 2A site during translation will produce a Ble$^r$-2A-mCherry chimera that will reside exclusively in the nucleus. **B)** Amino acid sequences of 2A peptides used in this study. TaV 2A: derived from the Thosea asigna Virus (TaV). Extended FMDV 2A: derived from the Foot and Mouth Disease Virus (FMDV). *FMDV 2A: the extended FMDV 2A made nonfunctional by replacement of two amino acids (NP) essential for 2A activity with noncanonical amino acids (TA).

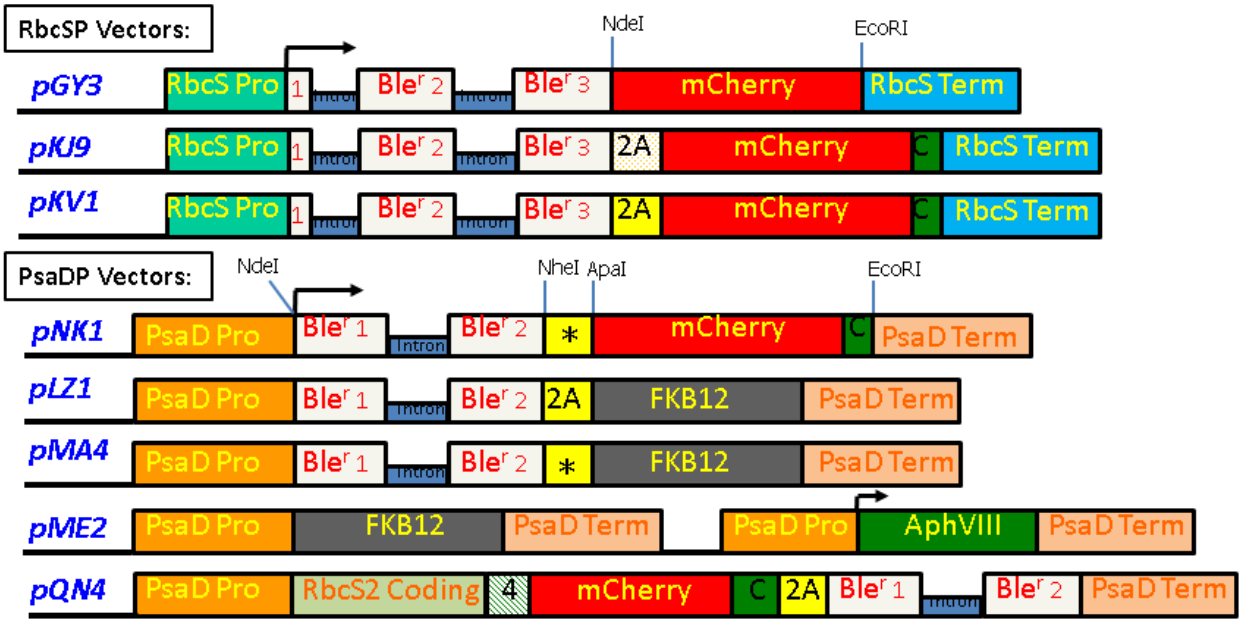



**Figure 2-2. Gene constructs used in the present study.** Genes in RbcSP vectors are driven by the *Chlamydomonas Rubisco small subunit 2* gene promoter and ended with the RbcS2 gene termination region. Genes in PsaDP vectors are driven by the *Chlamydomonas PsaD* gene promoter promoter and ended with the *PsaD* gene termination region. Ble$^r$, bleomycin resistance gene exons divided by one or two introns; mCherry, mCherry fluorescence protein gene; C, coding region for *CIA5* C-terminal peptide epitope, 4, coding region for four repeats of the flexible peptide, EAAAR; FKB12, coding region for the *Chlamydomonas FKB12* gene; *RbcS2* Coding, coding region of the *Rubisco small subunit 2* gene; AphVIII, paromomycin resistance gene driven by the *Chlamydomonas PsaD* gene promoter.

Indeed, when the Ble$^r$ and mCherry proteins were fused without the incorporation of a 2A peptide (Fig. 2-2, pGY3), the overwhelming majority of the mCherry signal was detected in the nucleus (Fig. 2-3), as was previously observed with a GFP/Ble$^r$ chimeric protein (110). When the TaV and FMDV 2A peptides were inserted between the Ble$^r$ and mCherry proteins (Fig. 2-2, pKJ9 and pKV1, respectively), mCherry signal was detected in abundance in the cytoplasm

indicating that the 2A peptides were being processed properly and that the two marker proteins were localizing independently as intended (Fig. 2-3). The mCherry protein remaining in the nucleus of these transformants could have either been unprocessed full length Ble$^r$-2A-mCherry protein or free mCherry protein that diffused through the nuclear pores from the cytoplasm due to its relatively small size (~27 kDa).

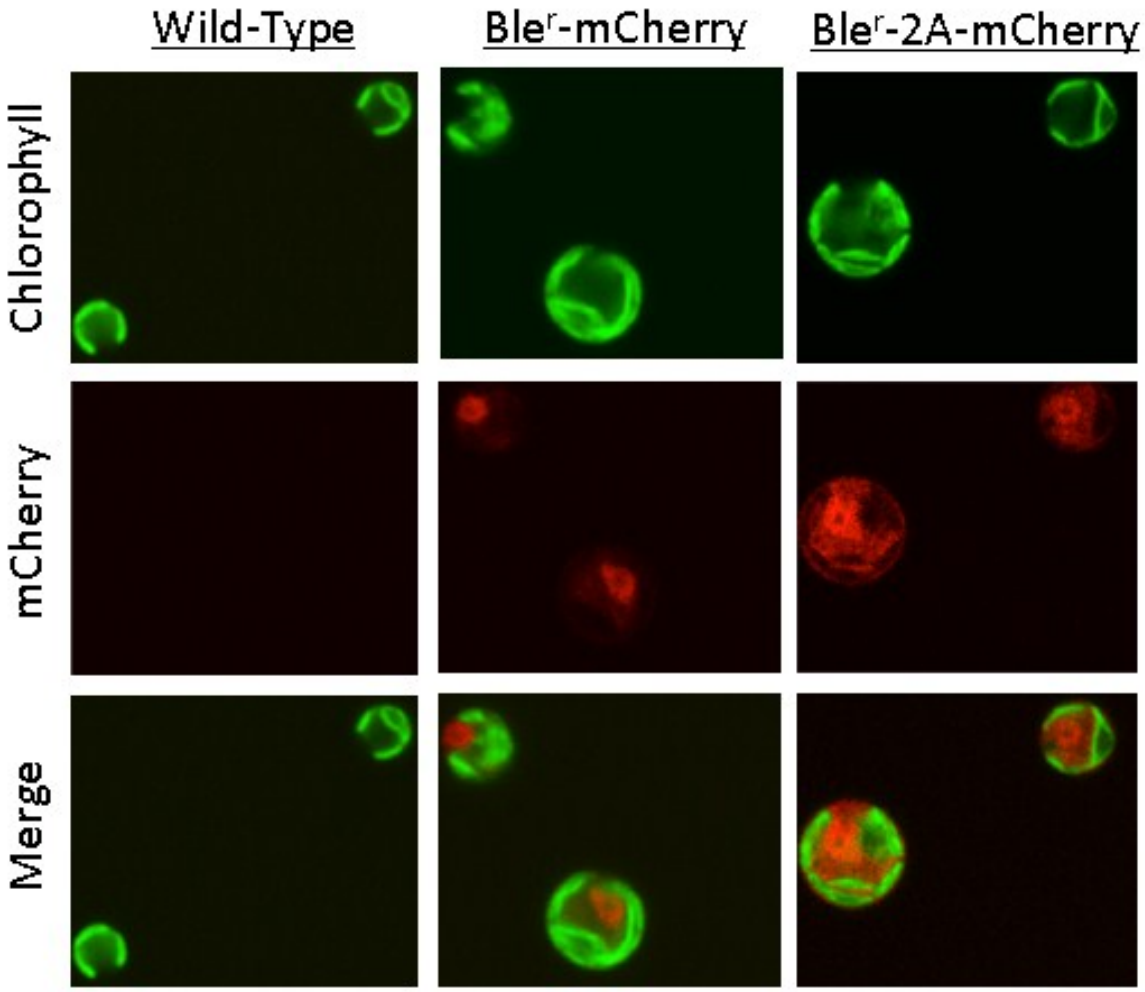



**Figure 2-3: Fluorescence imaging of *Chlamydomonas* cells**. WT cells were transformed either with a Ble$^r$-mCherry gene construct (center panels) or a Ble$^r$-TaV 2A-mCherry gene construct (right panels). Chlorophyll autofluorescence is depicted in green and mCherry fluorescence in red.

To confirm these visual observations at the protein level and assess the processing efficiency of the extended FMDV 2A peptide compared to the minimal TaV 2A peptide in *Chlamydomonas*, an epitope tag derived from the C-terminus of the *Chlamydomonas* transcription activator, CrCIA5 (118), was added to the C terminus of the mCherry protein to allow for its detection during protein blot analysis (Fig. 2-2, EpCIA5). Previously, a highly sensitive and specific polyclonal antibody was raised against this epitope (118). In addition to the sensitivity of the CrCIA5 antibody, an additional advantage of using this epitope is that the CIA5 protein is constitutively expressed under standard growth conditions and can serve as an internal loading control during protein blot experiments (118,119). Utilization of this epitope to tag the mCherry protein enabled detection of processed and unprocessed mCherry protein species. Accordingly, protein blot analysis of proteins extracted from a number of pKV1 transformants confirmed by fluorescence microscopy to be expressing mCherry revealed in 10 out of 11 cases distinct and prominent bands corresponding in size to free, tagged mCherry and far weaker bands corresponding in size to unprocessed full length Ble$^{r}$-2A-mCherry (EpCIA5) (Fig. 2-4A). Although precise quantification of the 2A peptide processing efficiency in this experiment was obscured because several lanes contained an overabundance of mature mCherry protein, estimates from more lightly loaded lanes indicate processing efficiencies of the extended FMDV 2A peptide approaching 100% (Fig. 2-4A). In comparison, cleavage efficiency of the TaV 2A peptide in a strain expressing the pKJ9 gene (Fig. 2-4A, Ble$^{r}$-TaV 2A-mCherry; third

lane from the left) was markedly poorer than that observed in lanes containing equivalent loading of products from the FMDV 2A-containing gene (Fig. 2-4A, KV1-1, 3, 4 and 8). No prominent bands other than the native WT CIA5 protein were detected in nontransformed cells or cells expressing a Ble$^{r}$-mCherry fusion protein (Fig. 2-4A, first and second lanes, respectively). Western blot analysis of cells expressing a Ble$^{r}$-FMDV 2A*-mCherry gene (Fig. 2-2, pNK1) in which essential residues in the 2A active domain were mutated (Fig. 2-1B), revealed that all of the detectable mCherry protein corresponded in size to the full length unprocessed protein (Fig 2-4B). The data from these experiments demonstrate that while both the extended FMDV 2A peptide and the TaV 2A peptide are processed well in *Chlamydomonas*, processing of the extended FMDV 2A is nearly quantitative. Therefore, for expression of 2A-containing dicistronic genes in *Chlamydomonas*, the extended FMDV 2A peptide should be the favored choice in most circumstances.

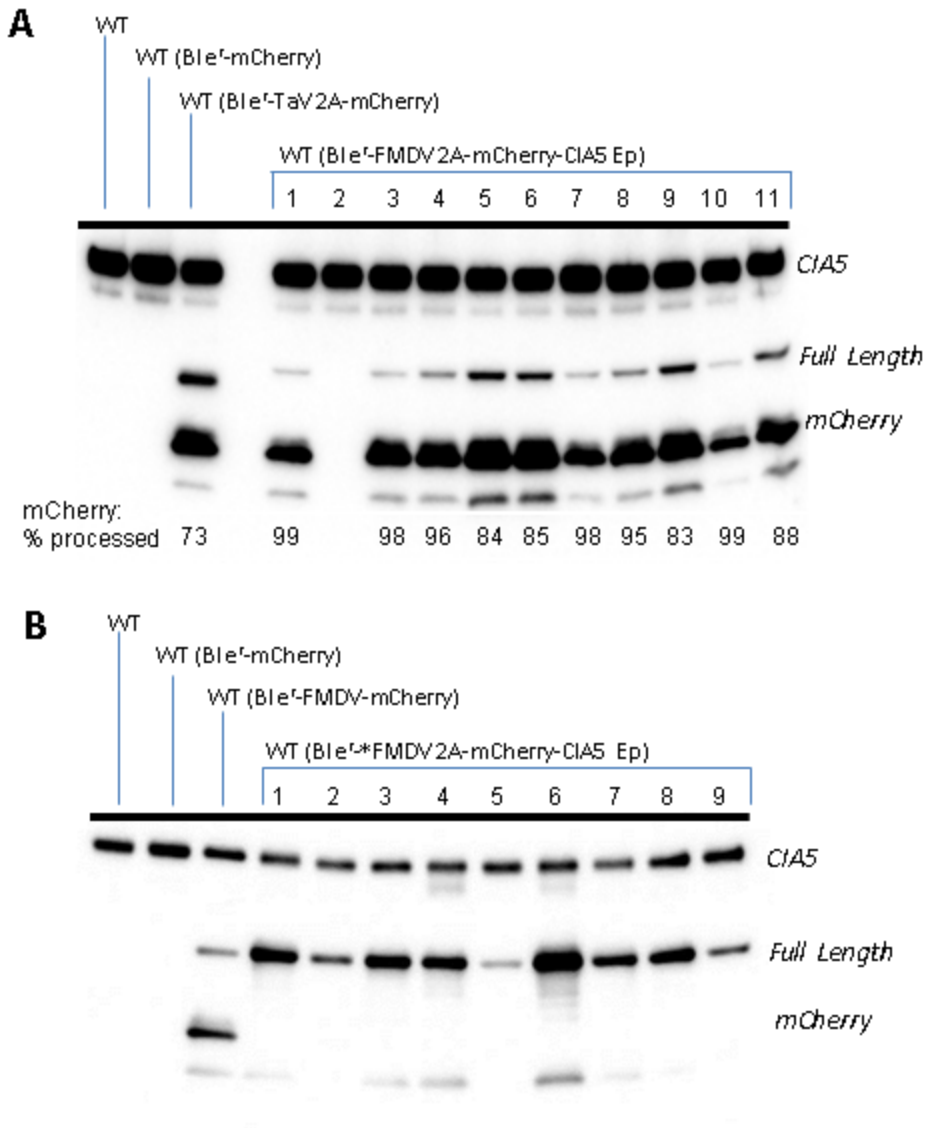



**Figure 2-4: Protein blot analysis of the efficiency of FMDV 2A processing in *Chlamydomonas* cells.** Expected sizes of indicated bands are as follows: CIA5 = ~100 kDa, full length unprocessed Ble$^r$- FMDV 2A-mCherry-CIA5 Ep = 48.6 kDa, mCherry-CIA5 Ep = 30.4 kDa. A) Proteins from transformed cells expressing an mCherry Ble$^r$-FMDV 2A-mCherry-CIA5 Ep gene construct detected on a protein blot using polyclonal antibodies raised against a synthetic CIA5 epitope (CIA5 Ep). B) Proteins from transformed cells expressing an mCherry Ble$^r$-*FMDV 2A-mCherry-CIA5 Ep gene construct producing a mutant, nonfunctional FMDV 2A peptide sequence.

**The 2A expression platform allows expression of GOIs at dependably high levels and significantly outperforms adjacent selection with respect to GOI expression frequency**

Although we have shown that a marker gene can be expressed at easily detectable levels with the Ble$^r$-2A system, it is perhaps more significant to be able achieve dependable and adequate expression of native GOIs at physiologically relevant levels in *Chlamydomonas*. To determine if the highly expressed Ble$^r$ marker is capable of accomplishing this goal, the Ble$^r$-FMDV 2A gene and its mutant counterpart possessing an inactive 2A peptide (FMDV 2A*) were coupled with the *CrFKB12* coding region to allow expression of the CrFKB12 protein *in vivo* (Fig. 2-2 pLZ1, pMA4). The small (~12 kDa) FKB12 protein is a cis-trans prolyl isomerase that assists with, among other things, protein folding (120,121). When the antibiotic rapamycin is present, it forms a complex with FKB12 that, in turn, binds and inhibits the target of rapamycin (TOR) protein. The TOR protein is a highly conserved eukaryotic protein that plays an essential role in coordinating cell growth and development in response to nutrient availability (120,121). Therefore, rapamycin treatment severely inhibits growth in most eukaryotic organisms. In *Chlamydomonas*, the FKB12 protein is nonessential and *fkb12* mutants are resistant to the effects of rapamycin (121). Complementation of *fkb12* mutants therefore cannot be directly selected by restoration of the ability of transformed cells to grow. Rather, complementation is detected by the restoration of sensitivity to rapamycin (121) determined by replica plating of transformants on plates containing media with and without rapamycin. In addition, FKB12 localizes predominantly to the endoplasmic reticulum and cytoplasm and is considered to be a relatively high

abundance protein (121,122). Therefore, this protein offered an opportunity for examining the utility of the 2A expression platform in gaining dependable, high-level expression of a specific native GOI at physiologically relevant levels. The pLZ1 and pMA4 constructs were used to transform an *fkb12* mutant derived from *Chlamydomonas* strain CC3491. Transformants were selected on zeocin and then randomly screened via replica plating for *fkb12* complementation (i.e., restoration of rapamycin sensitivity). A large proportion of randomly selected zeocin resistant transformants screened displayed WT sensitivity to rapamycin (Fig. 2-5). This indicated that the 2A peptide was being processed and that the FKB12 protein was being expressed at functional levels and properly targeted within the cell. When the mutant version of the 2A was used to fuse the $Ble^r$ and the *FKB12* gene (pMA4), restoration of rapamycin sensitivity was not observed, likely due to disruption of the FKB12-rapamycin-TOR complex due to steric hindrance and/or to improper FKB12 localization (Fig. 2-5).

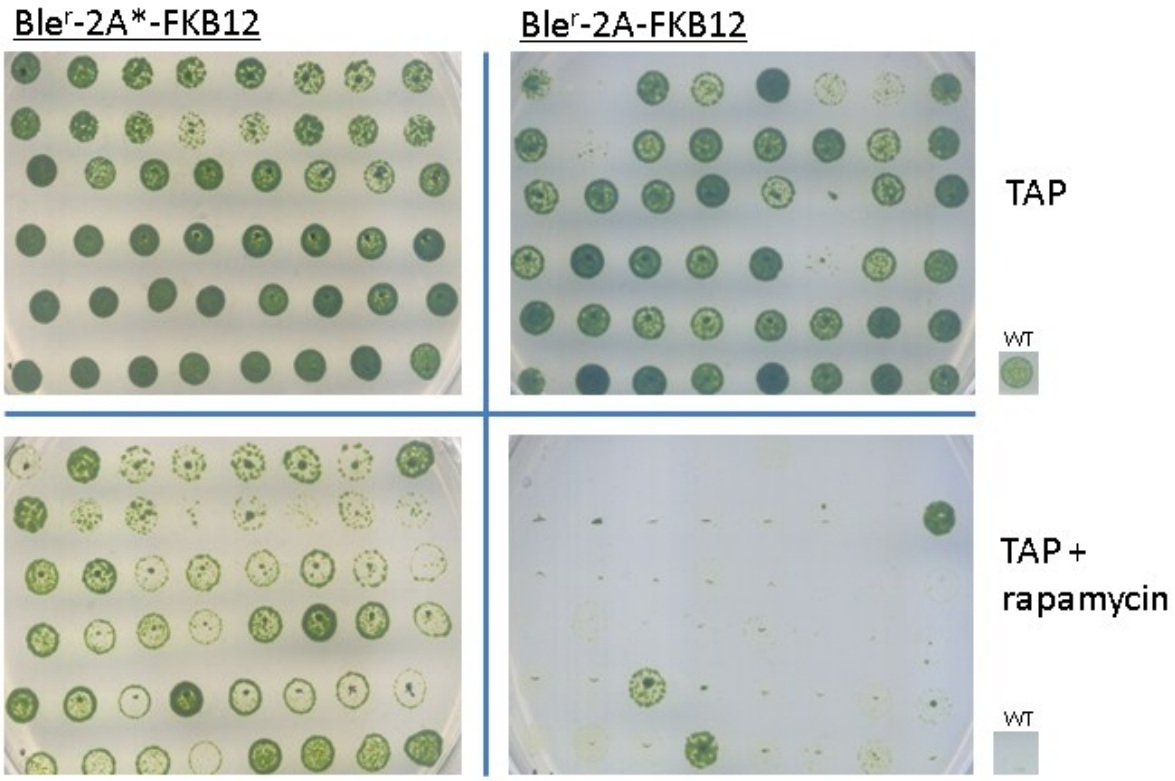



**Figure 2-5: *FKB12* complementation assay**. Zeocin resistant colonies of originally rapamycin-resistant *Chlamydomonas fkb12* mutants transformed with plasmid pMA4 containing a nonfunctional FMDV 2A coding sequence (left column) or with plasmid pLZ1 containing a functional FMDV 2A coding sequence (right column) were randomly picked to TAP liquid medium, mixed and replica plated onto TAP plates (top row) and TAP plates containing rapamycin (bottom row). Successful complementation of *fkb12* mutants with a functional *FKB12* gene results in loss of rapamycin resistance.

Currently, the most commonly employed method for expressing a GOI in *Chlamydomonas* is to locate an independent selectable marker gene adjacent to a GOI expression cassette on a single plasmid that is subsequently used for transformation (30,115,123). This method is referred to as adjacent selection. As previously mentioned, adjacent selection is unreliable and inefficient at achieving

GOI expression, often necessitating labor intensive and tedious screening procedures. To directly compare rates of co-expression of a GOI and a selectable marker gene linked through a 2A peptide coding region to that obtained with adjacent selection, an additional plasmid was constructed containing independent FKB12 and selectable marker gene expression cassettes. The selectable marker gene included in this plasmid was the commonly used *AphVIII* that provides resistance to the antibiotic paromomycin (33) (Fig. 2-2, pME2). Randomly chosen $Par^r$ transformants generated in the *fkb12* mutant were subsequently assayed for *fkb12* complementation. The complementation rate obtained using this adjacent selection construct was approximately 10% while the rate achieved with the 2A expression platform (pLZ1) was approximately 93% (Fig. 2-6).

