Molecular techniques to interrogate and edit the
*Chlamydomonas* nuclear genome

Robert E. Jinkerson and Martin C. Jonikas*
Department of Plant Biology, Carnegie Institution for Science, 260 Panama Street, Stanford, CA 94305, USA

Received 17 November 2014; revised 13 February 2015; accepted 16 February 2015.
*For correspondence (e-mail mjonikas@carnegiescience.edu).

**SUMMARY**

The success of the green alga *Chlamydomonas reinhardtii* as a model organism is to a large extent due to the wide range of molecular techniques that are available for its characterization. Here, we review some of the techniques currently used to modify and interrogate the *C. reinhardtii* nuclear genome and explore several technologies under development. Nuclear mutants can be generated with ultraviolet (UV) light and chemical mutagens, or by insertional mutagenesis. Nuclear transformation methods include biolistic delivery, agitation with glass beads, and electroporation. Transforming DNA integrates into the genome at random sites, and multiple strategies exist for mapping insertion sites. A limited number of studies have demonstrated targeted modification of the nuclear genome by approaches such as zinc-finger nucleases and homologous recombination. RNA interference is widely used to knock down expression levels of nuclear genes. A wide assortment of transgenes has been successfully expressed in the *Chlamydomonas* nuclear genome, including transformation markers, fluorescent proteins, reporter genes, epitope tagged proteins, and even therapeutic proteins. Optimized expression constructs and strains help transgene expression. Emerging technologies such as the CRISPR/Cas9 system, high-throughput mutant identification, and a whole-genome knockout library are being developed for this organism. We discuss how these advances will propel future investigations.

Keywords: *Chlamydomonas reinhardtii*, molecular tools, genome editing, transformation, nuclear genome, reporters, markers, genetics.

**INTRODUCTION**

The green alga *Chlamydomonas reinhardtii* (hereafter, *Chlamydomonas*) is an excellent model system for the study of a variety of topics including photosynthesis, phototaxis, flagellar/ciliary biology, the algal carbon concentrating mechanism, and algal lipid biology. *Chlamydomonas* shares common ancestry and many biological features with land plants; thus, discoveries made in this alga are often directly applicable to plants. There is also significant interest for the use of this organism in biotechnology applications, such as the production of therapeutic proteins and as a model for optimizing algal metabolic pathways for biofuel production (Hannon et al., 2010; Jinkerson et al., 2011; Rasala and Mayfield, 2011; Scranton et al., 2015; Terashima et al., 2015).

Several key features of this organism make it a useful platform for such a variety of topics. *Chlamydomonas* can grow photoautotrophically in the light, but is also able to grow completely heterotrophically in the dark if supplied with acetate as a carbon and energy source. Owing to this metabolic flexibility, *Chlamydomonas* enables the facile study of mutations in genes essential for photosynthesis, which would be lethal in many other systems. The unicellular nature of *Chlamydomonas* simplifies the study of responses to environmental perturbations, by eliminating the complexity associated with the variety of different tissues found in higher plants. Liquid cell cultures can be grown easily in the laboratory, and their cell cycles can be synchronized by growth in a light/dark diurnal cycle. Vegetative cells are haploid, so recessive mutations immediately show a phenotype. *Chlamydomonas* has a well-characterized sexual cycle, and gametes can be crossed to yield diploid zygotes that allow for tetrad analysis.
The haploid nuclear genome of *Chlamydomonas* has a size of approximately 111.1 Mb (Blaby et al., 2014). The *Chlamydomonas* genome sequence was first available in 2002 (Grossman et al., 2003), and the completed draft was released in 2007 (Merchant et al., 2007). Through classical linkage analysis, it was established that *Chlamydomonas* has 17 linkage groups that display Mendelian segregation (Dutcher et al., 1991). The *Chlamydomonas* genome is relatively G+C rich, with an average G+C content of 64% throughout the nuclear genome and 68% in coding sequences (Merchant et al., 2007). The nuclear genome is predicted to have approximately 17,741 gene models (Blaby et al., 2014). In addition, the approximately 203 kb chloroplast genome contains 99 genes (Maul et al., 2002) and the approximately 16 kb mitochondrial genome contains eight genes (Remacle et al., 2006). Both of these organellar genomes can be genetically transformed through highly efficient homologous recombination techniques—a feature, that along with the availability of numerous nuclear mutants and nuclear genome transformation, greatly facilitates studies of nuclear/organellar interactions.