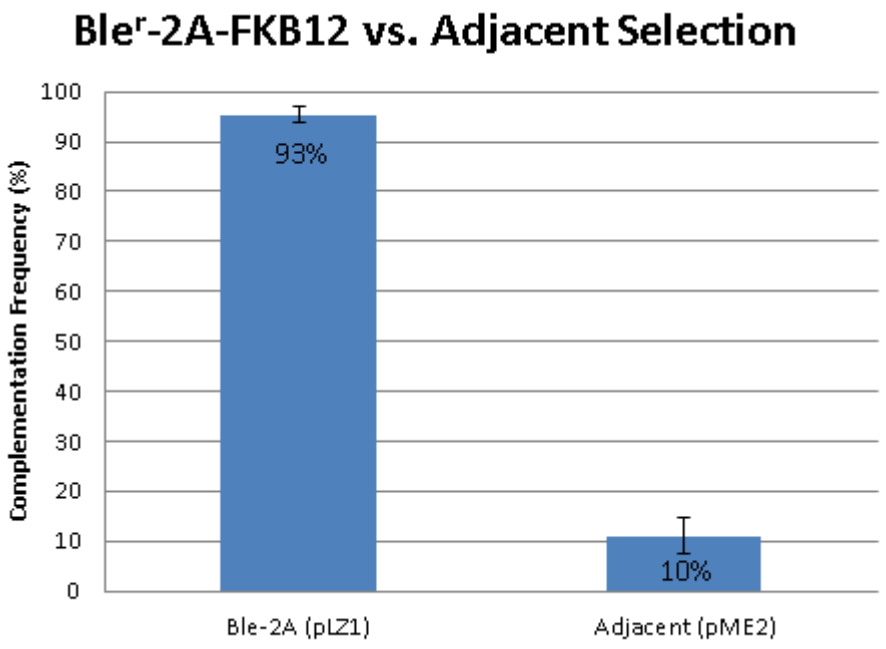



**Figure 2-6: Comparison of complementation frequencies obtained with the 2A system and adjacent selection.** For each of three independent repetitions, at least 96 zeocin or paromomycin resistant colonies were randomly screened for *FKB12* complementation via replica plating onto rapamycin-containing TAP plates.

**<u>The 2A expression platform facilitates subcellular localization studies in *Chlamydomonas*</u>**

One of the most useful and widely implemented applications of fluorescent proteins (FPs) is to serve as localization markers for novel or unstudied proteins of interest. Such studies are difficult to conduct in *Chlamydomonas* however due to the poor control of transgene expression discussed above. By coupling a highly expressed selectable marker such as Ble$^r$ to the FP-tagged POI through a 2A peptide bridge, unreliable POI expression levels resulting from position effects, gene silencing and transgene fragmentation can be mitigated or eliminated. Moreover, because in most cases localization of the 2A fusion partner is completely independent of localization of the selectable marker protein (72), a single highly

expressed marker gene such as Ble can be used to conduct diverse protein localization studies.

To demonstrate the utility of 2A based vectors in allowing precise localization of POIs in *Chlamydomonas*, coding regions for the Rubisco small subunit protein (RBCS2), the CrCIA5 epitope (Fig. 2-2) and mCherry were joined and expressed as a fusion with the FMDV 2A-Ble$^r$ coding regions (Fig. 2-2, pQN4). Previously, it has been shown by immune-gold labeling techniques that RBCS2 localizes solely to the chloroplast and, more specifically, to the pyrenoid under ambient and $CO_2$ limiting conditions (19). Following transformation of WT cells with plasmid pQN4 and selection on zeocin-containing agar plates, randomly selected colonies were screened for mCherry fluorescence. Of 14 colonies examined, 12 (86%) exhibited exclusive localization of mCherry fluorescence to the pyrenoid (Fig. 2-7A). These results should be contrasted with the results depicted in Figure 2-3 demonstrating that the Ble$^r$ selectable marker protein localized almost exclusively to the nucleus while “free” mCherry protein was distributed throughout the nucleus and the cytoplasm. Free mCherry and Ble$^r$ - mCherry protein species were undetectable in the chloroplast in these experiments (Fig. 2-3). To confirm expression and proper processing of the pQN4 translation products, protein blot analysis was conducted with whole cell extracts prepared from four independent mCherry-positive transformants. In all 4 cases, the vast majority of the pQN4 protein product had been efficiently cleaved at the 2A site to yield the expected ~50 kDa RBCS2 chimera (Fig. 2-7B). As expected, only the native CIA5 protein was detected in WT control cells (Fig. 2-7B).

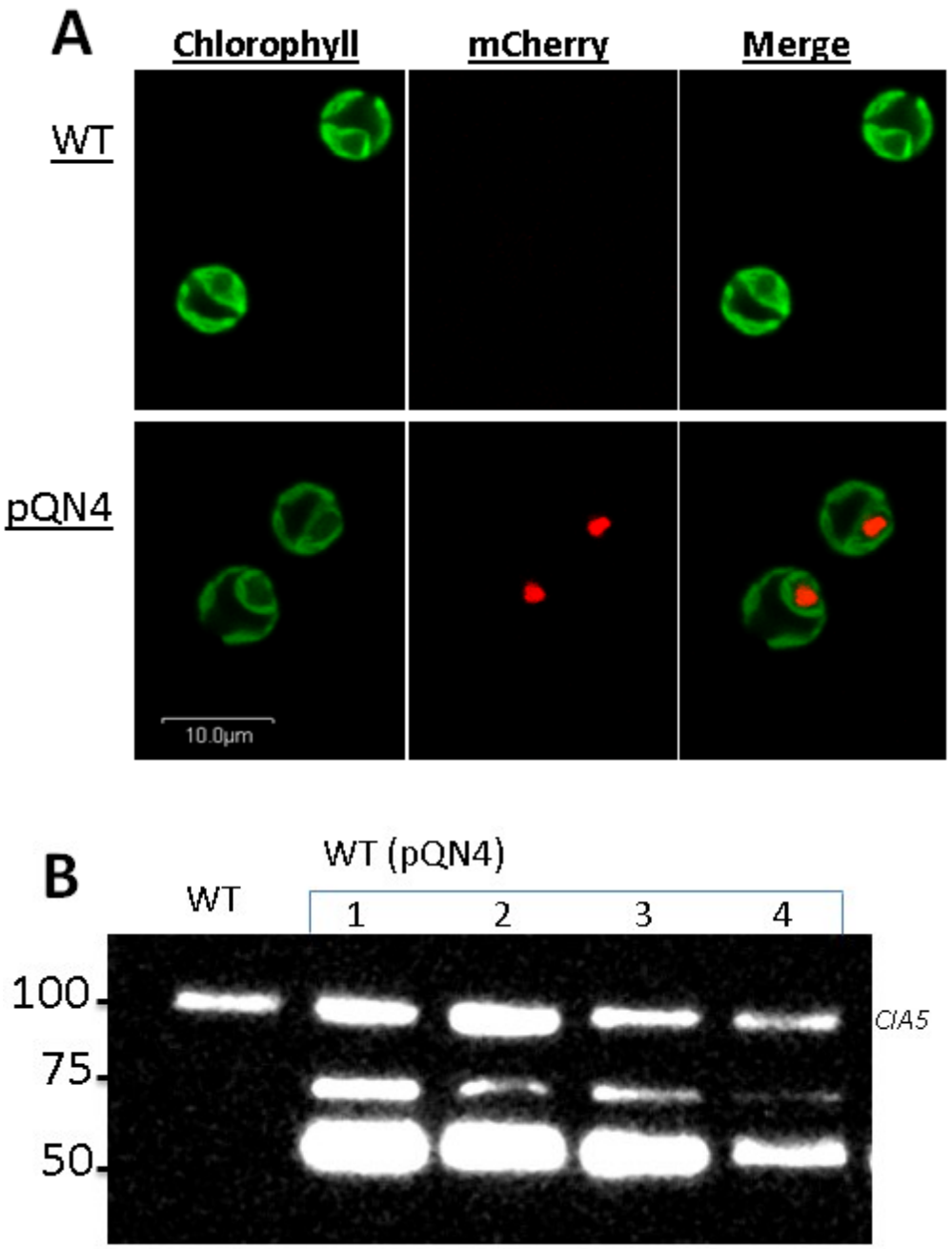



**Figure 2-7: Analysis of WT cells expressing mCherry-tagged RbcS2 expressed from plasmid pQN4**. **A)** Fluorescence imaging of WT cells transformed with the plasmid pQN4. Chlorophyll autofluorescence is depicted in green and mCherry fluorescence in red. **B)** Protein blot analysis of whole cell extracts of 4 individual pQN4 transformants in which CIA5 and RbcS2-mCherry-CIA5 Ep are detected with an antibody to the CIA5 epitope. The expected sizes of the CIA5, processed RbcS2-mCherry-CIA5 Ep and unprocessed full length RbcS2-mCherry-CIA5 Ep-2A-Ble$^{r}$ are ~100 kDa, ~55 kDa and ~70 kDa, respecitvely.

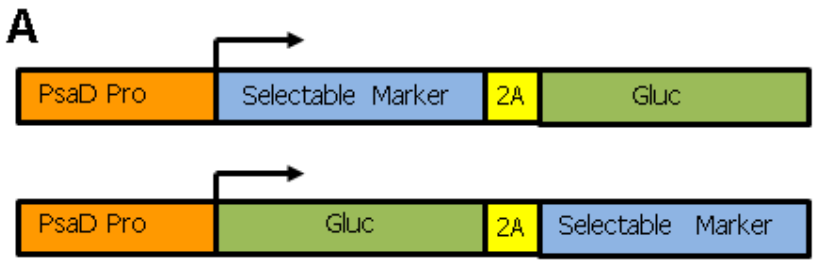
**Evaluation of the functionality and efficacy of various selection marker/2A peptide combinations**

As demonstrated above, use of the Ble$^r$ selectable marker gene in combination with the 2A system allows for dependable, high-level co-expression of various GOIs. However, there are numerous circumstances in which use of other or additional selectable marker genes in combination with the 2A-GOI system are likely to be advantageous (e.g. when high constitutive levels of expression of a GOI cannot be tolerated). Furthermore, situations may arise in which it is desirable to transform a single cell with multiple polycistronic 2A constructs – perhaps each with its own strength and type (e.g., conditional/nonconditional) of promoter. For these reasons, a number of selectable markers that are commonly used in *Chlamydomonas* research were evaluated for compatibility with the extended FMDV 2A peptide and for expression level under the selection conditions employed. In addition, because there are benefits and drawbacks to placing a GOI at either the C or N terminal end of a 2A peptide, both of these orientations were tested for many of the selectable marker genes tested.

To evaluate these selectable marker/2A combinations in a facile manner, the gene encoding the Gaussia luciferase (Gluc) protein (124,125) was used as the fusion partner in all cases. When this protein is expressed it metabolizes the substrate, coelenterazine, to generate visible light that can be easily quantified using a luminometer. Gluc-2A-selectable marker and selectable marker-2A-Gluc expression vectors were constructed (Fig. 2-8A) using the following selectable marker genes: Ble$^r$ (zeocin resistance), AphVIII (paromomycin resistance) (33),

AphVII (hygromycin resistance) (34), *AadA* (spectinomycin resistance) (35), *ARG7* (arginino succinate lyase, for the restoration of arginine auxotrophy to *arg7* mutants) (36) and *RBCS2* (Rubisco small subunit 2, for the restoration of photoautotrophic growth to *rbcs* mutants) (11). As results summarized in Figure 2-8B demonstrate, expression levels and compatibilities of these selectable marker/2A combinations vary considerably. The $Ble^r$ protein functions well and promotes robust Gluc expression in either orientation (Gluc-2A-$Ble^r$ or $Ble^r$-2A-Gluc). In contrast, the Aph8, Aph7 and AadA markers (all acetyl-transferase enzymes) all behave similarly in that functional dicistronic gene constructs are not obtained with these markers in position 1. When they occupy position 2, Gluc expression levels are nearly undetectable. When the *ARG7* cDNA occupies position 2, there is approximately 10 fold greater Gluc expression as compared to the expression levels obtained with the AphVII, AphVIII and *AadA* selectable markers. Dicistronic genes consisting of *ARG7* in position 1 were not evaluated. Although *RBCS2* is not a commonly used selection marker, it was chosen for evaluation because it is one of the most highly expressed proteins in photosynthetic organisms and *rbcs* mutants exist that can be complemented with *RBCS2* genomic DNA (11). A dicistronic gene consisting of Gluc in position 1 and the *RBCS2* gene in position 2 however did not complement the *rbcs* mutant, T60-3 (11). Interestingly, when the Gluc marker was excised from this plasmid, leaving only an FMDV 2A-RBCS2 fusion, the resulting gene was able to complement T60-3 cells as efficiently as the complete WT *RBCS2* gene (Fig. 2-8B). Although fusion of GLUC to 2A-RBCS2 prevented proper RBCS2 funcionality in this particular transgenic context, future

experiments will be required to determine if it is possible to express other GOIs using the 2A-*RBCS2* vector. In summary, 6 different selectable markers available for use in *Chlamydomonas* were assessed for compatibility with a 2A peptide. Among the 4 commonly used antibiotic resistance markers tested (Ble, AphVII, AphVIII, *AadA*), all were viable when placed downstream of the FMDV 2A peptide but only the $Ble^r$ protein performed well when placed upstream of a 2A peptide (Fig. 2-8). Interestingly, the AphVII, AphVIII and *AadA* proteins appear not to tolerate fusions to their C-termini (Fig. 2-8). The active sites of the homologous AphVII and AphVIII proteins overlap with their C-terminal domains (33,34) which likely renders this region sensitive to modification.




**B**

| Construct Tested | Negative Effect on CFU/µg DNA | GOI Linkage Maintained | Approximate Luciferase RLUs |
|---|---|---|---|
| Ble$^r$-2A-Gluc | Medium | Yes | 10,000,000s |
| Gluc-2A-Ble$^r$ | Medium | Yes | 10,000,000s |
| Aph8-2A-Gluc | Extreme | No | --- |
| Gluc-2A-Aph8 | Slight | Yes | 1,000s |
| AadA-2A-Gluc | Extreme | No | --- |
| Gluc-2A-AadA | Slight | Yes | 1,000s |
| Aph7-2A-Gluc | Extreme | No | --- |
| Gluc-2A-Aph7 | Slight | Yes | 1,000s |
| Gluc-2A-Arg7 | Medium | Yes | 100,000s |
| Gluc-2A-RbcS2 | Extreme | No | --- |
| 2A-RbcS2 | None | --- | --- |

**Figure 2-8: Evaluation of various selectable markers for compatibility with the extended FMDV 2A peptide. A)** Constructs for testing effects of opposite orientations of a GOI and selectable markers in 2A-containing dicistronic genes. Gluc, Gaussia Luciferase; 2A, FMDV 2A peptide; PsaD Pro, promoter from *Chlamydomonas* PsaD gene. **B)** Performance of each construct in regard to negative effect on transformation rates [relative to the standard pSP124 plasmid (8) conferring Ble$^r$], maintenance or loss of linkage between the GOI (Gluc) and the selectable marker gene, and the approximate relative luminescence units (RLUs) measured in cultures of at least 10 randomly chosen transformants. Linkage was defined as detection of Gluc signal in at least 70% of transformants tested. Ble$^r$ (Zeocin$^r$); Aph8 (Paromomycin$^r$); Aph7 (Hygromycin$^r$); AadA (Spectinomycin$^r$); Arg7 (Arginnino Succinate lyase; complementation of arginine auxotrophy in arg7

mutants); RbcS2 (Rubisco Small Subunit 2, complementation of photoautotrophy in rbcs2 mutants).

When designing a 2A expression cassette, it is important to consider the benefits and drawbacks of placing a GOI either on the 5' or 3' side of the 2A peptide coding region. In the case of *Chlamydomonas* that has a propensity for cleaving and rearranging newly introduced DNA molecules prior to chromosomal integration, placement of the GOI upstream of the 2A and selectable maker gene sequences has the advantage of helping to insure that transformants will express the GOI. On the negative side, the protein produced from a GOI placed upstream of a 2A peptide coding sequence will always possess the residual 2A peptide (minus the 2A C-terminal proline) at its C-terminus. Conversely, proteins encoded by GOIs placed downstream of a 2A peptide coding region will always retain the 2A-derived proline at their N-termini (69). The degree to which a protein will tolerate these relatively minor alterations ultimately must be determined empirically. However, it appears that instances in which the C-terminal 2A polypeptide or the N-terminal 2A proline interfere with POI function are limited (72).

**Utilization of the Heat Shock 70A gene promoter allows robust inducible expression of a 2A-dicistron when coupled with a low expression marker gene**

Numerous circumstances exist in which constitutive expression of a particular protein can be inhibitory or toxic. For example, certain proteins such as custom nucleases may cause toxicity (126) if constitutively expressed but might be tolerated if their expression can be persistently suppressed to low levels and only transiently induced. Because several commonly available selectable markers used

for transformation of *Chlamydomonas* (Fig. 2-8) rely on enzymatic inactivation of antibiotics and, therefore, are effective even when expressed at very low levels, we reasoned that coupling such selectable marker genes to a weak constitutive, but strongly inducible, promoter could allow us to initially select for stable transformation under non-inducing conditions where expression of many potentially toxic GOIs would be very low. We further reasoned that utilizing such promoters to drive expression of 2A polycistrons containing low expression requiring marker genes would allow us achieve stable transformation under non-induction conditions and further allow for transient, high level expression of coupled GOIs. With this strategy in mind, we designed and evaluated an inducible *Chlamydomonas* 2A-containing vector (Fig. 2-9A, pPY1) that coupled expression of a potent selectable marker gene (*AphVIII*, whose product is needed in small amounts) to the Gluc gene (whose gene product, luciferase, can be readily measured) using the heat inducible *Chlamydomonas* heat shock protein 70A (*HSP70A)* gene promoter (127) to control transcription.

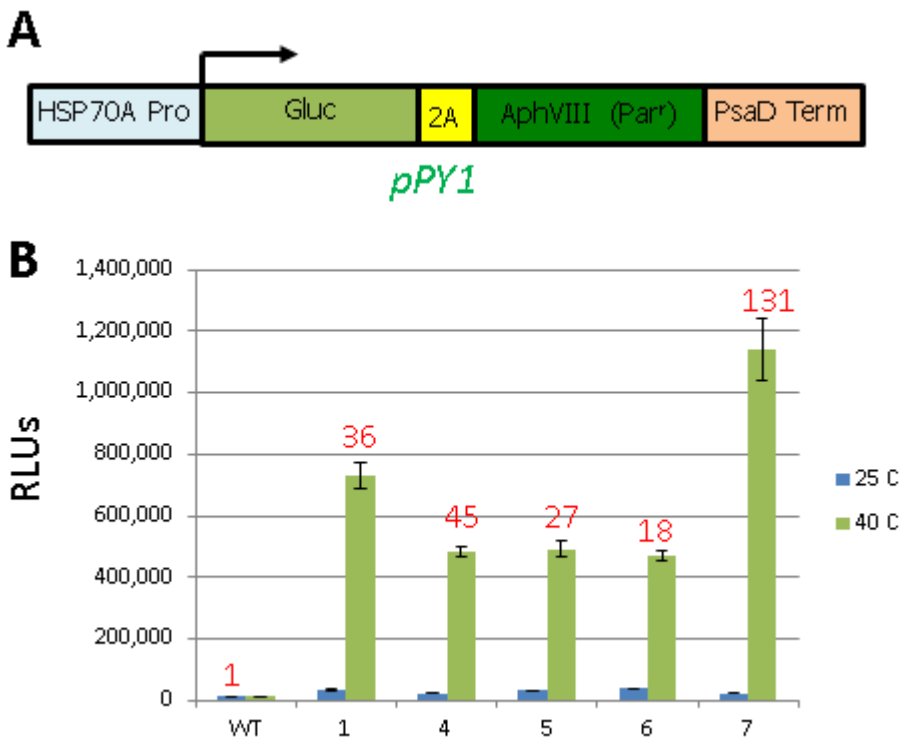



**Figure 2-9: Transcriptional control of a 2A dicistron with the Heat Shock Protein 70A gene promoter. A)** Diagram of the expression cassette used to assay heat shock inducibility. The FMDV 2A peptide was used in this construct. **B)** Heat responsive expression of the Gluc protein. Gluc activity was measured from samples prior to the initiation and 15 minutes following the completion of a heat shock. Sample densities were normalized via OD750 prior to Gluc activity measurements. The numbers above the value bars represent heat dependent Gluc signal induction factors. Raw RLU measurements are provided here.

To evaluate the ability of the *HSP70A* promoter in the pPY1 plasmid to function first as a low-level constitutive promoter for the paromomycin resistance gene during selection of transformed cells and subsequently as a strong, inducible promoter to drive high-level expression of the Gluc gene, we subjected WT cells to electroporation in the presence of pPY1 and selected transformants at 23°C on solid paromomycin-containing TAP medium. From hundreds of transformants, 12

colonies were randomly selected and grown at 23°C to mid-log phase in liquid TAP medium containing paromomycin. After shifting the cultures to 40°C for 45 min, 9 of 12 cultures screened expressed the GLUC protein in a heat responsive manner. The degree to which exposure to elevated temperature increased GLUC expression was examined more thoroughly with four of these cultures. This experiment demonstrated induction levels ranging from 18 to 131 fold above noninduced levels (Fig. 2-9B). As expected, nontransformed WT cells produced only background levels of luminescence under both heat shock and nonheat shock conditions.

These experiments provide proof of concept that a *Chlamydomonas* inducible promoter can be successfully used to initially allow selection of transformed cells with low-level expression of a selectable marker gene and subsequently gain high-level expression of a GOI (coupled to the selectable marker gene through a 2A linkage) from the same promoter when cells are subjected to conditions that force strong promoter activity. Although we chose the *HSP70A* promoter for these initial experiments, a number of other *Chlamydomonas* conditional promoters are available, including those from the *CYC6* (128), *NIT1* (129), and *CAH1* (130) genes. While the *NIT1* promoter, for instance, can be induced more strongly than the *HSP70A* promoter (127,129), the *HSP70A* promoter was used in our proof of concept experiments due to the ease and rapidity with which its activity can be controlled. And finally, because the utility of using low expression marker proteins to promote low level expression of POIs with the 2A system is limited, employing inducible promoters such as these to control

expression of such polycistrons significantly increases their usefulness and augments the overall utility and versatility of the 2A expression system.

**Development of an inducible tricistronic expression cassette for use in *Chlamydomonas***

One of the benefits of viral 2A peptides is their ability to allow the expression of complex polycistronic genes in eukaryotic cells. To date, only dicistronic genes have been produced and tested for use in *Chlamydomonas*, with one gene necessarily being a selectable marker. However, cellular operations and metabolic pathways often involve complicated, multicomponent processes controlled by multiple genes. Therefore, to fully exploit the potential of 2A peptide technology to genetically engineer *Chlamydomonas* cells for academic and commercial purposes, it would be of great potential value to have available 2A-containing polycistronic gene constructs consisting of three or more GOIs. To test the feasibility of such a goal, the tricistronic plasmid pQV7 (Fig. 2-10A was constructed and evaluated. For demonstration purposes, the *Hsp70A* gene promoter was again selected as an inducible promoter to drive both low and high-level expression of three downstream genes: *Chlamydomonas* aryl sulfatase II gene (Cr*ARSII*), Gluc and the bacterial *AadA* spectinomycin resistance gene. The Cr*ARSII* and *Gluc* genes were chosen because their enzymatic activities are easily detectable and can be accurately measured. Production of the CrARSII protein is induced by sulfate starvation and is responsible for scavenging sulfur from sulfate-containing aromatic compounds. CrARSII activity can be easily measured in sensitive colorimetric assays using 5-bromo-4-chloro-3-indolyl-β-$SO_4$ (X-$SO_4$) or alpha naphthyl sulfate as substrates

(129). Importantly, CrARSII activity is normally undetectable in cultures grown under sulfur replete conditions. This means that increases in CrARSII activity caused by activation of a promoter driving *CrARSII* transgene expression can easily be detected and measured in *Chlamydomonas* cultures even when sulfate is plentiful in the external medium (131). Because expression of most proteins is more robust from natural, intron-containing coding regions compared to cDNA sequences (123), the *CrARSII* transcript (lacking UTRs)was used in construction of the pQV7 plasmid. The *AadA* gene (Fig. 2-8B) that needs only low-level expression to provide spectinomycin resistance to transgenic *Chlamydomonas* cells (Fig. 2-8) was used as the selectable marker in the pQV7 construct.

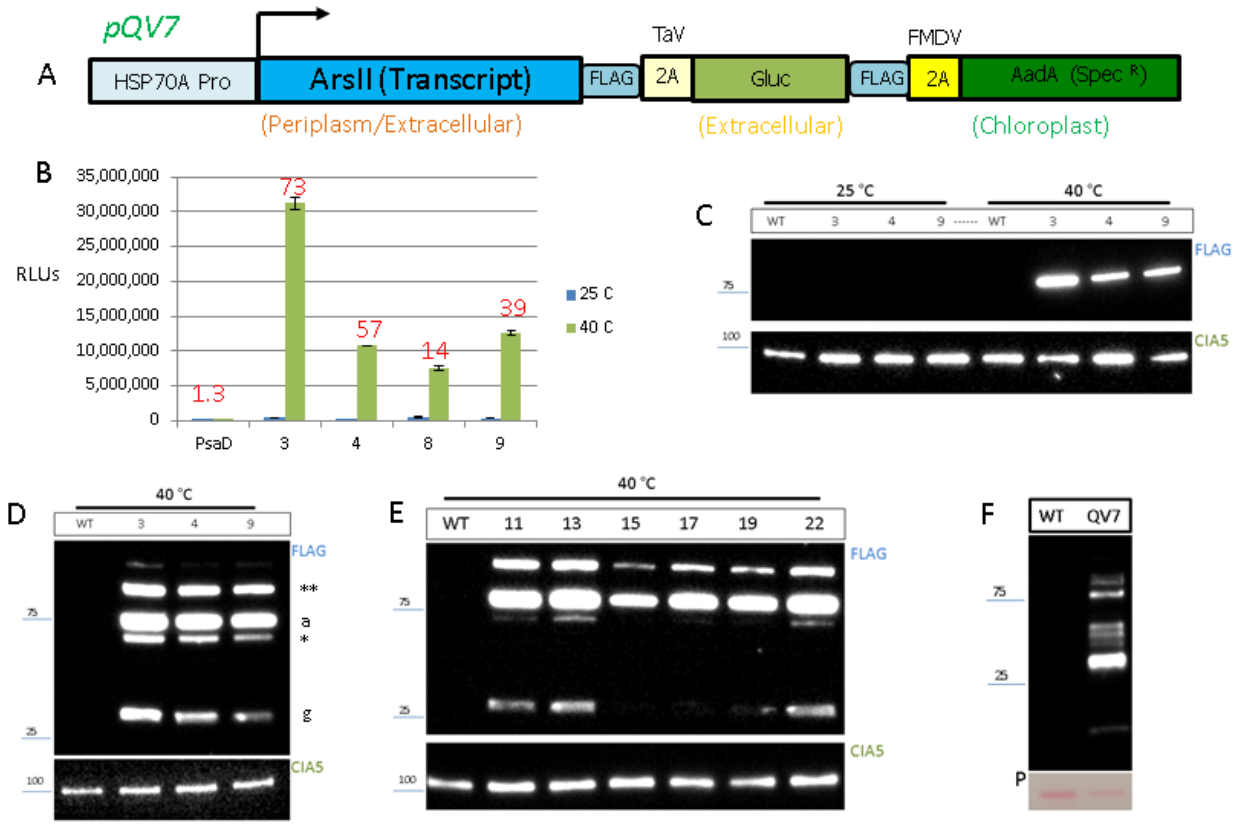



**Figure 2-10: Functional evaluation of the pQV7 tricistronic expression cassette. A)** The pQV7 cassette composed of the *HSP70A* gene promoter, the CrArsII coding region fused with a FLAG-Tag coding region, a TaV 2A coding region, the Gaussia Luciferase (Gluc) cDNA fused with a FLAG-Tag coding region, a FMDV 2A coding region and the spectinomycin resistance gene (AadA). The first 2A is the TaV 2A while the second is the FMDV 2A. 3xFLAG tags were included to allow for the analysis of 2A peptide processing. **B)** Heat inducible expression of Gluc. Gluc activity (in relative luminescence units (RLUs)) was measured from four independent pQV7 transformants (numbers 3, 4, 8 and 9) prior to (blue bars) and 15 minutes following the completion of a 45 min 42°C heat shock treatment (green bars). Gluc activity levels were also measured using cells transformed with an analogous tricistronic construct (PsaD) in which the *PsaD* gene promoter was inserted in place of the *HSP70A* gene promoter. Cell concentrations were normalized prior to Gluc activity measurements. Red numbers located above the value bars represent the ratio of Gluc activity after heat shock to Gluc activity before heat treatment. **C)** Protein blot analysis of media-free, whole cell extracts of the pQV7 transformants (3, 4 and 9) used in the experiment depicted in **B** using antibodies to the FLAG-tag peptide (upper panel) and antibodies to the CIA5 protein (lower panel). Expected sizes of proteins under study: ArsII-3xFLAG-TaV 2A protein, ~78 kDa; Gluc-3xFLAG-FMDV 2A protein, ~28 kDa; CIA5, ~100 kDa. **D)** Protein blot analysis of extracts of media-free, pelleted WT cells and QV7 transformants depicted in **C** using more sensitive detection conditions. (a) ArsII-

3xFLAG-TaV 2A protein; (g) FLAG-tagged Gluc; (* and **) intermediates resulting from incomplete 2A processing. **E)** Protein blot analysis of post heat shock extracts from 6 additional pV7 transformants that induced Gluc expression in a heat responsive manner **F)** Protein blot analysis of FLAG-tagged proteins present in cell-free media from cultures of QV7 transformant #3 and WT cells. After heat shock treatment, cell-free media from WT and QV7-3 cultures were collected and concentrated ~150 fold prior to separation by SDS-PAGE. P = Ponceau staining, post transfer, for assessing relative protein load sizes.

To analyze production and processing of proteins from the *CrARSII*-Gluc-*AadA* polycistronic gene in pQV7, twenty-two spectinomycin resistant transformants were randomly picked and transferred to liquid TAP media containing spectinomycin. When cultures had grown to mid-log phase, Gluc activity measurements were conducted with culture aliquots both prior to and after heat shock treatment at 40°C. Preliminary analyses indicated that Gluc expression was temperature responsive to varying degrees in 16 of the 22 transformants. Closer examination of four of cultures displaying temperature sensitive expression of Gluc revealed marked induction of Gluc activity by heat shock and fold increases ranging from 14 to 73 in heat shocked compared to nonheat shocked cultures (Fig. 2-10B). To verify that the 2A peptides were being processed properly during translation, both the ARSII and the Gluc proteins were tagged with a 3xFLAG epitope (Fig. 2-10A) to allow their detection on protein blots. Whole cell lysates of three pQV7 transformants displaying heat responsive Gluc induction were prepared from culture aliquots taken both prior to and after heat treatment for 45 minutes at 40°C. Protein blot analysis using a FLAG antibody detected no FLAG tagged proteins in the samples not subjected to heat treatment whereas a protein with the predicted size of tagged ARSII (~78 kDa) was clearly visible in lysates of heat shocked cells (Fig.

2-10C). A subsequent protein blot analysis conducted with the same samples from heat shocked cells but using more sensitive detection conditions also revealed the presence of FLAG tagged proteins corresponding in size to the Gluc protein (~28 kDa) and intermediate proteins resulting from incomplete 2A processing (Fig. 2-10D). Using the same protein blot conditions to analyze 6 additional pQV7 transformants that expressed the GLUC protein in a temperature sensitive manner revealed that in every sample, the FLAG tagged ARSII protein could also be detected (Fig. 2-10E). Therefore, 9 out of 9 samples examined that displayed temperature responsive GLUC expression appeared to possess an intact and functional pQV7 transgene.

The results of these experiments demonstrate successful production and processing of three independent proteins from a single tricistronic gene in *Chlamydomonas*. The efficiency of protein cleavage at the two 2A sites (FMDV 2A and TaV 2A) separating the three proteins was high as judged by the large abundance of full-sized mature proteins relative to much more limited quantities of semiprocessed Flag-tagged translation products (Fig 2-10D). However, this success was tempered by the observation that the Cr*ARSII* and Gluc gene products, that are usually secreted in large part into the growth medium (124,131) by wall-less *Chlamydomonas* cells, were found primarily associated with the cell bodies (Fig. 2-10C and 2-10D) and in only limited quantities in the external medium (Fig. 2-10F). These observations suggest that proper localization of the CrARSII and Gluc proteins may have been prevented by the presence of the residual 2A peptides or bulky FLAG tags at their C-termini. Because previous efforts have

revealed that 2A based vectors are capable of allowing excretion of proteins to the growth medium (111), it cannot be concluded from the present study that expression of proteins from 2A based polycistronic genes precludes efficient secretion from the cell. With respect to the functionality of the Gluc and ARSII proteins expressed from the pQV7 gene, although the Gluc protein was functional and easily detectable via luciferase assay, ARSII protein activity could not be detected convincingly with colorimetric assays. The residual 2A peptide as well as the FLAG tag may have prevented proper function of this particular protein. Therefore, the selection of proteins of interest, protein combinations, protein placement in a polycistron, selection of epitope tags and experimental conditions must be optimized to allow for ideal performance of 2A polycistrons.

## Conclusions

The present study provides a clear demonstration of the power of 2A expression systems in allowing efficient generation of transgenic lines as well as reliable and controllable expression of recombinant protein from the nuclear genome of *Chlamydomonas*. The versatility afforded by the large array of selectable marker configurations and the expression control strategies described here provide a user-friendly molecular tool box that makes durable, multigene engineering in *Chlamydomonas* feasible. Utilization of this tool box has the potential to expedite a broad range of algal genetic engineering and basic research endeavors.

# Materials and Methods

## Chemicals and reagents

All standard chemicals and reagents were purchased from SigmaAldrich (http://www.sigmaaldrich.com). Restriction enzymes were purchased from New England Biolabs (https://www.neb.com).

## Plasmid Constructions

The pSP124 vector (115), consisting of the $Ble^r$ gene containing two copies of the RbcS2 1st intron under the control of the *RBCS2* Promoter, was used in the assembly of all RbcS promoter-based constructs and was obtained from the laboratory of Dr. Saul Purton. To allow for construction of $Ble^r$-GOI fusions, an in-frame NdeI site (5` -CATATG GTCCTGCTCCTCGGCCAC) as well as an EcoRI site (5 ` - GAATTC CCGACGTCGACCCACTC) were added directly after the last codon of the $Ble^r$ gene. A plasmid containing an mCherry cDNA was obtained from Dr. Edgar Cahoon, University of Nebraska-Lincoln, Department of Biochemistry. The mCherry cDNA was modified via PCR to add an ApaI restriction site to the 5` terminus (5` - AAC CCG GGC CCC ATG GTG AGC AAG G) and EcoRI and NheI sites to the 3` terminus (5` -GAT GAA TTC TTA GTA GGT ACC GTT GCT AGC CTT GTA CAG CTC GTC CAT G). TaV and FMDV 2A cDNAs codon optimized for expression in *Chlamydomonas* were synthesized via megaprimer PCR and included 5` NheI, NdeI and 3` ApaI restriction sites that were used for ligation into

the NdeI and ApaI sites at the 3' end of the Ble$^r$ gene and the 5' end of the mCherry cDNA. To create the pGenD-based Ble$^r$-2A expression vectors containing PsaD promoter and termination regions, the original pGenD plasmid (123) was digested with ApaI, blunt ended with T4 DNA Polymerase and religated to ensure that the ApaI at the end of the 2A genes would be unique. The Ble$^r$ gene (containing the second RbcS2 intron, but not the first) was amplified by PCR from pSP124 using oligonucleotide primers that added in-frame 5` NdeI and 3` NheI restriction sites. This PCR product, along with the FMDV 2A cDNA and the various fusion genes were then ligated into the PsaD vector using the unique NdeI, NheI, ApaI and EcoRI restriction sites (Fig. 2-2). Plasmid pME2 was constructed by placing the CrFKB12 cDNA into the NdeI and EcoRI sites in the plasmid pJR38 (132) that contains a PsaDP cassette for expressing cDNAs, adjacent to an AphVIII expression cassette. The plasmid pQN4 was constructed by ligating together, in-frame, a RT-PCR amplified RbcS2 mRNA coding region, the coding regions for a 4xEAAAR flexible peptide linker (133), mCherry, CIA5 Ep, FMDV 2A and the Ble$^r$ gene (containing a single intron) and inserting this multigene chimera between the PsaD promoter and terminator regions of the pGenD vector (Fig. 2-2). The plasmids constructed for experiments depicted in Figure 2-8 were all based on the plasmid, pLZ1 (Fig. 2-2). The various marker genes and Gluc gene were amplified via PCR with oligonucleotide primers inserting either 5` NdeI/3'NheI or 5` ApaI/3' EcoRI restriction sites to construction of the various gene combinations examined. To construct plasmid pPY1 (Fig. 2-9) the following forward PCR primer, 5` TCTAGA AG GCT TGA CAT GAT TGG TGC, and reverse primer, 5` CAT ATG TAC TGA

CTC TTA AGC GAG TTG AGT GG, were used to amplify the *Chlamydomonas HSP70A* gene promoter including the full 5` UTR from WT genomic DNA. The NdeI and XbaI sites were used to replace the PsaD promoter in the Gluc-2A-AphVIII plasmid constructed for experiments depicted in Figure 2-8, putting the *HSP70A* promoter start codon in frame with the Gluc cDNA. The vector pQV7, used in experiments depicted in Figure 2-10, was constructed by amplifying the WT CrARSII coding region from genomic DNA via PCR and ligating it in frame with the 3xFLAG tags, 2A peptides, Gluc and the AadA coding region (35). The two 3xFLAG peptides were synthesized via megaprimer PCR to maximize codon diversity and minimize repetitive DNA sequences while maintaining optimized *Chlamydomonas* codon usage.

## ***Chlamydomonas* Growth Conditions and Transformation Protocol**

The *Chlamydomonas* wall-less strain, CC3491, was used for all experiments. Standard electroporation transformation conditions (30) were employed. Briefly, cells were grown in TAP medium (134) under continuous illumination of ~80 µE Particle bombardment is primarily used to transform the *C. reinhardtii* chloroplast genome, where foreign DNA is primarily incorporated via homologous recombination (25). Particle bombardment is primarily used to transform the *C. reinhardtii* chloroplast genome, where foreign DNA is primarily incorporated via homologous recombination (25). /$m^2$/sec and rotation at 140 RPM to a density of ~$3x10^6$ cell/mL. Cells were pelleted by centrifuged and brought to a concentration of

$4x10^8$ cells/mL in TAP + 60 mM sucrose. For each individual transformation, 1 ug of circular plasmid DNA was used to transform $1x10^8$ cells. Cells were allowed to recover in TAP + 60 mM sucrose medium with shaking at 140 RPM at room temperature under constant illumination for 20-24 hours before plating onto solidified TAP medium containing appropriate levels of antibiotics. Plates were maintained at 25°C under constant illumination. Antibiotic resistance genes and antibiotic concentration used were: AphVIII (20 µg /mL Paromomycin), $Ble^r$ (10 µg/mL Zeocin), AadA (150 µg /mL Spectinomycin), Aph7 (10 µg /mL Hygromycin). The arginine-requiring mutant, strain CC1618 , was transformed with vectors containing a WT *ARG7* gene followed by selection for complementation of arginine auxotrophy on plain TAP plates, under constant illumination at room temperature. *Chlamydomonas* mutant strain, T60-3 (11) obtained from Robert Spreitzer, lacking functional *RBCS1* and *RBCS2* genes, was transformed with *RBCS2* constructs followed by selection for photoautotrophic growth on TP plates lacking a carbon source at room temperature, under constant illumination in 5% $CO_2$.

## ***FKB12* complementation assay**

A rapamycin resistant strain of CC3491 was generated and used for complementation experiments using a WT *Chlamydomonas FKB12* gene. Following transformation and selection on the appropriate antibiotic, colonies 7-10 days old were picked into 100 uL of TAP media in 96 well plates and allowed to acclimate for 12-24 hours in the light at room temperature. Samples were thoroughly mixed via

pipetting and then replica plated in 5 µl volumes onto plates containing solidified TAP or TAP + 10 µg/mL rapamycin. CC3491 and the rapamycin resistant CC3491 mutant were plated as controls. The plates were wrapped in Saran wrap and kept in continuous light at room temperature. Colonies on plates were photographed 3-5 days after spotting.

**Protein Blotting**

Samples of *Chlamydomonas* colonies to be analyzed were picked into 2 mL of TAP media in test tubes and shaken at 285 RPM, 25 C° under continuous illumination until the density reached ~3-6 x $10^6$ cells/mL. An appropriate volume of cells were centrifuged and resuspended in 1/$10^{th}$ volume of water. The resuspended cells were then diluted with 1 volume of 2xSDS-PAGE buffer (1 M Tric-Cl pH 6.8, 4% SDS, 2% glycerol, 20 mg bromophenol blue, 14 µl/ml 2-mercaptoethanol) and boiled for 5 minutes. 10µL of each sample was loaded into wells of a Tris-Glycine SDS-PAGE gel (www.bio-rad.com/) and fractionated by electrophoresis. For western blotting, the membrane was blocked with TBST (Tris Buffered Saline containing 0.05% Tween 20) in 5% not-fat dry milk. Primary CIA5 antibody (Ab153) was added at a 1:8,000 dilution into TBST in 5% non-fat dry milk. HRP conjugated donkey anti-rabbit secondary antibody (GE Healthcare – ECL, http://www.gelifesciences.com) was added at a dilution of 1:2,500. The membrane was then treated with SuperSignal West Pico Chemilumiscent Substrate (Thermo Scientific, *www.thermoscientific.com*) followed by image capture using the Odyssey FC Imaging System from LI-COR Biosciences (www.licor.com).

## Confocal Fluorescence Microscopy

Images of live *Chlamydomonas* cells were captured using a Nikon A1 confocal imaging system mounted on a Nikon Eclipse 90i microscope with a 100x objective. mCherry and chloroplast fluorescence signals were acquired sequentially with a 561.5nm excitation and 570-620nm emission and 641nm excitation and 662-737nm emission, respectively, and pseudo-colored red and green for visualization.

## Heat shock treatment

Prior to heat shock treatment, samples were grown under continuous light to mid log phase (~3-6 x $10^6$ cells/mL) in TAP media and normalized using $OD_{750}$ readings. A volume of 2 mL of each culture were shaken in a test tube at 275 rpm, under constant illumination of ~80 µmole/$m^2$/sec, at either 25°C or 40°C for 45 minutes. Immediately following the completion of the heat shock treatment, cell samples were transferred to room temperature and allowed to equilibrate/recover for 15 minutes with shaking and illumination conditions (identical to those above) prior to assessing gene expression levels and the efficiency of 2A-mediated protein cleavage by protein blot analysis.

**Luciferase Assay**

Prior to luciferase assays, 7-9 day old colonies of transformed *Chlamydomonas* cells grown on TAP plates under antibiotic selection and constant illumination at 25 C°, were picked and suspended thoroughly in 99 µl of TAP liquid media. A volume of 1 µL of 1 mM Coelenterazine was added with mixing to each sample immediately prior to measurement of luminescence with a luminometer (Turner Biosystems Instruments, Model: 2030-000; www.promega.com) An integration period of 5 seconds was used for each measurement. The addition of substrate directly to nontransgenic WT *Chlamydomonas* cultures at a density of ~6 x $10^6$ cells/mL produces a background reading of 10,000 RLUs or less.

# CHAPTER 3: Enhanced methods for the generation and characterization of insertional mutant libraries in *Chlamydomonas reinhardtii* containing high percentages of genic insertions

## Abstract

Given the current lack of methods available for achieving targeted gene modification in *Chlamydomonas*, creation of large mutant libraries containing a high percentage of knockout mutants and rapid characterization of library individuals are vital to maintaining the value of this single-cell alga as a model eukaryote for studying various processes of fundamental and practical importance such as photosynthesis and biofuel production. Toward this end, we report the development of several new and efficient methods for creating, isolating, and characterizing mutants with insertionally inactivated genes. The first of these methods involves use of a small, promoter-less, selectable marker gene coding region to bias selection of transformants bearing insertions in transcribed regions of the genome. The second involves use of a cocktail of restriction enzymes to digest genomic DNA from insertional mutants to create marker gene containing DNA fragments of an optimal size range to allow efficient inverse PCR amplification of genomic DNA flanking marker insertion events. We also provide verification that when foreign DNA is delivered to cells via electroporation the site of foreign DNA insertion is generally subject only to small deletions and insertions of nucleotides typical of that occurring during nonhomologous end joining DNA repair. This contrasts markedly with previously described situations in which foreign DNAs are delivered via particle bombardment or the glass bead vortexing method. Finally, we report discovery that low concentrations of the Fermentas Fast Digest Green® buffer can be used to increase by approximately 8-fold the rate of transformation of *Chlamydomonas* with low amounts of foreign DNAs, such as those available in the presently described

methods for efficient generation and characterization of mutants containing insertionally inactivated genes.

## Introduction

*Chlamydomonas* is a unicellular, photosynthetic, eukaryotic alga with a fully sequenced and annotated genome (2,8). It is a useful model for studying processes such as photosynthesis (106), cellular motility (45) and more recently biofuel production (51). Because *Chlamydomonas* is haploid during vegetative growth, a powerful strategy that can be used to study these biological processes is to characterize mutants that have been generated by various means. Unfortunately, unlike bacteria and certain yeast where targeted gene knockout and replacement by homologous recombination can be achieved with ease, homology mediated modification of the nuclear genome is highly inefficient in *Chlamydomonas* (as with most eukaryotic organisms) (56,135,136). Because of this, various random mutagenesis methodologies such as chemical mutagenesis or exposure to UV radiation (87,137) have been relied upon in these organisms to generate mutants. However, the mutants generated by these treatments are often difficult to study because individuals with an interesting mutant phenotype may contain many other mutations throughout its genome. This makes identification of the gene(s) responsible for a particular phenotype cumbersome and time consuming. The introduction of genetic transformation methods enabled the development of another technique known as insertional mutagenesis that does not suffer from this limitation

(90,100). Insertional mutagenesis is the process by which a marker DNA fragment (often an antibiotic resistance gene) is introduced into host cells resulting in random insertion events throughout the genome, thereby generating a pool of mutants harboring genes disrupted by the transforming DNA. Because a marker DNA of known sequence is inserted into the genome with this technique, genomic regions flanking insertion sites can be determined to reveal the nature of the insertion events as well as the identities of the disrupted genes. Due to the beneficial features of this strategy, it has become the favored means of generating mutants in *Chlamydomonas* (58,100,138).

A number of methods have been developed to enable identification of genomic DNA sequences flanking loci disrupted by insertion of foreign DNA (89,105). Some examples of the more well-known and successful techniques used are Thermal Asymmetric Interlaced (TAIL) PCR, adapter-ligation mediated PCR and Inverse PCR (89,104,139). Many commonly utilized methods however are simple variations of the basic TAIL PCR procedure (89,104,105,140). With this technique, a DNA fragment consisting of a portion of the known marker and a portion of the genomic DNA immediately flanking this marker is amplified via PCR by an assortment of semi-random, degenerate primers that are used in conjunction with a marker specific primer. Often subsequent nested PCR reactions with additional random and marker specific primers are required to enhance specificity and produce an adequate supply of PCR product. The PCR product obtained is then sequenced to reveal one of the genomic DNA regions flanking a marker insertion site. Although this is a relatively simple and generally successful strategy,

there are significant limitations associated with its implementation. First, designing the random primers that are utilized for this technique requires that great care be taken to optimize performance. Second, the design, efficacy and specificity of the primers are highly empirical and often times many of them need to be tested before success is achieved – a costly and time-consuming effort (89,140,141). Third, because insertional mutagenesis can lead to gross genomic rearrangements or deletion of large genomic DNA fragments in the vicinity of marker insertion events (142), it is important to be able to identify the genomic DNA regions flanking both sides of an insertion event. This requires that two successful TAIL PCR procedures be conducted for each individual mutant. Similar limitations also plague gene identification methods based on ligation of adapter oligonucleotides to restriction enzyme cut sites adjacent to integrated marker DNA, otherwise known as adapter ligation mediated PCR (89,139). These hurdles become significant drawbacks when the goal is to analyze large pools of mutants consisting of hundreds or thousands of individuals.

Inverse PCR is an alternative, straightforward method for identifying genomic DNA sequences flanking marker insertion events. Most importantly, this method enables the amplification of DNA sequences flanking both sides of a foreign DNA insert using only a single PCR reaction (89,143). The first step in this technique is to fragment the genome of an insertional mutant with a restriction enzyme that does not cut within the inserted marker DNA. Next, this restriction fragment pool is treated with DNA ligase to generate circular DNA molecules, one of which contains the known marker DNA as well as the flanking genomic sequences. At this point,

PCR is conducted using two marker-specific primers (5` anti-sense and 3` sense) to PCR amplify DNA from the insert and the two adjacent flanking regions. The amplified DNA can subsequently be sequenced to identify the genomic loci into which the foreign DNA was inserted and the extent to which these loci were truncated or rearranged. There are, however, two significant limitations associated with this technique that significantly reduce its efficiency in many instances. First, because large DNA fragments are more difficult to amplify by PCR than smaller fragments, the sizes of the genomic DNA flanking regions resulting from fragmentation of the genome with restriction enzymes must be minimized. However, in many cases the restriction enzyme used results in the production of fragments that are difficult to amplify due to their excessive size. Second, if DNA sequences not essential for proper marker gene functionality are included in the transforming DNA, they are often lost during transformation. Because of this, the primers used for inverse PCR should bind to essential marker DNA regions. This can drastically increase the final size of inverse PCR products if whole plasmids are used to generate insertional mutants. Nonetheless, such molecules are often used as mutagens (89,105). If these limitations are successfully addressed, the efficacy of inverse PCR can be significantly increased.

When conducting insertional mutagenesis, it is important to consider that the degree to which significant phenotypic disruptions can be obtained depends on the type of genomic loci that are disrupted by marker insertion events. For many experiments that have described the generation of insertional mutants of *Chlamydomonas*, intact antibiotic resistance genes containing their own

independent promoters and terminators have been used as selectable marker DNAs (144-148). Such genes, if inserted intact, can therefore function nearly anywhere they are inserted in the genome. Bulk analysis of plant and algal insertional mutant pools generated with such markers has revealed that a significant number of the insertion events generated by transformation with full-length genes occur in inter-genic regions or 3` UTRs. Such inserts are less likely to cause significant functional disruption of gene activity than inserts into genic regions such as exons or introns. A recent study conducted in *Chlamydomonas* described a reverse genetics approach for isolating individuals that contain disruptions in a defined set of genes of interest following insertional mutagenesis. Of the mutants analyzed in these experiments, the marker gene had integrated into intergenic or 3` UTR regions in 14 out of 47 instances (138). Furthermore, in plants, one study that examined over 88,000 unique *Agrobacterium tumefaciens*-mediated mediated T-DNA insertion events determined that roughly 60% of these T-DNAs had landed in intergenic regions, 3` UTRs or putative promoter regions – and not in coding regions (90).

In the present study, optimized methods for generating insertional mutants and characterizing the resulting transformants by inverse PCR were developed. By transforming with a minimal marker gene lacking non-essential DNA sequences (i.e., promoter and terminator regions) we reasoned that this strategy was likely to yield a much higher percentage of transformants with bona fide gene knockouts and fewer transformants bearing insertions in regions of the genome of dubious interest. Additionally, careful selection of the restriction enzymes used for the initial

fragmentation of the genome prior to inverse PCR was found generate optimally sized inverse PCR templates that could be amplified with ease from most mutants. As reported below, utilization of a promoter-less marker gene to generate transformants caused a pronounced bias towards genic insertions and, not surprisingly, produced fewer transformants than its full-length counterpart. However, we report that a commercially available buffer is capable of significantly improving transformation efficiency when incorporated into the standard *Chlamydomonas* electroporation protocol (30), allowing large pools of mutants to be generated with a promoter-less marker gene in a facile manner. Together, this newly devised insertional mutagenesis system provides significantly improved success rates for generating *Chlamydomonas* mutants with insertions in transcribed regions of the genome and for identifying the genes that have been disrupted.

## Results

### Utilization of a restriction enzyme cocktail to fragment the genome prior to inverse PCR generates a favorable restriction fragment size range for inverse PCR

When deciding which restriction enzymes to use to fragment the genome of an insertional mutant prior to inverse PCR, it is important to consider how frequently one expects an enzyme to cut. The most commonly employed restriction enzymes have 6 bp recognition sequences. This is because there is an abundance of such enzymes commercially available and it is usually difficult to construct selectable marker genes that do not contain recognition sites for the relatively fewer commercially available restriction enzymes that recognize 4 bp or 5 bp sequences. Even though a random 6 bp sequence is expected to occur every 4096 bp in a

genome of 50% GC, 50% AT, genomes do not consist of random distributions of nucleotides (8). Therefore, some restriction enzymes with 6 bp recognition will cut more frequently than others in a particular genome. Consequently, prior to inverse PCR, the genome of an insertional mutant is usually digested separately with a number of different enzymes because many of them may not produce small enough fragments to enable efficient PCR amplification of the target DNA sequence (89,143). To analyze the possibility of going beyond this semi-random approach to enzyme selection, the 1$^{st}$ chromosome of the *Chlamydomonas* nuclear genome (64% GC content), representing 9.8 MB of the total 111 MB genome, was downloaded and parsed for the number of times various 6 bp restriction sites occur. It was striking to find that the number of times that restriction sites with identical GC content (NruI, FspI, BamHI, PstI; 66%) occur in the genome can differ by more than 10 fold (e.g. PstI vs BamHI) (Table 3-1A). After examining the cutting frequency of a number of enzymes, some were found that cut much more frequently than would be predicted by assuming random distribution of restriction sites throughout the genome (Table 3-1B). However, even these enzymes did not seem suitable alone as *in silico* digestions of the 1$^{st}$ chromosome revealed large numbers of fragments larger than 1 kb in size for each enzyme analyzed (Figs. 3-1 & 3-2). An inverse PCR template containing two 1 kb flanking sequences would generate a 2 kb PCR product that should be relatively easy to amplify in most cases. Larger PCR products can be successfully amplified but the risk of failure increases significantly as amplicon size increases. To alleviate this restriction fragment size issue that exists even with the some of the most frequently cutting enzymes with 6 bp

recognition sequences, it would be beneficial to simultaneously digest genomic DNA with multiple enzymes to lower the average size of DNA regions flanking the DNA insert. Accordingly, a cocktail was designed consisting of enzymes that do not cut within the selectable marker gene and that all produce blunt ends, enabling all restriction fragments to circularize upon treatment with DNA ligase regardless of which particular enzyme(s) cuts on either side of a marker DNA insert (Table 3-2). *In silico* digestion of portions of the 1$^{st}$ chromosome with this cocktail revealed that very few fragments larger than 1 kb were produced (Figs. 3-1 & 3-2). When the actual fragment size ranges for various digestion scenarios were analyzed, it was found that the odds of the individual cocktail enzymes alone producing fragments 1 kb or larger ranged from 23% to 48% (Table 3-2). In contrast, only about 2% of restriction fragments generated by the complete cocktail were larger than 1 kb.

When dealing with large genomes, it is also imperative that the flanking regions amplified by inverse PCR remain large enough to enable unambiguous identification. This consideration becomes significant with some 4 bp recognizing enzymes such as HaeIII (GGCC) where the odds of producing fragments less than 13 bp in size for a single flanking region approaches 10% (data not shown). A random 13 bp fragment is expected to occur roughly once every 67 MB, or twice in the 111 MB *Chlamydomonas* genome. Therefore, precise mapping of any fragment 13 bp or larger would be relatively simple, especially when the opposite flanking sequence is large enough to be easily mapped. However, a random 12 mer would theoretically occur once every 17 MB or roughly 6 to 7 times in the *Chlamydomonas* genome, rendering precise mapping of sequences 12 bp or smaller more

problematic. The odds of the cocktail producing fragments 12 bp or smaller is only 6% (Tabe 3-2), which means that approximately 0.36% (or 1 in 278) of insertion events will result in the production of inverse PCR templates that possess genomic DNA flanking regions on both sides of the marker smaller than 13 bp in size.

**A**

| Enzyme | Cut Site | Sites in Chr #1 | Avg. Frag Size |
|---|---|---|---|
| Random (66% GC) | nnnnnn | ~2,437 | ~4,096 |
| | | | |
| XbaI | tctaga | 353 | 28,278 |
| ClaI | atcgat | 440 | 22,687 |
| NotI | gcggccgc | 2,732 | 3,654 |
| | | | |
| NruI | tcgcga | 1,352 | 7,383 |
| FspI | tgcgca | 5,420 | 1,842 |
| | | | |
| BamHI | ggatcc | 1,062 | 9,399 |
| PstI | ctgcag | 10,698 | 933 |

**B**

| Enzyme | Cut Site | Sites in Chr #1 | Est Sites in Genome | Avg. Frag Size | Sites in Aph8 |
|---|---|---|---|---|---|
| Random (66% GC) | nnnnnn | ~2,437 | ~27,343 | ~4,096 | --- |
| FspI | tgc gca | 5,420 | 60,831 | 1,842 | 0 |
| BsaAI | yac gtr | 9,194 | 103,187 | 1,086 | 0 |
| PvuII | cag ctg | 10,295 | 115,544 | 970 | 0 |
| NaeI | gcc ggc | 13,032 | 146,263 | 766 | 0 |
| SfoI | ggc gcc | 13,267 | 148,900 | 752 | 0 |
| | | | | | |
| AluI | ag ct | 41,884 | 470,079 | 238 | 4 |
| HpyCH4V | tg ca | 65,535 | 735,522 | 152 | 8 |
| BstUI | cg cg | 71,027 | 797,160 | 141 | 9 |
| HaeIII | gg cc | 76,800 | 861,953 | 130 | 9 |

**Table 3-1: Restriction analysis of chromosome 1 of *Chlamydomonas*.** Chromosome 1 was parsed to determine restriction enzymes recognition sequence

frequency. **A)** Restriction analysis using select restriction enzymes. Average fragment sizes are calculated by dividing the total size of chromosome 1 by the number of cut sites detected for each enzyme. **B)** Comparative restriction analysis of the inverse PCR restriction enzyme cocktail enzymes and select restriction enzymes with 4 bp recognition sequences. Aph8 refers to the ΔPro AphVIII marker sequence described in the text.

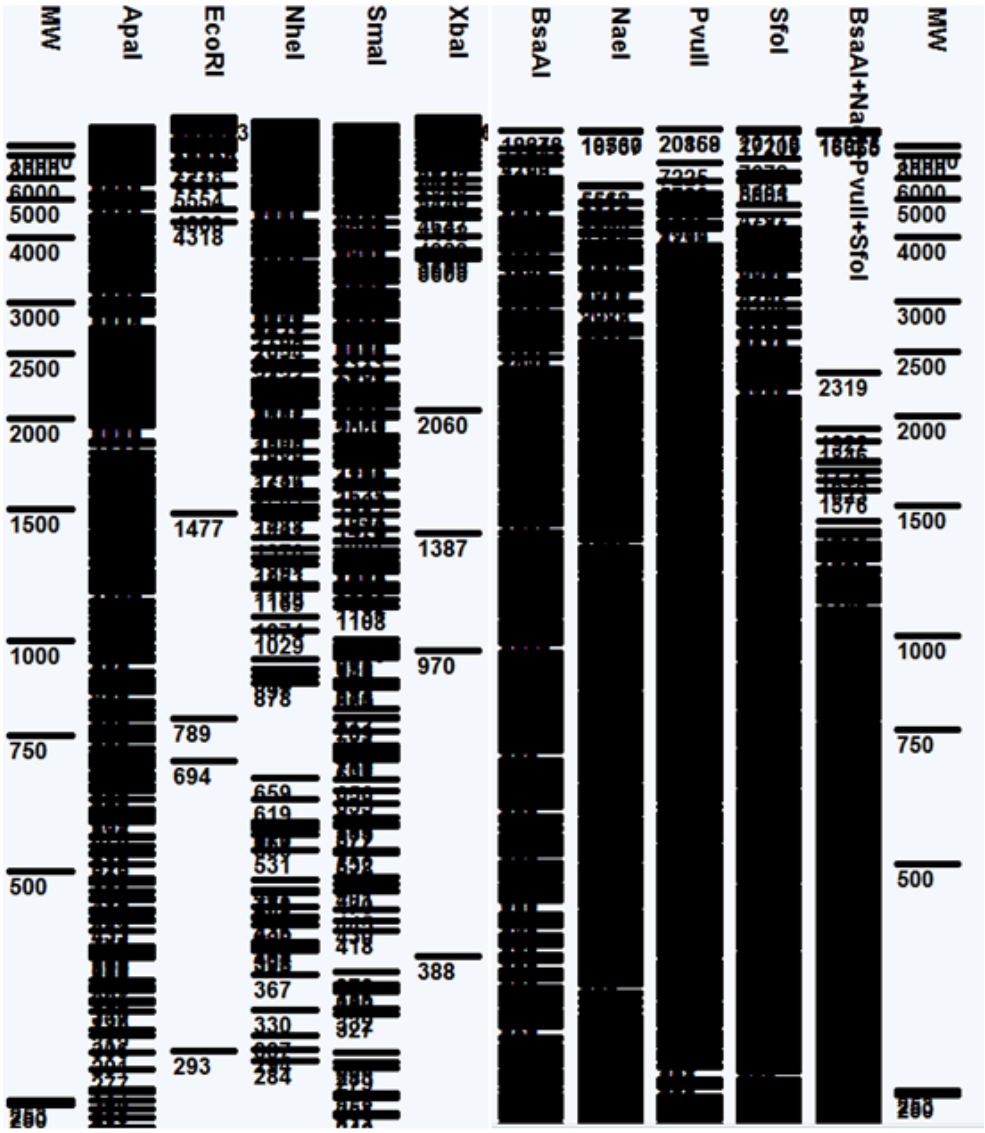



**Figure 3-1**: **Graphical representation of a virtual agarose gel electrophorogram of various restriction digests of ~700 kb of *Chlamydomonas* chromosome 1.** Graphics were generated by the pDRAW32 program. Each horizontal black band represents a unique restriction fragment. The number below each band represents its size in nucleotide base pairs. The second lane from the right represents the restriction pattern generated by digestion with BsaAI, NaeI, PvuII and SfoI simultaneously.

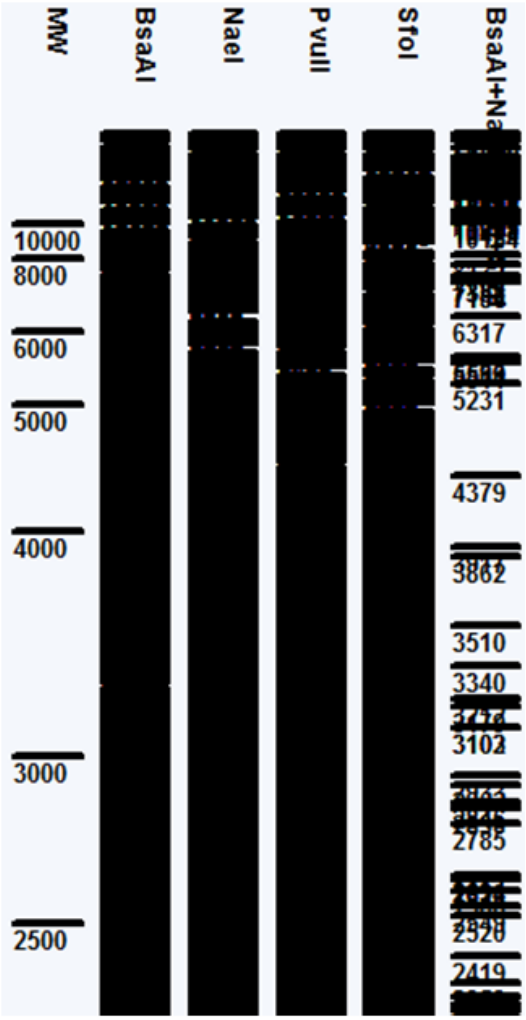



**Figure 3-2**: **Graphical representation of a virtual agarose gel electrophorogram of various restriction digests of chromosome 1 of *C. reinhardtii*.** Graphic was generated by the pDRAW32 program. Each horizontal black band represents a unique restriction fragment. The number below each band represents its size in nucleotide base pairs. Parts of some digests are almost solidly black due to the generation of many restriction fragments of overlapping size. The last lane to the right represents the restriction pattern generated by combined digestion with BsaAI, NaeI, PvuII and SfoI.

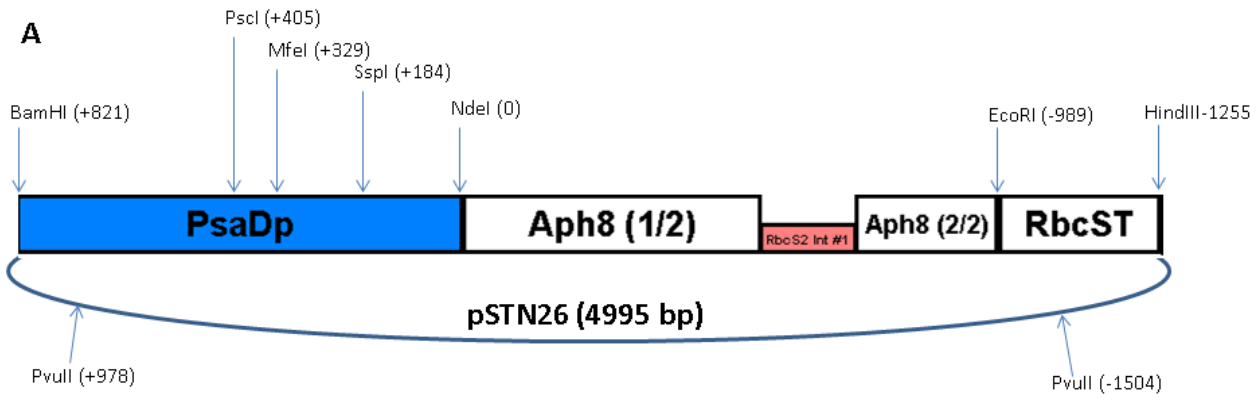
| Enzyme | <13 bp | >1 kb |
|---|---|---|
| FspI (TGC GCA) | 0.5% | 48.4% |
| NaeI (GCC GGC) | 2.3% | 24.1% |
| BsaAI (YAC GTR) | 1.1% | 35.1% |
| SfoI (GGC GCC) | 2.8% | 23.1% |
| | | |
| AluI (AG CT) | 5.8% | 1.3% |
| FspI+NaeI+BsaAI+SfoI | 6.0% | 1.9% |

**Table 3-2: Restriction fragment analysis.** Restriction fragment patterns produced *in silico* for each listed digestion scenario were analyzed to determine the odds of a random restriction fragment being either smaller than 13 bp or larger than 1,000 bp. Odds were calculated from actual restriction fragment pools generated by *in silico* digestion of ~700 kb of *Chlamydomonas* chromosome 1.

**Construction of a minimal marker gene and evaluation of an optimized inverse PCR method relying on the use of a restriction enzyme cocktail to fragment genomic DNA prior to inverse PCR**

To further reduce the size of amplified inverse PCR products and improve the efficiency of our procedure for recovering DNA sequences flanking marker gene insertion sites, we sought to trim our selectable marker gene (AphVIII) to only those sequences essential for efficient transformation. The AphVIII gene provides resistance to the antibiotic paromomycin and achieves high transformation efficiency when used as a selection marker in *Chlamydomonas* (33). A modified AphVIII gene was constructed using the PsaD gene promoter (123) and the RbcS2 gene terminator (115) to control transcription (Fig. 3-3A). In addition, to increase transformation efficiency further, an intron containing a known transcriptional enhancer element (RbcS2 1$^{st}$ intron (115)) was inserted into the AphVIII cDNA. To determine the minimal elements needed to produce high transformation efficiency, the terminator and the promoter were truncated to various lengths and the resulting linear fragments were used to generate *Chlamydomonas* transformants. When the terminator was removed, there was no observable decrease in transformation efficiency (Fig. 3-3B). The PsaD promoter on the other hand was essential for optimal transformation efficiency and could only be truncated to the MfeI site without causing a significant drop in transformation efficiency. Excision of all DNA sequences upstream of the SspI site resulted in transformation efficiency roughly equal to that of a AphVIII fragment completely lacking promoter and terminator elements (ΔPro/Ter AphVIII) (Fig. 3-3B). The ΔPro/Ter AphVIII gene (i. e., coding region) alone produced more than 10 times fewer colonies than the full-length

AphVIII gene (Fig. 3-3B). These results demonstrate that the AphVIII gene does not need a terminator or 3` UTR to achieve optimal transformation efficiency and that only a portion of the standard ~ 800 bp PsaD promoter is required for functionality.




**B**

| Fragment Transformed | Transformation Efficiency |
|---|---|
| STN26 (BamHI-HindIII) --- Full Gene | ~6000 |
| STN26 (BamHI-EcoRI) ---- ΔTer | ~6000 |
| STN26 (PscI-EcoRI) | ~4000 |
| STN26 (MfeI-EcoRI) | ~3200 |
| STN26 (SspI-HindIII) | ~440 |
| STN26 (NdeI-EcoRI) --- ΔPro/Ter | ~400 |

**Figure 3-3**: **Construction of a modified AphVIII gene for generating insertional mutant libraries. A)** Diagram of the AphVIII gene designed for this study (pSTN26). The RbcS2 1$^{st}$ intron, a known transcriptional enhancer element,was inserted into the AphVIII cDNA to increase transformation efficiency and transcriptional activity from low expression genomic loci. The PSAD promoter and 5` UTR and the RBCS2 3` UTR and terminator regions were used to control transcription. **B)** Transformation efficiency comparison of various truncated versions of the AphVIII gene depicted in **A**. Fragments were obtained by digesting 1.5 µg of pSTN26 and purifying the indicated fragments from an agarose gel following electrophoresis. The numbers in the right column represent the total number of colonies obtained from transformation with these fragments.

The ΔPro/Ter AphVIII selectable marker gene (coding region) was used to evaluate the efficacy of an optimized inverse PCR method relying on the use of this minimal marker and the aforementioned restriction enzyme cocktail (Table 3-2). First, genomic DNAs from 24 randomly chosen ΔPro/Ter AphVIII restriction fragment transformants were isolated and purified. This purified genomic DNA was then digested with the restriction enzyme cocktail to produce a collection of blunt-ended DNA fragments that were ligated to produce circular, inverse PCR templates. These samples were PCR amplified using two marker-specific oligonucleotide primers. PCR products were obtained for every sample, 12 of which are depicted in Figure 3-4. These products were then purified and sequenced. Two samples did not produce reliable sequence information, likely due to inadequate template concentration. Of the 22 samples that produced reliable sequence data, 21 could be mapped to the genome (Table 3-5, Samples 1-24). The single sample that could not be mapped to the genome (sample #5) appeared to be a PCR product derived from a contaminating vector containing no *Chlamydomonas* genomic sequence (Not Shown, sample 5). Unexpectedly, for two of the remaining samples that mapped to the genome, each of the genomic DNA regions flanking the marker aligned with high confidence to genomic loci either on different chromosomes or very distant loci on the same chromosome (Not shown, samples 3 and 15). These phenomena could possibly have been the result of genomic DNA rearrangement during transformation, which has been observed previously (142). The maximum observed size for any inverse PCR in this experiment was approximately 1.5 kb, with each product containing at most 230 bp of marker specific sequence.

Therefore, utilization of a multicomponent restriction enzyme cocktail as well as a minimal selectable marker gene to generate transformants was successful in producing small and manageable inverse PCR templates that could easily be amplified by PCR in at least 22 out of 24 samples (92%) analyzed. Moreover, 21 of the 22 PCR products could be unambiguously mapped to the genome. Importantly, for nearly every sample, only one prominent PCR product was amplified, whereas with other methods, significant background DNA fragments are a common occurrence (103,104,140).

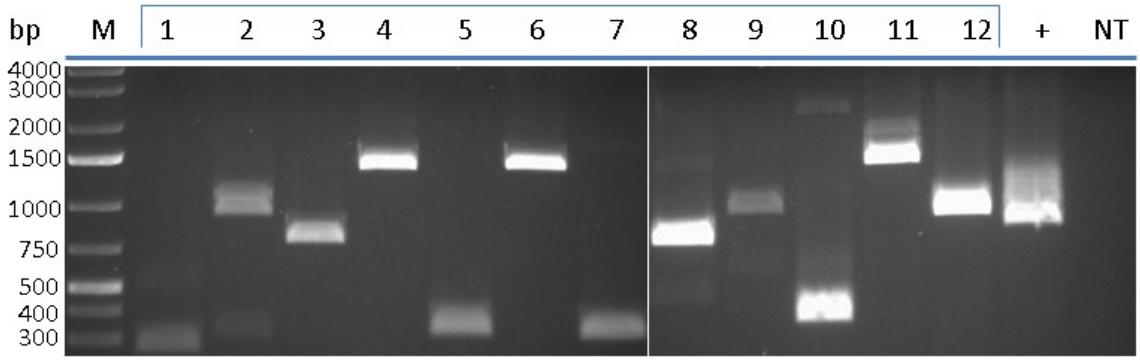



**Figure 3-4**: **Trial inverse PCR amplifications of DNA fragments containing insert DNA using optimized inverse PCR conditions**. Genomic DNAs purified from 12 randomly chosen paromomycin resistant transformants were digested with a cocktail of restriction enzymes producing blunt ends, ligated to form circular DNA molecules and PCR amplified with a set of oligonucleotide primers specific to the Δpro/Ter AphVIII insert DNA. Bp = base pairs; M = DNA markers. + represents a positive control sample containing gDNA that previously produced a verified inverse PCR product. NT = No Template control.

**Analysis of insertion events resulting from transformation with truncated marker genes of various lengths and compositions**

As previously mentioned, when a full-sized selectable marker gene is used to generate insertional mutants, a large percentage of insertion events occur in intergenic or 3` UTR regions (90,138). Even when in one study (103) the gene terminator region was removed from the marker gene, a high portion of inserts (11 out of 38) were still found in these regions. In addition, in an earlier study with *Chlamydomonas* it was shown that most insertion events produced by transformation with a PCR-generated minimal marker gene do not produce significant genomic deletions or insertions in the vicinity of marker insertion sites (138). In this particular study, the largest deletion observed among a population of 47 individuals was 59 bp. However, a potential bias existed in this study because only marker insertion events within pre-determined sections of pre-determined genes that could be successfully amplified were detected and analyzed. Analysis of a larger number of insertion events selected on a random basis would more accurately represent the nature of insertional mutagenesis events in *Chlamydomonas* using minimal selectable marker genes. To this end, one of the minimal AphVIII gene fragments containing a partial gene promoter sequence and no gene terminator sequence (Fig. 3-3) was excised from pSTN26 using the MfeI and EcoRI enzymes and the resulting linear fragment was introduced into cells via electroporation. Of the resulting paromomycin resistant colonies, 32 were randomly chosen for analysis. The data obtained from the resulting inverse PCR products corroborated the aforementioned observations made in *Chlamydomonas* (138) as in every sample, there were no significant genomic DNA deletions at the site of

integration. Moreover, 18 of the insertions (56%) occurred in intergenic or 3` UTR regions (Table 3-3).

**A**

| Gene Affected | Δ gDNA (bp) | Insertion Site |
|---|---|---|
| Cre17.g717850 | 4 | Promoter/Intergenic |
| Cre09.g395150 | 0 | 3' UTR |
| Cre10.g433700 | 1 | 3' UTR |
| Cre08.g384900 | --- | 3' UTR |
| Cre07.g346650 | 6 | 3' UTR |
| Cre10.g460850 | 0 | Promoter/Intergenic |
| Cre10.g460750 | 0 | 5' UTR |
| Cre12.g512788 | 0 | Promoter/Intergenic |
| Cre01.g033300 | 3 | 3' UTR |
| Cre06.g6388 | --- | 5' UTR |
| Cre04.g226700 | --- | Exon 1 of 18 |
| Cre01.g001350 | --- | 5' UTR |
| Cre14.g14763 | --- | Promoter/Intergenic |
| Cre07.g330300 | --- | 3' UTR |
| Cre08.g370550 | --- | Intron 1 of 12 |
| Cre08.g384900 | --- | 3' UTR |
| Cre01.g045400 | 0 | Promoter/Intergenic |
| Cre01.g045400 | 4 | Promoter/Intergenic |
| Cre09.g416900 | 0 | 5' UTR |
| Cre07.g7791 | 3 | 3' UTR |
| Cre05.g238250 | 0 | 5' UTR |
| Cre01.g022200 | --- | Promoter/Intergenic |
| Cre02.g111550 | 10 | Exon 16 of 25 |
| Cre14.g617400 | 1 | Exon 1 of 2 |
| Cre06.g305750 | --- | Intron |
| Cre06.g310200 | 3 | 5' UTR |
| Cre16.g16700 | 0 | 5' UTR |

| | | |
|---|---|---|
| Cre03.g190950 | --- | Terminator/Intergenic |
| Cre02.g105400 | 2 | 3' UTR |
| Cre01.g000100 | 2 | Exon 13 of 13 |
| Cre10.g433700 | 0 | 3' UTR |
| Cre03.g2901 | 2 | Exon 1 of 6 |

**B**

| Insertion Site | Events (n = 32) |
|---|---|
| Exon | 5 |
| Intron | 2 |
| 5` UTR | 7 |
| 3` UTR | 10 |
| Intergenic/Pro | 8 |

**Table 3-3: Analysis of inverse PCR products obtained from colonies transformed with an MfeI-EcoRI fragment derived from pSTN26. A)** The extent to which the gDNA at the insertion site was truncated (in bps) and the type of genomic domain disrupted is listed for each transformant analyzed. --- indicates that inadequate sequence data were obtained to fully characterize the sample. **B)** Tabulation of the number of times an insertion occurred in a particular type of genomic domain.

The observations made here and by others that marker insertion is not accompanied by extensive truncation or modification of the genomic insertion site under the prescribed circumstances are somewhat unexpected. Analysis of mutants generated by many previous insertional mutant studies conducted in *Chlamydomonas* has revealed much more extensive insertion site modifications in most mutants (58,138,144,149). However in all of these studies, extensive insertion site truncation always coincided with the use of entire plasmids that had been linearized prior to transformation. Furthermore, analysis of mutants generated by

the introduction of linear fragments into cells via electroporation (Tables 3-3, 3-5 and [138]) has revealed that extensive insertion site modification rarely occurs under these circumstances. In light of these data, it may have been that the large amounts of non-essential DNA present in the plasmids used in previous studies were responsible for the large insertion site deletions observed. Moreover, it would be of interest to determine if introduction of supercoiled, circular DNA molecules into cells via electroporation significantly affects the nature of resulting insertion events, relative to transformation with linear DNA molecules. Accordingly, the 2,482 bp PvuII portion of pSTN26 (Fig. 3-3A), bearing hundreds of non-essential nucleotides at both termini, was purified and treated with T4 DNA ligase. Only a portion of this plasmid was used as the mutagen to facilitate generation of minimally sized inverse PCR template molecules. This fragment also contained at either end 157 bp end 249 bp of vector-derived sequences of a lower GC content than the *Chlamydomonas* genome (8). The resulting DNA was then used to generate insertional mutants via electroporation. Analysis of inverse PCR products amplified from 20 randomly chosen colonies resulting from transformation with this molecule again revealed minimal genomic DNA modification at the site of integration and a high ratio of intergenic and 3` UTR integration events (9 out of 20 or 45%) (Table 3-4A & B). All deletions observed except one were smaller than 41 bp. The large 2 MB deletion observed was most likely not legitimate, or resulted from rearrangement of the genome during transformation, since such a large deletion would presumably lead to lethality. Importantly, the marker DNA termini were much more extensively truncated in almost every sample (Table 3-4A). The amount of

non-coding or otherwise expendable DNA sequence flanking the essential AphVIII coding region in the marker DNA used was 978 bp on the 5` promoter end and 515 bp on the 3` terminator end (Fig. 3-3). The promoter region was truncated at least 90 bp, and up to 779 bp, in all but 2 samples in which there were no apparent truncations. In general, the smaller terminator region was less extensively truncated with 3 samples showing deletions of less than 7 bp and 4 more showing truncations between 16 and 83 bp. The remaining 13 samples all had 3` truncations ranging from 178-495 bp. Therefore, DNA sequence that flanks essential selectable marker gene DNA are highly susceptible to truncation during transformation. Moreover, neither the inclusion of such extraneous sequences nor the circularity of transforming DNA appears to increase the extent to which genomic DNA flanking foreign DNA insertion sites is truncated during electroporation. Additionally, these data further fortify two conclusions made here and by others. Firstly, that the vast majority of insertion events generated by introduction of foreign DNA into *Chlamydomonas* via electroporation do not lead to extensive insertion site modification and, secondly, that a significant proportion of the insertion events generated with full genes markers occur in intergenic or 3` UTR regions (138).

**A**

| Gene Effected | Δ gDNA (bp) | Δ 5` (bp) | Δ 3` (bp) | Insertion Site |
|---|---|---|---|---|
| Cre12.g531850 | 0 | 90 | 344 | Promoter (5` UTR+400) |
| Cre03.g186500 | 8 | 420 | 5 | 3` UTR/Last Exon |
| Cre14.g618800 | 3 | 326 | 3 | Promoter (5` UTR+100) |
| Cre02.g147700 | 5 | 605 | 388 | 3` UTR |
| Cre06.g270050 | 7 | 2 | 6 | Promoter (5` UTR+350) |
| Cre12.g549550 (TUB2) | 40 | 620 | 353 | Exon - Last of TUB2 |
| Cre03.g190750 | 3 | 320 | 298 | 3` UTR |
| Cre03.g153400 | --- | 0 | 444 | Exon 13 of 13 |
| Cre03.g182500 (SRP72) | 5 | 406 | 178 | 3` UTR |
| Cre10.g440200 (SMC5B) | 22 | 94 | 30 | Intron 15 of 26 |
| Cre07.g346100 | 11 | 779 | 63 | 5` UTR |
| Cre17.g725150 | 10 | 479 | 366 | Exon 17 of 20 |
| Cre04.g215050 and g225750 | 2 MB | 304 | 398 | Exon 3 of 7<br>Exon 14 of 14 |
| Cre12.g487800 | 3 | 475 | 83 | Exon 5 of 5 |
| Cre16.g16818 | 13 | 575 | 418 | 3` UTR |
| Cre13.g605750 | --- | 625 | --- | 3` UTR |
| Cre05.g5380 (BCS1) | 31 | 335 | 477 | Exon 5 of 10 |
| Cre13.g584400 (FAP189) | 0 | 379 | 16 | Exon 7 of 9 |
| Cre02 (Intergenic) | 5 | 353 | 495 | Intergenic |
| Cre03.g157050 | 4 | 191 | 426 | Intron 13 of 15 |

**B**

| Insertion Site | Events (n = 20) |
|---|---|
| Exon | 8 |
| Intron | 2 |
| 5` UTR | 1 |
| 3` UTR | 5 |
| Intergenic | 4 |

**Table 3-4: Analysis of inverse PCRs obtained from colonies transformed with the PvuII portion of pSTN26**. **A)** The extent to which the gDNA at the insertion site and the termini of the marker fragment were truncated during transformation is indicated in the middle columns. The type of genomic element disrupted by the insertion event is also provided for each transformant. --- indicates that inadequate sequence data were obtained to fully characterize the sample. **B)** Tabulation of the number of times insertions occurred in each type of genomic element.

**Generation of insertional mutant**
**promoter positively biases insertion events towards transcribed genomic loci**

Because a large portion of insertion events can occur in regions that may not significantly affect the functionality of any gene, it would be beneficial to bias insertion events towards transcribed regions more likely to produce mutants bearing significant functional impediments when disrupted. It was documented above that the ΔPro/Ter AphVIII minimal selectable marker gene alone was able to generate a small but significant number of paromomycin resistant colonies (Fig. 3-3B). This positive outcome is likely the result of AphVIII landing downstream of a native promoter element or inserting to produce an in-frame chimeric gene capable of producing a fusion protein that has paromomycin phosphotransferase activity. This may mean that transformation with ΔPro/Ter AphVIII could significantly bias insertion events towards genic regions. To test this hypothesis, ΔPro/Ter AphVIII DNA restriction fragments or PCR products were introduced into cells via electroporation and primary transformants were selected on paromomycin-

containing medium. Sequence analysis of the inverse PCR products obtained from 44 random samples revealed that 42 insertions had occurred in genic regions (5` UTR, exon or intron) (Table 3-5). Because so few mutants generated in this fashion contained non genic or 3` UTR insertions, the value of the resulting mutant libraries are significantly higher compared to libraries generated with full-length marker genes.

**A**

| Mutagen | Gene Affected | Δ gDNA | Orientation | Insertion Site |
|---|---|---|---|---|
| Aph8 PCR | Cre10.g10637 | 7 | Sense | Exon 5 of 6 |
| Aph8 PCR | Cre08.g8723 | --- | Sense | Intron 1 of 4 |
| Aph8 PCR | Cre12.g12842 | 4 | Sense | Exon 1 of 3 |
| Aph8 PCR | Cre06.g6016 | 7 | Sense | Exon 20 of 20 |
| Aph8 PCR | Cre10.g10637 | 7 | Sense | Exon 5 of 6 |
| Aph8 PCR | Cre14.g619050 | 6 | Sense | Exon 5 of 5 |
| Aph8 PCR | Cre02.g093600 | 0 | Sense | Exon 1 of 4 |
| Aph8 PCR | Cre02.g1771 | 4 | Sense | Exon 4 of 5 |
| Aph8 PCR | Cre02.g1771 | 6 | Sense | Exon 4 of 5 |
| Aph8 PCR | Cre01.g044600 | 0 | Sense | Exon 15 of 16 |
| Aph8 PCR | Cre11.g11681 | 2 | Sense | Exon 6 of 6 |
| Aph8 PCR | Cre07.g342350 (PDE14) | 0 | Sense | Exon 6 of 13 |
| Aph8 PCR | Cre06.g6300 | 31 | N/A | Promoter |
| Aph8 PCR | Cre12.g12482 | 4 | Sense | Exon 2 of 3 |
| Aph8 PCR | Cre01.g054150 | 7 | Sense | 5' UTR |
| Aph8 PCR | Cre08.g8349 | 9 | Sense | Intron |
| Aph8 PCR | Cre17.g17537 | 14 | Sense | Exon 8 of 20 |
| 1 - Aph8 (R Frag) | Cre12.g544400.t1.2 | 3 | Sense | Intron 1 of 10 |
| 2 - Aph8 (R Frag) | Cre14.g617400 (HSP22F) | 4 | Sense | Exon 1 of 2 |
| 4 - Aph8 (R Frag) | Cre13.g580000.t2.1 | 0 | Sense | Intron 1 of 3 |
| 6 - Aph8 (R Frag) | Cre05.g5380.t1 (BCS1) | 61 | Sense | Exon 8 of 10 |
| 7 - Aph8 (R Frag) | Cre02.g105550.t1.3 | --- | Anti | Intron 4 of 12 |
| 8 - Aph8 (R Frag) | Cre09.g9176.t1 | --- | Sense | Intron 1 of 16 |
| 9 - Aph8 (R Frag) | Cre12.g502000.t1.2 (FAP253) | 14 | Sense | Exon 7 of 7 |
| 10 - Aph8 | Cre12.g521400.t1.2 | 5,032 | Sense | Exons (Most of |

| | | | | |
|---|---|---|---|---|
| (R Frag) | and CLPP2 | | | CLPP2) |
| 11 - Aph8 (R Frag) | Cre09.g9251.t1 | 8 | Sense | 5` UTR |
| 14 - Aph8 (R Frag) | Cre03.g158850.t1.3 | 10 | Sense | 5` UTR |
| 16 - Aph8 (R Frag) | Cre10.g443950.t1.3 | 0 | Sense | Exon 1 of 4 |
| 17 - Aph8 (R Frag) | Cre09.g395750.t1.3 | 11 | Sense | Exon 23 of 24 |
| 18 - Aph8 (R Frag) | Cre13.g589450.t1.2 | 8 | Sense | 5` UTR |
| 19 - Aph8 (R Frag) | Cre09.g394700.t1.3 | 11 | Sense | 5` UTR |
| 20 - Aph8 (R Frag) | Cre12.g557850.t1.3 | 0 | Sense | 5` UTR |
| 21 - Aph8 (R Frag) | Cre16.g653700.t1.3 | 0 | Sense | 5` UTR |
| 22 - Aph8 (R Frag) | Cre13.g577950.t1.2 (VSP60) | 6 | Sense | 5` UTR |
| 23 - Aph8 (R Frag) | Cre13.g589450.t1.2 | 2 | Sense | Exon 1 of 6 |
| 24 - Aph8 (R Frag) | Cre08.g379400.t1.3 | --- | Sense | Exon 1 of 1 |
| Aph8 (R Frag) | Cre12.g484650 | --- | Sense | Exon 8 of 9 |
| Aph8 (R Frag) | Cre13.g589450 | 0 | Sense | Exon 2 of 6 |
| Aph8 (R Frag) | Cre13.g589450 | 0 | Anti | 5' UTR |
| Aph8 (R Frag) | Cre01.g398 | 197 | Anti | 5' UTR |
| Aph8 (R Frag) | Cre05.g5380 | --- | Anti | Exon |
| Aph8 (R Frag) | Cre01.g045400 | --- | Sense | Intron 2 of 8 |
| Aph8 (R Frag) | Cre01 | 2 | N/A | Intergenic/Terminator |
| Aph8 (R Frag) | Cre12.g510034 | --- | Sense | Exon 3 of 6 |

**B**

| Insertion Site | Events (n = 44) |
|---|---|
| Exon | 25 |
| Intron | 7 |
| 5` UTR | 10 |
| 3` UTR | 0 |
| Intergenic/Pro | 2 |

**Table 3-5: Analysis of inverse PCRs obtained from ΔPro AphVIII fragment transformants. A)** The extent of insertion site modification, the type of genomic element interrupted and the orientation of the marker relative to the disrupted gene are provided for each transformant. Where --- occurs, inadequate sequence data was available to fully analyze the sample. 'R frag' samples were generated by transformation with restriction fragments. The numbered samples represent the initial samples used to evaluate the optimized methods described in the text (Electrophoresed Inverse PCR products shown in Figure 3-4) **B)** Summary of insertion events based on the type of genomic element that was disrupted.

## A commercial restriction enzyme buffer significantly improves transformation efficiency when included during electroporation of *Chlamydomonas* cells

Although insertional mutant libraries generated with the ΔPro/Ter AphVIII minimal selectable marker gene are likely more valuable than libraries generated with full-length selectable marker genes, transformation with this marker produces roughly one-tenth the number of colonies generated by the full-length AphVIII gene (Fig. 3-3B). This may be a limitation for those interested in generating insertional mutant libraries that provide extensive coverage of the entire genome. Thus, it would be useful to increase transformation efficiency beyond that which can be achieved using standard electroporation protocols (30).

During the course of numerous transformation experiments using electroporation, it was discovered that a commonly used restriction enzyme reaction buffer,Fermentas Fast Digest Green Buffer® (FDG), exerts a significant positive effect on the rate of *Chlamydomonas* genetic transformation. When the standard electroporation solution was supplemented with this buffer, bringing its final concentration to 0.7% (vol./vol.), the transformation efficiency achieved by the ΔPro/Ter AphVIII gene was increased by approximately 8 fold (Fig. 3-5). Transformation of several samples of *Chlamydomonas* with 250 ng of the ΔPro/Ter AphVIII in standard electroporation buffer lacking FDG buffer produced on average approximately 100 colonies whereas roughly 800 colonies were produced by the same amount of ΔPro/Ter AphVIII gene when FDG buffer was included during electroporation. Furthermore, the FDG buffer exerted an apparent "electroprotectant" effect on cells during electroporation as cells electroporated in the presence of FDG displayed higher survival rates compared to untreated cells 20 hours post transformation (Fig. 3-6). These observations demonstrate that significant improvements of *Chlamydomonas* transformation can be obtained by the simple addition of low concentrations of FDG buffer during electroporation. This suggests that in those cases in which maximal numbers of *Chlamydomonas* transformants are important to the success of an experiment, the incorporation of FDG buffer into the electroporation mixture will be a significant benefit.

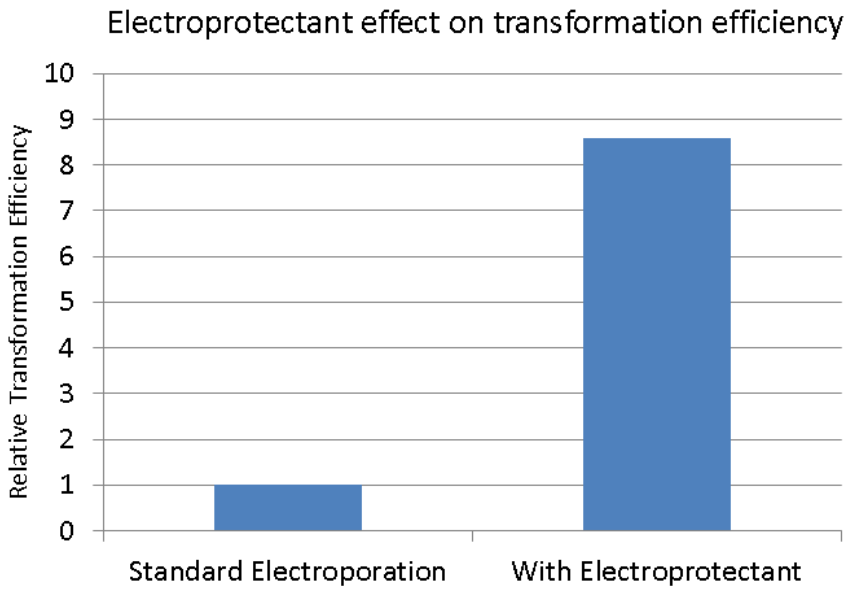



**Figure 3-5**: **Electroprotectant effect on transformation efficiency.** Comparison of transformation efficiencies obtained using standard electroporation conditions or by supplementation with FDG buffer prior to electroporation. The transformation efficiency of the standard electroporation protocol was set to 1.

**Figure 3-6: Inclusion of FDG buffer during electroporation affects cellular health post-transformation.** Four replicate samples were subjected to electroporation using standard conditions. The two samples on the right were treated with FDG buffer prior to transformation whereas the two on the left were not. This image was taken 20 hours post electroporation.

## Discussion

Insertional mutagenesis is an important tool that is used to create random knockouts of genes throughout the genome of an organism of experimental interest. Studies for such mutants can provide invaluable information in regard to the function of a particular gene and its role in various metabolic and physiological processes. Quickly establishing the identity of the gene involved in the knockout mutation is a critical first step in such studies and numerous techniques for accomplishing this task have been devised. Inverse PCR represents one of the best methods available because it is simple, cost effective and is the only PCR based method that amplifies in single PCR reaction both genomic DNA regions flanking the insertion site for the selectable marker gene employed(89). Other methods such as TAIL PCR require the synthesis and empirical testing of numerous semi-random primers which often produce nonspecific background products (104,140). For Adapter Ligation-mediated PCR to perform optimally, adapters containing nucleotides with costly chemical modifications must be utilized, making this method significantly more expensive than TAIL PCR or inverse PCR (89,139). Here, an optimized inverse PCR method was developed that is highly efficient. Through the use of a minimal selectable marker DNA fragment and a strategically designed restriction enzyme cocktail, the size range of the final inverse PCR templates was optimized. In *Chlamydomonas*, the average fragment size resulting from digestion with this cocktail is about 240 bp. This size approximates the average fragment size created by restriction enzymes such as AluI that have DNA recognition sites of only 4 bp (Table 3-1B). The use of enzymes that recognize 4 bp sequences to fragment

the genome prior to inverse PCR would eliminate the need to use a cocktail consisting of multiple enzymes. However, it is difficult to find selectable marker genes that do not contain such sites. The ΔPro/Ter AphVIII gene alone contains 4 AluI recognition sites (Table 3-1B). Most commonly used selectable marker genes, especially larger ones, likely would need to be synthesized *de novo* to avoid such sites. The restriction enzyme cocktail utilized in this study generates fragments 1 kb or larger <2% of the time and 12 bp or smaller <5% of the time (Table 3-2). Therefore, in the majority of cases, an inverse PCR product will be amplified with ease and both genomic DNA flanking regions will be large enough to map unambiguously. An additional benefit of using a multicomponent cocktail consisting of enzymes that recognize 6 bp sequencesrather than individual enzymes that recognize 4 bp sequences is that in the unlikely event that both flanking regions are too small to map or too large to amplify, the cocktail can be easily modified without risking digestion of the selectable marker DNA.

A trial of this method consisting of an analysis of 24 independent transformants revealed inverse PCR products were generated in every case. Additionally, all of these PCR products were smaller than 1.5 kb, indicating that the restriction cocktail was successful in minimizing the final inverse PCR product size (Fig. 3-4). Of these 24 samples, 21 out of the 22 samples for which reliable sequencing data was obtained could be mapped to the genome. However, the genomic DNA flanking regions observed in two inverse PCR reactions aligned to the genome but at two non-adjacent loci. This may have been due to recombination between genomic elements during transformation – a phenomenon that has been observed previously

(142). Thus, it is possible that these are valid inverse PCR products. In confusing situations such as these, subsequent analyses can be performed to determine which loci are actually disrupted. However, it should be emphasized that because inverse PCR amplifies both flanking regions simultaneously, in the vast majority of mutants, the available sequence data will minimize the occurrence of such phenomena.

Our optimized inverse PCR method for identification of foreign DNA insertion sites was also used to confirm past observations with respect to the extent of nucleotide deletions at marker gene insertion sites and to determine if alterations in marker gene length and composition affect the integrity of genomic DNA at insertion sites during transformation. Numerous studies in *Chlamydomonas* have demonstrated the generation of insertional mutants containing a high frequency of large genomic DNA deletions of several kbp at marker insertion sites (58,138,144,149). In all of these cases however, whole, linearized plasmids were used as mutagens. In addition, the method used to integrate the marker DNA into the genome was either particle bombardment (29) (used in (58)) or agitation of cells in the presence of DNA and glass beads (31) (used in (138,144,149)). From these studies it is difficult to ascertain whether it was the transformation method, nature of the transforming DNA or intrinsic properties of *Chlamydomonas* that most significantly contributed to the large insertion site deletions observed. In addition, it has been observed here and by others that introduction of linear DNA fragments into *Chlamydomonas* cells via electroporation does not result in extensive insertion site modification (Tables 3-3 and 3-5) [14]. To further investigate this phenomenon,

we used for electroporation of *C. reinhhardtii* a large DNA fragment (the 2, 482 bp PvuII fragment from pSTN26, Fig. 3-3) resembling a small plasmid in size and composition, in that it was circularized before use and contained hundreds of expendable nucleotides flanking the full-length selectable marker gene. As a point of caution in interpreting the data obtained, it should be noted that once this PvuII fragment was circularized, no measures were taken to ensure that all of the unligated, linear fragments were excluded from the population of DNA molecules subsequently used to generate transformants. Therefore, an unknown, but likely small number of the transformants analyzed (results summarized in Table 3-4) may have been generated by DNA molecules that were linear when they entered the cell. Regardless, we observed that electroporation with this large molecule, bearing a large number of non-essential nucleotides, did not increase the extent of insertion site DNA truncation (Table 3-4). To clarify these results, future experiments will be required to analyze the nature of transformation events that occur after introduction of purified, circular molecules into cells. Nonetheless, from the available data it appears that truncation of genomic DNA at marker insertion sites concomitant with transformation may be most heavily influenced by the method used to introduce DNA into cells (i.e. electroporation vs. glass bead or particle bombardment) rather than the nature of the transforming DNA.

In a recent study, Gonzalez-Ballester et al. (138), used a relatively small PCR product consisting of a minimal AphVIII gene to transform cells via electroporation. In this study, the largest insertion site deletion observed out of a population of 47 individuals was surprisingly only 34 bp. However, a potential bias

existed in this study because only mutants harboring insertions in pre-determined genomic loci were isolated. In our studies, the extent to which genomic DNA in the vicinity of selectable marker gene insertions was truncated could be determined for a total of 76 randomly selected transformants. All of these transformants were generated by introduction of either PCR products or restriction fragments into cells via electroporation. Only 4 of these individuals contained insertion site deletions larger than 40 bp. Therefore, our data corroborate the observations of Gonzalez-Ballester et al. and provide further evidence that transformation of *Chlamydomonas* cells via electroporation usually does not lead to large genomic DNA truncations irrespective of the size or composition of the marker gene used.

Although it is known that DNA introduced into *Chlamydomonas* cells by common transformation methods is often subject to rearrangement or truncation (29-31,58), the concomitant post-transformation fate of genomic DNA in the vicinity of insertion site has not been extensively explored. Interestingly, our studies show that when a selectable marker gene containing a few hundred bp of non-essential nucleotides of a GC content lower than that of the *Chlamydomonas* genome was used to generate transformants via electroporation, each of its termini were heavily truncated in nearly every case while the genomic DNA at the insertion was not (Table 3-4). It is possible that foreign DNA entering the cells via electroporation and other methods is vulnerable to various forms of nucleolytic attack prior to integration into the genome. The precise mechanism by which foreign DNA is integrated into the *Chlamydomonas* genome during transformation is unknown but in other

organisms where transforming DNA is integrated illegitimately, the process is thought to rely on the non-homologous end joining (NHEJ) DNA repair pathway (150,151). When this pathway encounters free DNA ends, it attempts to ligate them together regardless of their origins or homology (152,153). If two adjacent DNA ends happen to have complementary 5` or 3` ssDNA overhangs, they are efficiently ligated without mutation. Alternatively, if two incompatible ssDNA overhangs or two blunt ends are encountered that cannot be easily ligated, the NHEJ machinery will commonly truncate or less frequently extend the ends until successful ligation can be achieved (153,154). It light of this information, it is difficult to understand why during genetic transformation using electroporation there is not the same type of extensive modification of genomic DNA at the site of foreign DNA insertion as is seen when the particle bombardment and glass bead methods are employed. These findings suggest that extensive insertion site DNA modifications should rarely be expected during transformation of *Chlamydomonas* using electroporation while extensive truncation of exogenously supplied DNA is likely a feature common to all methods of genetic transformation used to obtain integration of transgenes into the *Chlamydomonas* nuclear genome.

We observed an interesting, but puzzling phenomenon when nonpurified PCR products were used to generate insertional mutants. Although only a small population of 15 transformants were analyzed, 7 of them contained at least two copies of the marker gene that had been inserted tandemly, in a head-to-tail fashion. Because the most prominent inverse PCR product obtained from each of these samples corresponded to the junction between the repeated markers, the

genomic loci of these insertion events could not be determined. Numerous attempts to obtain inverse PCR products from these samples using various conditions failed to achieve success. This phenomenon has been observed previously in *Chlamydomonas* and was reported to affect up to 20% of all insertional mutants (58,103). Transformation of other organisms such as fungi (94,95,155), plants (156), mammals (157,158) and even fish (159) can also result in a similarly high rate of tandem marker insertion events. In contrast, we never observed this phenomenon when the selectable marker gene fragment used for transformation was a plasmid-derived restriction fragment. Because this phenomenon occurs at high frequencies, rendering the resulting marker gene insertion events intransigent to PCR-based genetic analyses, further investigation of this intriguing phenomenon may be warranted to elucidate its cause and perhaps develop methods to avoid it.

Transformation with full-length selectable marker genes results in many insertions occurring in putative promoter regions, intergenic regions and 3` UTRs (90,138) that may or may not produce a measurable phenotype. Because it is difficult to accurately predict promoter elements, that can often be relatively small, it is likely that some insertions of foreign DNA into putative promoter regions do not result in observable mutant pheotypes. The same logic applies even more strongly to intergenic regions given that a much larger portion of the DNA in these regions likely can endure insertions without triggering a mutant phenotype. In addition, in this study it has been shown that 3` UTR and terminator sequences are completely expendable for optimal AphVIII gene transformation efficiency when the gene is under the control of the PsaD gene promoter region. This has also been observed

for many other native and foreign transgenes indicating that many 3` UTR disruptions are tolerable. Therefore, insertional mutant libraries generated with full gene markers may contain many individual mutants that will not yield significant information related to the function of any genetic elements. Because so many of the full-length marker gene insertion events characterized in the present study and by others occur in regions such as these, the minimal ΔPro/Ter AphVIII gene was tested as the mutagen to determine if a bias could be created for insertion into genomic regions containing introns and exons. Even without a promoter, AphVIII was able to generate a significant number of transformants (Fig. 3-3B) (33). When these transformants were analyzed, there emerged evidence for a strong bias for genic insertions. Most likely this bias was created because insertion events that did not result in linkage of the ΔPro/Ter AphVIII coding region to a native promoter were not viable under selection due to inadequate APHVIII gene expression. Of 44 individual mutants analyzed that were generated with the ΔPro/Ter AphVIII gene, only 2 contained an insertion that did not occur within a 5` UTR, intron or exon (Table 3-5). Among the 42 genic insertion events, all but 4 of them resulted in integration of the marker in the sense orientation with respect to the gene they disrupted. Therefore, it is likely that integration into exon sequences resulted in the production of fusion proteins displaying paromomycin phosphotransferase activity. In the instances in which the AphVIII coding region landed in intron or 5` UTR sequences, the native reading frames and/or splicing junctions obviously had to align to allow adequate APHVIII gene expression. These data strongly support the hypothesis that generating transformants with the ΔPro/Ter AphVIII coding region

strongly favors the integration of this “naked” coding region into an active gene to allow it to become a functional antibiotic resistance gene. At a practical level, our results clearly demonstrate that in attempts to recover mutants via insertional gene inactivation, use of a promoter-less, selectable marker gene coding region, such as the ΔPro/Ter AphVIII coding region, greatly improves the efficiency and frequency with which useful gene knockouts can be recovered.

Although insertional mutants generated with a promoter-less gene appear more likely to contain debilitating mutations, the naked ΔPro AphVIII coding region produces fewer colonies than its full-length gene counterpart (Fig. 3-3). However, our observations during multiple transformation experiments led to the discovery of the Fermentas Fast Digest Green Buffer® (FDG) as a presumed electroprotectant (EP) solution that increased transformation efficiency by more than 8-fold (Fig. 3-5). The presence of the EP buffer during electroporation of *Chlamydomonas* cells led to significantly improved cell viability compared to cells subjected to electroporation in the absence of the EP (Fig. 3-6). The mode of action of the EP in improving cell viability and transformation rates is unknown. However, we have noted that inclusion of the EP during electroporation significantly lowers the observed time constant. The time constant is the time it takes for the initial voltage of an electroporation sample to drop to 33%. Thus, the EP apparently lowers the electrical resistance of the solution and allows the current to flow through the sample faster. Other measures that lower the time constant by similar extents do not increase transformation efficiencies. Thus it is likely that the FDG buffer may contain one or more components that influence such parameters as the speed or

efficiency with which DNA enters to cell, the amount of time electroporation-generated pores in the cell membranes remain open, the number of pores generated per cell, the efficiency with which pores are sealed, the stability or mobility of the foreign DNA taken into the cell, or any of several other factors that might influence cell health and viability following electroporation. Unfortunately, the composition of FDG buffer is unknown and determination of buffer components responsible for increased transformation rates cannot be determined for the time being. Regardless of how this protective effect is achieved, a practical positive outcome of these observations is the ability of the EP buffer to increase transformation efficiency of *Chlamydomonas* cells used in the production of large insertional mutant libraries using selectable marker coding regions that lack promoters. From a broader perspective it can be noted that transformation rates surpassing $10^8$ CFU/µg of DNA can be achieved by electroporation of other microbes such as yeast (61) while transformation of *Chlamydomonas* cells via electroporation is approximately $10^4$ less efficient (30). Thus, it appears that there may be much potential to improve the transformation efficiency of *Chlamydomonas* and other important algal species via electroporation. Indeed, in a recent study, the combination of microfluidics techniques with electroporation resulted in a significant increase in transformation efficiency of *Chlamydomonas* by up to three orders of magnitude (160). Perhaps combining this advancement with FDG buffer or its components can increase transformation efficiency still further.

Although the total number of transformants that can be generated with full-length selectable marker genes will always be higher than with promoter-less

selectable marker gene coding regions, it is important to remember that a significant fraction of the transformants obtained with the full-length gene will not lead to knockout of a gene – whereas a high proportion of mutants obtained using the promoter-less selectable marker gene coding region will contain disruptions of transcribed regions of the genome. Conversely, insertion events resulting from use of full-length selectable marker gene constructs will be more recalcitrant to inverse PCR analysis and will require considerably more effort to identify within pools of transformants containing many individuals with inconsequential mutations. Finally, it should be noted that the improvements in strategies for generating and characterizing insertional mutant libraries described here likely can be adapted for use in any organism for which insertional mutagenesis is feasible.

## Materials and Methods

### Marker Construction

The AphVIII cDNA was isolated via PCR from pSI103 provided by Peter Hegemann. To potentially enhance transcriptional activity in the event this cDNA landed in a conditionally inactive or low expression gene, the RbcS2 1$^{st}$ intron, a known transcriptional enhancer element (115) was placed into the AphVIII cDNA via overlap extension PCR (161) to avoid the introduction of artificial restriction sites. This intron was placed into the AphVIII cDNA at a position that enabled the creation of ideal intron exon boundaries at each end of the intron without altering any AphVIII cDNA sequence. This marker was also placed under the control of the PsaD promoter and RbcS2 3` UTR and terminator region by first adding via PCR

NdeI and EcoRI sites to the 5` and 3` ends, respectively, of the AphVIII gene and EcoRI and ApaI sites to the 5` and 3` ends, respectively, of the RbcS2 terminator. These two PCRs were then digested and ligated together into the pGenD [45] cassette to create pSTN26. Fragments of this gene were used to generate the mutants analyzed in this study.

**Genomic Restriction Enzyme Analyses**

The *Chlamydomonas* genome sequence was downloaded from www.phytozome.net in October 2011. Because the genome file was relatively large, only the 1$^{st}$ chromosome, representing ~9.8 MB or ~9% of the ~111 MB genome, was selected for further analysis. The number of restriction enzyme cut sites found in chromosome #1 by a simple motif search was used to calculate the number of expected restriction sites in the whole genome for each restriction scenario. Subsequent restriction pattern analyses were conducted *in silico* with the program pDRAW32 (http://www.acaclone.com/). To analyze and produce images representing these *in silico* restriction digest patterns, a 700 kb portion of Chromosome #1 was used in addition to the entire 1$^{st}$ chromosome. It is important to note that the average restriction fragment size calculated from actual restriction patterns of this 700 kb section closely approximates the average restriction fragment size calculated from dividing the total size of chromosome #1 by the number of cut sites a particular restriction enzyme produces in the 64% GC *Chlamydomonas* genome. Therefore, the 700 kb fragment used for the *in silico* analyses appears to be representative of Chromosome #1. When calculating the

odds of actual restriction fragments for various digestion schemes falling into specific size classes, certain unrealistic restriction fragments were excluded or otherwise accounted for. For instance, there remain large unsequenced regions of the genome of multiple kb that are indicated as N nts that are not digested *in silico* by any enzyme. This resulted in the *in silico* production of many fragments larger than 1 kb that were ignored. In addition, restriction fragments produced several *in silico* would have required digestion of precursor fragments that are smaller than the minimum sized DNA fragment needed for efficient digestion. These too were ignored.

**Cell strains, culturing conditions and transformation procedure**

The *Chlamydomonas* strain CC3491 was used for the generation of all insertional mutants. Standard electroporation conditions were used for all transformations [41]. Briefly, cells were grown in TAP medium (1) under continuous illumination of ~80 μE/m$^2$/sec and rotation at 140 RPM to a density of ~3 x 10$^6$ cell/mL. The cells were then concentrated via centrifugation and brought to a concentration of 4 x 10$^8$ cells/ml in TAP + 60 mM sucrose. For each individual transformation, various DNA fragments were used to transform 1 x 10$^8$ cells in a volume of 250 μL. Cells were recovered in TAP + 60 mM sucrose medium shaking at 140 RPM at room temperature under constant illumination for 20-24 hours before plating. When the standard electroporation buffer containing FDG buffer was used for transformation, 20 μl of 1X Fermentas Fast Digest Green ® buffer was added along with DNA to a standard 250 μl sample (containing 1 x 10$^8$ cells) less than 10

minutes before electroporation. No other aspect of the standard protocol outline above was altered when using the FDG buffer. For selection of cells stably transformed with the AphVIII gene, TAP plates were supplemented with 20 µg/mL paromomycin. Selection occurred at 25° C under constant illumination. Colonies were picked to TAP media 7-10 days post transformation for growth and subsequent collection of genomic DNA by previously described methods (1).

**Inverse PCR**

Approximately 1-4 µg of RNAse treated genomic DNA was used for each inverse PCR reaction. In a total volume of 15 µL of 1X FDG buffer, the genomic DNA as well as 0.25 µl each of FspI (or PvuII), NaeI, BsaAI and SfoI (all Fermentas fast digest enzymes ®) were combined. These samples were incubated for 2 hours at 37° C followed by a 10 min heat denaturation step at 80° C to inactivate the restriction enzymes. Because T4 DNA Ligase also functions in FDG buffer, no purification or buffer exchange was required before addition of this enzyme to the samples. Therefore, the next step consisted of addition of 5 µL of a T4 DNA ligase (Fermentas) solution containing 2.5 units of T4 DNA ligase and ATP at a concentration of 2 mM to the heat inactivated restriction digests. These samples were incubated for 1 hour at room temperature followed by a second denaturation step for 10 minutes at 75° C to inactivate the T4 DNA ligase. Next, the samples are diluted 1:10 with water and 1 µl of each was used as template for inverse PCR. At least 30 cycles with Phusion DNA polymerase was used for each PCR reaction. Two step PCR cycle was used consisting of 10 seconds at 98° C followed by 20

seconds per kb at 72° C. As most inverse PCR products created with this procedure were less than 2-3 kb in size, 40-60 seconds was generally the maximum extension time required. Inverse PCR reaction products were then subjected to electrophoreses on 1.2% agarose gels and amplified DNA bands were excised from the gel, purified and sequenced with the AphVIII gene-specific primers that also were used originally to generate the inverse PCR products: Aph8 R Met+130 (5` CCA GAG CTG CCA CCT TGA CAA ACA AC) and Aph8 F Met+753 (5` GGT ATC GGA GGA AAA GCT GGC GTT TTA CC).

## REFERENCES:


1. Harris, E. H. (1989) **The *Chlamydomonas* sourcebook: a comprehensive guide to biology and laboratory use**. *Academic Press, San Diego, CA*, 780 pp
2. Harris, E. H. (2001) ***Chlamydomonas* as a model organism**. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 363-406
3. Rochaix, J. D. (1995) ***Chlamydomonas reinhardtii* as the photosynthetic yeast**. *Annu Rev Genet* **29**, 209-230
4. Govorunova, E. G., and Sineshchekov, O. A. (2005) **Chemotaxis in the green flagellate alga *Chlamydomonas***. *Biochemistry (Mosc)* **70**, 717-725
5. Pan, J., Misamore, M. J., Wang, Q., and Snell, W. J. (2003) **Protein transport and signal transduction during fertilization in *Chlamydomonas***. *Traffic* **4**, 452-459
6. te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubesova, M., and Pysek, P. (2012) **The more the better? The role of polyploidy in facilitating plant invasions**. *Ann Bot* **109**, 19-45
7. Weiss-Schneeweiss, H., Emadzade, K., Jang, T. S., and Schneeweiss, G. M. (2013) **Evolutionary consequences, constraints and potential of polyploidy in plants**. *Cytogenet Genome Res* **140**, 137-150
8. Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Marechal-Drouard, L., Marshall, W. F., Qu, L. H., Nelson, D. R., Sanderfoot, A. A., Spalding, M. H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C. L., Cognat, V., Croft, M. T., Dent, R., Dutcher, S., Fernandez, E., Fukuzawa, H., Gonzalez-Ballester, D., Gonzalez-

Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P. A., Lemaire, S. D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A. M., Niyogi, K., Novoselov, S. V., Paulsen, I. T., Pazour, G., Purton, S., Ral, J. P., Riano-Pachon, D. M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S. L., Allmer, J., Balk, J., Bisova, K., Chen, C. J., Elias, M., Gendler, K., Hauser, C., Lamb, M. R., Ledford, H., Long, J. C., Minagawa, J., Page, M. D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A. M., Yang, P., Ball, S., Bowler, C., Dieckmann, C. L., Gladyshev, V. N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R. T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martinez, D., Ngau, W. C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. (2007) **The *Chlamydomonas* genome reveals the evolution of key animal and plant functions**. *Science* **318**, 245-250

9. Levine, R. P. (1968) **Genetic dissection of photosynthesis**. *Science* **162**, 768-771

10. Spreitzer, R. J., and Mets, L. (1981) **Photosynthesis-deficient mutants of *Chlamydomonas reinhardtii* with associated light-sensitive phenotypes**. *Plant Physiol* **67**, 565-569

11. Khrebtukova, I., and Spreitzer, R. J. (1996) **Elimination of the *Chlamydomonas* gene family that encodes the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase**. *Proc Natl Acad Sci U S A* **93**, 13689-13693

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

13. Dent, R. M., Han, M., and Niyogi, K. K. (2001) **Functional genomics of plant photosynthesis in the fast lane using *Chlamydomonas reinhardtii***. *Trends Plant Sci* **6**, 364-371

14. Rochaix, J. D. (2002) ***Chlamydomonas*, a model system for studying the assembly and dynamics of photosynthetic complexes**. *FEBS Lett* **529**, 34-38

15. Gutman, B. L., and Niyogi, K. K. (2004) ***Chlamydomonas* and *Arabidopsis*. A dynamic duo**. *Plant Physiol* **135**, 607-610

16. Wang, Y., and Spalding, M. H. (2006) **An inorganic carbon transport system responsible for acclimation specific to air levels of $CO_2$ in *Chlamydomonas reinhardtii***. *Proc Natl Acad Sci U S A* **103**, 10110-10115

17. Spalding, M. H. (2008) **Microalgal carbon-dioxide-concentrating mechanisms: *Chlamydomonas* inorganic carbon transporters**. *J Exp Bot* **59**, 1463-1473

18. Fukuzawa, H., Miura, K., Ishizaki, K., Kucho, K. I., Saito, T., Kohinata, T., and Ohyama, K. (2001) **Ccm1, a regulatory gene controlling the induction of a carbon-concentrating mechanism in *Chlamydomonas reinhardtii* by sensing $CO_2$ availability**. *Proc Natl Acad Sci U S A* **98**, 5347-5352

19. Borkhsenious, O. N., Mason, C. B., and Moroney, J. V. (1998) **The intracellular localization of ribulose-1,5-bisphosphate Carboxylase/Oxygenase in *Chlamydomonas reinhardtii***. *Plant Physiol* **116**, 1585-1591

20. Brueggeman, A. J., Gangadharaiah, D. S., Cserhati, M. F., Casero, D., Weeks, D. P., and Ladunga, I. (2012) **Activation of the carbon concentrating mechanism by $CO_2$ deprivation coincides with massive transcriptional restructuring in *Chlamydomonas reinhardtii***. *Plant Cell* **24**, 1860-1875

21. Fang, W., Si, Y., Douglass, S., Casero, D., Merchant, S. S., Pellegrini, M., Ladunga, I., Liu, P., and Spalding, M. H. (2012) **Transcriptome-wide changes in *Chlamydomonas reinhardtii* gene expression regulated by carbon dioxide and the $CO_2$-concentrating mechanism regulator CIA5/CCM1**. *Plant Cell* **24**, 1876-1893

22. Wang, Y., Duanmu, D., and Spalding, M. H. (2011) **Carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*: inorganic carbon transport and $CO_2$ recapture**. *Photosynth Res* **109**, 115-122

23. Michaelis, G., Vahrenholz, C., and Pratje, E. (1990) **Mitochondrial DNA of *Chlamydomonas reinhardtii*: the gene for apocytochrome b and the complete functional map of the 15.8 kb DNA**. *Mol Gen Genet* **223**, 211-216

24. Maul, J. E., Lilly, J. W., Cui, L., dePamphilis, C. W., Miller, W., Harris, E. H., and Stern, D. B. (2002) **The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats**. *Plant Cell* **14**, 2659-2679

25. Boynton, J. E., Gillham, N. W., Harris, E. H., Hosler, J. P., Johnson, A. M., Jones, A. R., Randolph-Anderson, B. L., Robertson, D., Klein, T. M., Shark, K. B. (1988)

**Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles**. *Science* **240**, 1534-1538

26. Remacle, C., Cardol, P., Coosemans, N., Gaisne, M., and Bonnefoy, N. (2006) **High-efficiency biolistic transformation of *Chlamydomonas* mitochondria can be used to insert mutations in complex I genes**. *Proc Natl Acad Sci U S A* **103**, 4771-4776
27. Randolph-Anderson, B. L., Boynton, J. E., Gillham, N. W., Harris, E. H., Johnson, A. M., Dorthu, M. P., and Matagne, R. F. (1993) **Further characterization of the respiratory deficient *dum-1* mutation of *Chlamydomonas reinhardtii* and its use as a recipient for mitochondrial transformation**. *Mol Gen Genet* **236**, 235-244
28. Purton, S. (2007) **Tools and techniques for chloroplast transformation of *Chlamydomonas***. *Adv Exp Med Biol* **616**, 34-45
29. Kindle, K. L., Schnell, R. A., Fernandez, E., and Lefebvre, P. A. (1989) **Stable nuclear transformation of *Chlamydomonas* using the *Chlamydomonas* gene for nitrate reductase**. *J Cell Biol* **109**, 2589-2601
30. Shimogawara, K., Fujiwara, S., Grossman, A., and Usuda, H. (1998) **High-efficiency transformation of *Chlamydomonas reinhardtii* by electroporation**. *Genetics* **148**, 1821-1828
31. Kindle, K. L. (1990) **High-frequency nuclear transformation of *Chlamydomonas reinhardtii***. *Proc Natl Acad Sci U S A* **87**, 1228-1232

32. Stevens, D. R., Rochaix, J. D., and Purton, S. (1996) **The bacterial phleomycin resistance gene ble as a dominant selectable marker in *Chlamydomonas***. *Mol Gen Genet* **251**, 23-30

33. Sizova, I., Fuhrmann, M., and Hegemann, P. (2001) **A *Streptomyces rimosus aphVIII* gene coding for a new type phosphotransferase provides stable antibiotic resistance to *Chlamydomonas reinhardtii***. *Gene* **277**, 221-229

34. Berthold, P., Schmitt, R., and Mages, W. (2002) **An engineered *Streptomyces hygroscopicus aph 7"* gene mediates dominant resistance against hygromycin B in *Chlamydomonas reinhardtii***. *Protist* **153**, 401-412

35. Meslet-Cladiere, L., and Vallon, O. (2011) **Novel shuttle markers for nuclear transformation of the green alga *Chlamydomonas reinhardtii***. *Eukaryot Cell* **10**, 1670-1678

36. Debuchy, R., Purton, S., and Rochaix, J. D. (1989) **The argininosuccinate lyase gene of *Chlamydomonas reinhardtii*: an important tool for nuclear transformation and for correlating the genetic and molecular maps of the *ARG7* locus**. *EMBO J* **8**, 2803-2809

37. Kovar, J. L., Zhang, J., Funke, R. P., and Weeks, D. P. (2002) **Molecular analysis of the acetolactate synthase gene of *Chlamydomonas reinhardtii* and development of a genetically engineered gene as a dominant selectable marker for genetic transformation**. *Plant J* **29**, 109-117

38. Nelson, J. A., Savereide, P. B., and Lefebvre, P. A. (1994) **The *CRY1* gene in *Chlamydomonas reinhardtii*: structure and use as a dominant selectable marker for nuclear transformation**. *Mol Cell Biol* **14**, 4011-4019

39. Rosales-Mendoza, S., Paz-Maldonado, L. M., and Soria-Guerra, R. E. (2012) ***Chlamydomonas reinhardtii* as a viable platform for the production of recombinant proteins: current status and perspectives**. *Plant Cell Rep* **31**, 479-494

40. Leon-Banares, R., Gonzalez-Ballester, D., Galvan, A., and Fernandez, E. (2004) **Transgenic microalgae as green cell-factories**. *Trends Biotechnol* **22**, 45-52

41. Surzycki, R., Greenham, K., Kitayama, K., Dibal, F., Wagner, R., Rochaix, J. D., Ajam, T., and Surzycki, S. (2009) **Factors effecting expression of vaccines in microalgae**. *Biologicals* **37**, 133-138

42. Mayfield, S. P., Manuell, A. L., Chen, S., Wu, J., Tran, M., Siefker, D., Muto, M., and Marin-Navarro, J. (2007) ***Chlamydomonas reinhardtii* chloroplasts as protein factories**. *Curr Opin Biotechnol* **18**, 126-133

43. Rosales-Mendoza, S. (2013) **Future directions for the development of *Chlamydomonas*-based vaccines**. *Expert Rev Vaccines* **12**, 1011-1019

44. Silflow, C. D., and Lefebvre, P. A. (2001) **Assembly and motility of eukaryotic cilia and flagella: lessons from *Chlamydomonas reinhardtii***. *Plant Physiol* **127**, 1500-1507

45. Pedersen, L. B., Veland, I. R., Schroder, J. M., and Christensen, S. T. (2008) **Assembly of primary cilia**. *Dev Dyn* **237**, 1993-2006

46. Snell, W. J., Pan, J., and Wang, Q. (2004) **Cilia and flagella revealed: from flagellar assembly in *Chlamydomonas* to human obesity disorders**. *Cell* **117**, 693-697

47. Matsuo, T., and Ishiura, M. (2011) ***Chlamydomonas reinhardtii* as a new model system for studying the molecular basis of the circadian clock**. *FEBS Lett* **585**, 1495-1502

48. Matsuo, T., and Ishiura, M. (2010) **New insights into the circadian clock in *Chlamydomonas***. *Int Rev Cell Mol Biol* **280**, 281-314

49. Schulze, T., Prager, K., Dathe, H., Kelm, J., Kiessling, P., and Mittag, M. (2010) **How the green alga *Chlamydomonas reinhardtii* keeps time**. *Protoplasma* **244**, 3-14

50. Grossman, A. R., Catalanotti, C., Yang, W., Dubini, A., Magneschi, L., Subramanian, V., Posewitz, M. C., and Seibert, M. (2011) **Multiple facets of anoxic metabolism and hydrogen production in the unicellular green alga *Chlamydomonas reinhardtii***. *New Phytol* **190**, 279-288

51. Radakovits, R., Jinkerson, R. E., Darzins, A., and Posewitz, M. C. (2010) **Genetic engineering of algae for enhanced biofuel production**. *Eukaryot Cell* **9**, 486-501

52. Merchant, S. S., Kropat, J., Liu, B., Shaw, J., and Warakanont, J. (2012) **TAG, you're it! *Chlamydomonas* as a reference organism for understanding algal triacylglycerol accumulation**. *Curr Opin Biotechnol* **23**, 352-363

53. Orr-Weaver, T. L., Szostak, J. W., and Rothstein, R. J. (1981) **Yeast transformation: a model system for the study of recombination**. *Proc Natl Acad Sci U S A* **78**, 6354-6358

54. Sawitzke, J. A., Thomason, L. C., Costantino, N., Bubunenko, M., Datta, S., and Court, D. L. (2007) **Recombineering: *in vivo* genetic engineering in *E. coli, S. enterica*, and beyond**. *Methods Enzymol* **421**, 171-199

55. Rochaix, J. D., van Dillewijn, J., and Rahire, M. (1984) **Construction and characterization of autonomously replicating plasmids in the green unicellular alga *Chlamydomonas* reinhardii**. *Cell* **36**, 925-931

56. Sodeinde, O. A., and Kindle, K. L. (1993) **Homologous recombination in the nuclear genome of *Chlamydomonas reinhardtii***. *Proc Natl Acad Sci U S A* **90**, 9199-9203

57. Zorin, B., Hegemann, P., and Sizova, I. (2005) **Nuclear-gene targeting by using single-stranded DNA avoids illegitimate DNA integration in *Chlamydomonas reinhardtii***. *Eukaryot Cell* **4**, 1264-1272

58. Dent, R. M., Haglund, C. M., Chin, B. L., Kobayashi, M. C., and Niyogi, K. K. (2005) **Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of *Chlamydomonas reinhardtii***. *Plant Physiol* **137**, 545-556

59. Schroda, M. (2006) **RNA silencing in *Chlamydomonas*: mechanisms and tools**. *Curr Genet* **49**, 69-84

60. Cerutti, H., Ma, X., Msanne, J., and Repas, T. (2011) **RNA-mediated silencing in Algae: biological roles and tools for analysis of gene function**. *Eukaryot Cell* **10**, 1164-1172

61. Benatuil, L., Perez, J. M., Belk, J., and Hsieh, C. M. (2010) **An improved yeast transformation method for the generation of very large human antibody libraries**. *Protein Eng Des Sel* **23**, 155-159

62. Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) **High efficiency transformation of *E. coli* by high voltage electroporation**. *Nucleic Acids Res* **16**, 6127-6145

63. Ngoi, S. M., Chien, A. C., and Lee, C. G. (2004) **Exploiting internal ribosome entry sites in gene therapy vector design**. *Curr Gene Ther* **4**, 15-31

64. Douin, V., Bornes, S., Creancier, L., Rochaix, P., Favre, G., Prats, A. C., and Couderc, B. (2004) **Use and comparison of different internal ribosomal entry sites (IRES) in tricistronic retroviral vectors**. *BMC Biotechnol* **4**, 16

65. Licursi, M., Christian, S. L., Pongnopparat, T., and Hirasawa, K. (2011) ***In vitro* and *in vivo* comparison of viral and cellular internal ribosome entry sites for bicistronic vector expression**. *Gene Ther* **18**, 631-636

66. Donnelly, M. L., Hughes, L. E., Luke, G., Mendoza, H., ten Dam, E., Gani, D., and Ryan, M. D. (2001) **The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences**. *J Gen Virol* **82**, 1027-1041

67. Ryan, M. D., King, A. M., and Thomas, G. P. (1991) **Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence**. *J Gen Virol* **72 ( Pt 11)**, 2727-2732

68. Ryan, M. D., and Drew, J. (1994) **Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein**. *EMBO J* **13**, 928-933

69. Donnelly, M. L., Luke, G., Mehrotra, A., Li, X., Hughes, L. E., Gani, D., and Ryan, M. D. (2001) **Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'**. *J Gen Virol* **82**, 1013-1025

70. Doronina, V. A., Wu, C., de Felipe, P., Sachs, M. S., Ryan, M. D., and Brown, J. D. (2008) **Site-specific release of nascent chains from ribosomes at a sense codon**. *Mol Cell Biol* **28**, 4227-4239

71. Donnelly, M. L., Gani, D., Flint, M., Monaghan, S., and Ryan, M. D. (1997) **The cleavage activities of aphthovirus and cardiovirus 2A proteins**. *J Gen Virol* **78 ( Pt 1)**, 13-21

72. de Felipe, P., Luke, G. A., Hughes, L. E., Gani, D., Halpin, C., and Ryan, M. D. (2006) ***E unum pluribus*: multiple proteins from a self-processing polyprotein**. *Trends Biotechnol* **24**, 68-75

73. Luke, G., Escuin, H., De Felipe, P., and Ryan, M. (2010) **2A to the fore - research, technology and applications**. *Biotechnol Genet Eng Rev* **26**, 223-260

74. Szymczak, A. L., Workman, C. J., Wang, Y., Vignali, K. M., Dilioglou, S., Vanin, E. F., and Vignali, D. A. (2004) **Correction of multi-gene deficiency *in vivo* using a single 'self-cleaving' 2A peptide-based retroviral vector**. *Nat Biotechnol* **22**, 589-594

75. Fang, J., Qian, J. J., Yi, S., Harding, T. C., Tu, G. H., VanRoey, M., and Jooss, K. (2005) **Stable antibody expression at therapeutic levels using the 2A peptide**. *Nat Biotechnol* **23**, 584-590

76. Lacroix, B., Li, J., Tzfira, T., and Citovsky, V. (2006) **Will you let me use your nucleus? How *Agrobacterium* gets its T-DNA expressed in the host plant cell**. *Can J Physiol Pharmacol* **84**, 333-345

77. Vasquez, K. M., Marburger, K., Intody, Z., and Wilson, J. H. (2001) **Manipulating the mammalian genome by homologous recombination**. *Proc Natl Acad Sci U S A* **98**, 8403-8410

78. Jiang, D., Zhu, W., Wang, Y., Sun, C., Zhang, K. Q., and Yang, J. (2013) **Molecular tools for functional genomics in filamentous fungi: Recent advances and new strategies**. *Biotechnol Adv* **8**, 1562-74

79. Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J. D., Bussey, H., Chu, A. M., Connelly, C., Davis, K., Dietrich, F., Dow, S. W., El Bakkoury, M., Foury, F., Friend, S. H., Gentalen, E., Giaever, G., Hegemann, J. H., Jones, T., Laub, M., Liao, H., Liebundguth, N., Lockhart, D. J., Lucau-Danila, A., Lussier, M., M'Rabet, N., Menard, P., Mittmann, M., Pai, C., Rebischung, C., Revuelta, J. L., Riles, L., Roberts, C. J., Ross-MacDonald, P., Scherens, B., Snyder, M., Sookhai-Mahadeo, S., Storms, R. K., Veronneau, S., Voet, M., Volckaert, G., Ward, T. R., Wysocki, R., Yen, G. S., Yu, K., Zimmermann, K., Philippsen, P., Johnston, M., and Davis, R. W. (1999) **Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis**. *Science* **285**, 901-906

80. Jasin, M. (1996) **Genetic manipulation of genomes with rare-cutting endonucleases**. *Trends Genet* **12**, 224-228

81. Lambowitz, A. M., and Belfort, M. (1993) **Introns as mobile genetic elements**. *Annu Rev Biochem* **62**, 587-622

82. Porteus, M. H., and Baltimore, D. (2003) **Chimeric nucleases stimulate gene targeting in human cells**. *Science* **300**, 763

83. Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M. H., and Chandrasegaran, S. (2005) **Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells**. *Nucleic Acids Res* **33**, 5978-5990

84. Bogdanove, A. J., and Voytas, D. F. (2011) **TAL effectors: customizable proteins for DNA targeting**. *Science* **333**, 1843-1846

85. Gaj, T., Gersbach, C. A., and Barbas, C. F., 3rd. (2013) **ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering**. *Trends Biotechnol* **31**, 397-405

86. Fauser, F., Roth, N., Pacher, M., Ilg, G., Sanchez-Fernandez, R., Biesgen, C., and Puchta, H. (2012) **In planta gene targeting**. *Proc Natl Acad Sci U S A* **109**, 7535-7540

87. Labrou, N. E. (2010) **Random mutagenesis methods for *in vitro* directed enzyme evolution**. *Curr Protein Pept Sci* **11**, 91-100

88. Casadaban, M. J., and Cohen, S. N. (1979) **Lactose genes fused to exogenous promoters in one step using a Mu-lac bacteriophage: *in vivo* probe for transcriptional control sequences**. *Proc Natl Acad Sci U S A* **76**, 4530-4533

89. Leoni, C., Volpicella, M., De Leo, F., Gallerani, R., and Ceci, L. R. (2011) **Genome walking in eukaryotes**. *FEBS J* **278**, 3953-3977

90. Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C. C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R.,

Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D. E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W. L., Berry, C. C., and Ecker, J. R. (2003) **Genome-wide insertional mutagenesis of *Arabidopsis thaliana***. *Science* **301**, 653-657

91. Jeon, J. S., Lee, S., Jung, K. H., Jun, S. H., Jeong, D. H., Lee, J., Kim, C., Jang, S., Yang, K., Nam, J., An, K., Han, M. J., Sung, R. J., Choi, H. S., Yu, J. H., Choi, J. H., Cho, S. Y., Cha, S. S., Kim, S. I., and An, G. (2000) **T-DNA insertional mutagenesis for functional genomics in rice**. *Plant J* **22**, 561-570

92. May, B. P., Liu, H., Vollbrecht, E., Senior, L., Rabinowicz, P. D., Roh, D., Pan, X., Stein, L., Freeling, M., Alexander, D., and Martienssen, R. (2003) ***Maize*-targeted mutagenesis: A knockout resource for *maize***. *Proc Natl Acad Sci U S A* **100**, 11541-11546

93. Xu, T., Bharucha, N., and Kumar, A. (2011) **Genome-wide transposon mutagenesis in *Saccharomyces cerevisiae* and *Candida albicans***. *Methods Mol Biol* **765**, 207-224

94. Brown, J. S., Aufauvre-Brown, A., and Holden, D. W. (1998) **Insertional mutagenesis of *Aspergillus fumigatus***. *Mol Gen Genet* **259**, 327-335

95. Meng, Y., Patel, G., Heist, M., Betts, M. F., Tucker, S. L., Galadima, N., Donofrio, N. M., Brown, D., Mitchell, T. K., Li, L., Xu, J. R., Orbach, M., Thon, M., Dean, R. A., and Farman, M. L. (2007) **A systematic analysis of T-DNA insertion events in *Magnaporthe oryzae***. *Fungal Genet Biol* **44**, 1050-1064

96. Broussolle, V., Pandiani, F., Haddad, N., Michaud, C., Carlin, F., Nguyen-the, C., and Brillard, J. (2010) **Insertional mutagenesis reveals genes involved in**

***Bacillus cereus ATCC 14579*** **growth at low temperature**. *FEMS Microbiol Lett* **306**, 177-183

97. Kool, J., and Berns, A. (2009) **High-throughput insertional mutagenesis screens in mice to identify oncogenic networks**. *Nat Rev Cancer* **9**, 389-399

98. Golling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, E., Chen, W., Burgess, S., Haldi, M., Artzt, K., Farrington, S., Lin, S. Y., Nissen, R. M., and Hopkins, N. (2002) **Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development**. *Nat Genet* **31**, 135-140

99. Kaiser, K., and Goodwin, S. F. (1990) **"Site-selected" transposon mutagenesis of *Drosophila***. *Proc Natl Acad Sci U S A* **87**, 1686-1690

100. Galvan, A., Gonzalez-Ballester, D., and Fernandez, E. (2007) **Insertional mutagenesis as a tool to study genes/functions in *Chlamydomonas***. *Adv Exp Med Biol* **616**, 77-89

101. Stanford, W. L., Cohn, J. B., and Cordes, S. P. (2001) **Gene-trap mutagenesis: past, present and beyond**. *Nat Rev Genet* **2**, 756-768

102. Cecconi, F., and Meyer, B. I. (2000) **Gene trap: a way to identify novel genes and unravel their biological function**. *FEBS Lett* **480**, 63-71

103. Meslet-Cladiere, L., and Vallon, O. (2012) **A new method to identify flanking sequence tags in *Chlamydomonas* using 3'-RACE**. *Plant Methods* **8**, 21

104. Wang, Z., Ye, S., Li, J., Zheng, B., Bao, M., and Ning, G. (2011) **Fusion primer and nested integrated PCR (FPNI-PCR): a new high-efficiency strategy for rapid chromosome walking or flanking sequence cloning**. *BMC Biotechnol* **11**, 109

105. Tonooka, Y., and Fujishima, M. (2009) **Comparison and critical evaluation of PCR-mediated methods to walk along the sequence of genomic DNA**. *Appl Microbiol Biotechnol* **85**, 37-43

106. Moroney, J. V., Husic, H. D., Tolbert, N. E., Kitayama, M., Manuel, L. J., and Togasaki, R. K. (1989) **Isolation and characterization of a mutant of *Chlamydomonas reinhardtii* deficient in the $CO_2$ concentrating mechanism**. *Plant Physiol* **89**, 897-903

107. Grossman, A. R. (2005) **Paths toward algal genomics**. *Plant Physiol* **137**, 410-427

108. Wu-Scharf, D., Jeong, B., Zhang, C., and Cerutti, H. (2000) **Transgene and transposon silencing in *Chlamydomonas reinhardtii* by a DEAH-box RNA helicase**. *Science* **290**, 1159-1162

109. Jeong Br, B. R., Wu-Scharf, D., Zhang, C., and Cerutti, H. (2002) **Suppressors of transcriptional transgenic silencing in *Chlamydomonas* are sensitive to DNA-damaging agents and reactivate transposable elements**. *Proc Natl Acad Sci U S A* **99**, 1076-1081

110. Fuhrmann, M., Oertel, W., and Hegemann, P. (1999) **A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii***. *Plant J* **19**, 353-361

111. Rasala, B. A., Lee, P. A., Shen, Z., Briggs, S. P., Mendez, M., and Mayfield, S. P. (2012) **Robust expression and secretion of Xylanase1 in *Chlamydomonas reinhardtii* by fusion to a selection gene and processing with the FMDV 2A peptide**. *PLoS One* **7**, e43349

112. Rasala, B. A., Barrera, D. J., Ng, J., Plucinak, T. M., Rosenberg, J. N., Weeks, D. P., Oyler, G. A., Peterson, T. C., Haerizadeh, F., and Mayfield, S. P. (2013) **Expanding the spectral palette of fluorescent proteins for the green microalga *Chlamydomonas reinhardtii***. *Plant J* **74**, 545-556

113. Thomas, B. J., and Rothstein, R. (1989) **The genetic control of direct-repeat recombination in *Saccharomyces*: the effect of *rad52* and *rad1* on mitotic recombination at *GAL10*, a transcriptionally regulated gene**. *Genetics* **123**, 725-738

114. Ma, C., and Mitra, A. (2002) **Intrinsic direct repeats generate consistent post-transcriptional gene silencing in tobacco**. *Plant J* **31**, 37-49

115. Victoria Lumbreras, D. R. S. a. S. P. (1998) **Efficient foreign gene expression in *Chlamydomonas reinhardtii* mediated by an endogenous intron**. *Plant J* **14**, 441–447

116. Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., and Tsien, R. Y. (2004) **Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma sp.* red fluorescent protein**. *Nat Biotechnol* **22**, 1567-1572

117. Tsien, R. Y. (1998) **The green fluorescent protein**. *Annu Rev Biochem* **67**, 509-544

118. Wang, Y., Sun, Z., Horken, K. M., Im, C.-S., Xiang, Y., Grossman, A. R., and Weeks, D. P. (2005) **Analyses of CIA5, the master regulator of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*, and its control of gene expression**. *Canadian Journal of Botany* **83**, 765-779

119. Xiang, Y., Zhang, J., and Weeks, D. P. (2001) **The *Cia5* gene controls formation of the carbon concentrating mechanism in *Chlamydomonas reinhardtii***. *Proc Natl Acad Sci U S A* **98**, 5341-5346

120. Crespo, J. L., and Hall, M. N. (2002) **Elucidating TOR signaling and rapamycin action: lessons from *Saccharomyces cerevisiae***. *Microbiol Mol Biol Rev* **66**, 579-591

121. Crespo, J. L., Diaz-Troya, S., and Florencio, F. J. (2005) **Inhibition of target of rapamycin signaling by rapamycin in the unicellular green alga *Chlamydomonas reinhardtii***. *Plant Physiol* **139**, 1736-1749

122. Nigam, S. K., Jin, Y. J., Jin, M. J., Bush, K. T., Bierer, B. E., and Burakoff, S. J. (1993) **Localization of the FK506-binding protein, FKBP 13, to the lumen of the endoplasmic reticulum**. *Biochem J* **294 ( Pt 2)**, 511-515

123. Fischer, N., and Rochaix, J. D. (2001) **The flanking regions of *PsaD* drive efficient gene expression in the nucleus of the green alga *Chlamydomonas reinhardtii***. *Mol Genet Genomics* **265**, 888-894

124. Ruecker, O., Zillner, K., Groebner-Ferreira, R., and Heitzer, M. (2008) ***Gaussia*-luciferase as a sensitive reporter gene for monitoring promoter activity in the nucleus of the green alga *Chlamydomonas reinhardtii***. *Mol Genet Genomics* **280**, 153-162

125. Shao, N., and Bock, R. (2008) **A codon-optimized luciferase from *Gaussia princeps* facilitates the *in vivo* monitoring of gene expression in the model alga *Chlamydomonas reinhardtii***. *Curr Genet* **53**, 381-388

126. Szczepek, M., Brondani, V., Buchel, J., Serrano, L., Segal, D. J., and Cathomen, T. (2007) **Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases**. *Nat Biotechnol* **25**, 786-793

127. Schroda, M., Blocker, D., and Beck, C. F. (2000) **The *HSP70A* promoter as a tool for the improved expression of transgenes in *Chlamydomonas***. *Plant J* **21**, 121-131

128. Ferrante, P., Catalanotti, C., Bonente, G., and Giuliano, G. (2008) **An optimized, chemically regulated gene expression system for *Chlamydomonas***. *PLoS One* **3**, e3200

129. Ohresser, M., Matagne, R. F., and Loppes, R. (1997) **Expression of the arylsulphatase reporter gene under the control of the *nit1* promoter in *Chlamydomonas reinhardtii***. *Curr Genet* **31**, 264-271

130. Kucho, K., Ohyama, K., and Fukuzawa, H. (1999) **$CO_2$-responsive transcriptional regulation of *CAH1* encoding carbonic anhydrase is mediated by enhancer and silencer regions in *Chlamydomonas reinhardtii***. *Plant Physiol* **121**, 1329-1338

131. de Hostos, E. L., Togasaki, R. K., and Grossman, A. (1988) **Purification and biosynthesis of a derepressible periplasmic arylsulfatase from *Chlamydomonas reinhardtii***. *J Cell Biol* **106**, 29-37

132. Neupert, J., Karcher, D., and Bock, R. (2009) **Generation of *Chlamydomonas* strains that efficiently express nuclear transgenes**. *Plant J* **57**, 1140-1150

133. Paulmurugan, R., and Gambhir, S. S. (2005) **Novel fusion protein approach for efficient high-throughput screening of small molecule-mediating protein-protein interactions in cells and living animals**. *Cancer Res* **65**, 7413-7420

134. Gorman, D. S., and Levine, R. P. (1965) **Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardtii***. *Proc Natl Acad Sci U S A* **54**, 1665-1669

135. Zorin, B., Lu, Y., Sizova, I., and Hegemann, P. (2009) **Nuclear gene targeting in *Chlamydomonas* as exemplified by disruption of the *PHOT* gene**. *Gene* **432**, 91-96

136. Sizova, I., Greiner, A., Awasthi, M., Kateriya, S., and Hegemann, P. (2013) **Nuclear gene targeting in *Chlamydomonas* using engineered zinc-finger nucleases**. *Plant J* **73**, 873-882

137. Ikehata, H., and Ono, T. (2011) **The mechanisms of UV mutagenesis**. *J Radiat Res* **52**, 115-125

138. Gonzalez-Ballester, D., Pootakham, W., Mus, F., Yang, W., Catalanotti, C., Magneschi, L., de Montaigu, A., Higuera, J. J., Prior, M., Galvan, A., Fernandez, E., and Grossman, A. R. (2011) **Reverse genetics in *Chlamydomonas*: a platform for isolating insertional mutants**. *Plant Methods* **7**, 24

139. O'Malley, R. C., Alonso, J. M., Kim, C. J., Leisse, T. J., and Ecker, J. R. (2007) **An adapter ligation-mediated PCR method for high-throughput mapping of T-DNA inserts in the *Arabidopsis* genome**. *Nat Protoc* **2**, 2910-2917

140. Liu, Y. G., and Whittier, R. F. (1995) **Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking**. *Genomics* **25**, 674-681

141. Luo, P., Su, T., Hu, C., and Ren, C. (2011) **A novel and simple PCR walking method for rapid acquisition of long DNA sequence flanking a known site in microbial genome**. *Mol Biotechnol* **47**, 220-228

142. Zhu, C., Wu, J., and He, C. (2010) **Induction of chromosomal inversion by integration of T-DNA in the rice genome**. *J Genet Genomics* **37**, 189-196

143. Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) **Genetic applications of an inverse polymerase chain reaction**. *Genetics* **120**, 621-623

144. Tam, L. W., and Lefebvre, P. A. (1993) **Cloning of flagellar genes in *Chlamydomonas reinhardtii* by DNA insertional mutagenesis**. *Genetics* **135**, 375-384

145. Pazour, G. J., and Witman, G. B. (2000) **Forward and reverse genetic analysis of microtubule motors in *Chlamydomonas***. *Methods* **22**, 285-298

146. Tuncay, H., Findinier, J., Duchene, T., Cogez, V., Cousin, C., Peltier, G., Ball, S. G., and Dauvillee, D. (2013) **A forward genetic approach in *Chlamydomonas reinhardtii* as a strategy for exploring starch catabolism**. *PLoS One* **8**, e74763

147. Tran, P. T., Sharifi, M. N., Poddar, S., Dent, R. M., and Niyogi, K. K. (2012) **Intragenic enhancers and suppressors of phytoene desaturase mutations in *Chlamydomonas reinhardtii***. *PLoS One* **7**, e42196

148. Davies, J. P., Yildiz, F., and Grossman, A. R. (1994) **Mutants of *Chlamydomonas* with aberrant responses to sulfur deprivation**. *Plant Cell* **6**, 53-63

149. Tanaka, A., Ito, H., Tanaka, R., Tanaka, N. K., Yoshida, K., and Okada, K. (1998) **Chlorophyll a oxygenase (*CAO*) is involved in chlorophyll b formation from chlorophyll a**. *Proc Natl Acad Sci U S A* **95**, 12719-12723

150. Li, J., Vaidya, M., White, C., Vainstein, A., Citovsky, V., and Tzfira, T. (2005) **Involvement of KU80 in T-DNA integration in plant cells**. *Proc Natl Acad Sci U S A* **102**, 19231-19236

151. Gelvin, S. B., and Kim, S. I. (2007) **Effect of chromatin upon *Agrobacterium* T-DNA integration and transgene expression**. *Biochim Biophys Acta* **1769**, 410-421

152. Lieber, M. R. (2008) **The mechanism of human nonhomologous DNA end joining**. *J Biol Chem* **283**, 1-5

153. Daley, J. M., Palmbos, P. L., Wu, D., and Wilson, T. E. (2005) **Nonhomologous end joining in yeast**. *Annu Rev Genet* **39**, 431-451

154. Bahmed, K., Nitiss, K. C., and Nitiss, J. L. (2010) **Yeast *Tdp1* regulates the fidelity of nonhomologous end joining**. *Proc Natl Acad Sci U S A* **107**, 4057-4062

155. Flowers, J. L., and Vaillancourt, L. J. (2005) **Parameters affecting the efficiency of *Agrobacterium tumefaciens*-mediated transformation of Colletotrichum graminicola**. *Curr Genet* **48**, 380-388

156. Rios, G., Lossow, A., Hertel, B., Breuer, F., Schaefer, S., Broich, M., Kleinow, T., Jasik, J., Winter, J., Ferrando, A., Farras, R., Panicot, M., Henriques, R., Mariaux, J. B., Oberschall, A., Molnar, G., Berendzen, K., Shukla, V., Lafos, M., Koncz, Z., Redei, G. P., Schell, J., and Koncz, C. (2002) **Rapid identification of *Arabidopsis***

**insertion mutants by non-radioactive detection of T-DNA tagged genes**. *Plant J* **32**, 243-253

157. Folger, K. R., Wong, E. A., Wahl, G., and Capecchi, M. R. (1982) **Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules**. *Mol Cell Biol* **2**, 1372-1387

158. Suemizu, H., Muguruma, K., Maruyama, C., Tomisawa, M., Kimura, M., Hioki, K., Shimozawa, N., Ohnishi, Y., Tamaoki, N., and Nomura, T. (2002) **Transgene stability and features of *rasH2* mice as an animal model for short-term carcinogenicity testing**. *Mol Carcinog* **34**, 1-9

159. Wu, B., Sun, Y. H., Wang, Y. W., Wang, Y. P., and Zhu, Z. Y. (2005) **Characterization of transgene integration pattern in F4 hGH-transgenic common carp (*Cyprinus carpio L.*)**. *Cell Res* **15**, 447-454

160. Qu, B., Eu, Y. J., Jeong, W. J., and Kim, D. P. (2012) **Droplet electroporation in microfluidics for efficient cell transformation with or without cell wall removal**. *Lab Chip* **12**, 4483-4488

161. Bryksin, A. V., and Matsumura, I. (2010) **Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids**. *Biotechniques* **48**, 463-465