This review aims to provide an overview of molecular techniques that are available to alter and interrogate the nuclear genome of *Chlamydomonas* (Figure 1). We review classical random mutagenesis, insertional mutagenesis, and methods of identifying the site of mutation. We then discuss RNAi and strategies for targeted mutagenesis. Finally, we discuss expression of transgenes in the nuclear genome. Where possible, we present exciting emerging techniques that may become powerful additions to the molecular toolbox of this important model system.

**GENERATING AND MAPPING MUTANTS IN THE NUCLEAR GENOME WITH MUTAGENS**

**Mutagens can be used to generate mutants**

Mutagens have been used to generate mutants with aberrations in photosynthesis and carbon fixation (Eversole, 1956; Levine, 1960a,b; Bennoun and Levine, 1967), motility (Randall et al., 1964; Warr et al., 1966), mating (Goodenough et al., 1976), nitrogen assimilation (Ebersold, 1956; Eversole, 1956), respiration (Wiseman et al., 1977), and S-phase (Lee and Jones, 1973; Hawks and Lee, 1976). One of the potential side effects of using these mutagens is that they may produce closely linked mutations due to preferential mutagenesis near the replication fork (Lee and Jones, 1976). UV mutagenesis on log phase cells shows no evidence of linked mutations from whole-genome sequencing of mutants and may be less effective on cells in stationary phase (Personal communication Susan Dutcher).

**Mutated nuclear loci can be identified by traditional genetic mapping**

Nuclear mutations are first identified by analysis of their inheritance pattern. Mutations transmitted by the nuclear genome follow Mendelian inheritance, while mutations found within the organellar genomes are uniparentally inherited (Sager, 1954; Grant et al., 1980; Mets and Geist, 1983). Once identified to the nuclear genome, crossing the mutants to strains with known markers can allow for mutations to be mapped to a linkage group and map units calculated for distances between markers (for a detailed example of mapping mutations, see Rymarquis et al., 2005). Over 266 genetic loci have been mapped to linkage groups in *Chlamydomonas* and a molecular map with over 500 markers that is linked to the genetic map has been created (Bowers et al., 2003; Kathir et al., 2003; Rymarquis et al., 2005). Markers include cleavable amplified polymorphic sequences (CAPS), insertions/deletions (InDel), restriction fragment length polymorphisms (RFLP), sequence tagged sites (STS), and single nucleotide polymorphisms (SNP). *Chlamydomonas* mapping can be done in a high-throughput manner by bulked segregant analysis and marker duplexing (Rymarquis et al., 2005). In bulked segregant analysis, multiple segregating progeny are combined for marker analysis. In duplexing, multiple markers are evaluated at once by combining PCR primer sets in one reaction. Both methods save time and costs during mapping. A kit is
even available from the *Chlamydomonas* Resource Center containing primers for 96 PCR-based markers. Once mapped to a locus, the mutation can be identified by positional cloning using the indexed *Chlamydomonas* Bacterial Artificial Chromosome (BAC) library (Nguyen et al., 2005). The library has approximately eightfold coverage of the genome with an average insert size of 75 kb. BAC clones covering the mapped locus can be transformed into mutants, and transformants can be screened for complementation.

*Chlamydomonas* genes can be mutated by transposons. *Chlamydomonas* contains several transposable elements, including class I retrotransposon TOC1 (Day et al., 1988) and REM1 (Pérez-Alegre et al., 2005) and class II elements *Gulliver* (Ferris, 1989), *Pioneer1* (Graham et al., 1995), *Tcr1* (Ferris et al., 1996), *Tcr3* (Wang et al., 1998),...
and Bill (Kim et al., 2006). Transposable elements that transpose during mitotic growth may randomly integrate throughout the genome. One study of spontaneous mutants found signatures of five different transposable elements at the AMT4 locus (encoding an NH3 channel) (Kim et al., 2006).

To identify genes disrupted by a transposon, a DNA fragment containing the transposon is isolated and sequenced. To determine if a transposition has occurred in a mutant of interest, a Southern blot is performed with transposon-specific probes to identify possible changes in the location or number of transposons relative to the wild-type. The mutant is then crossed to the wild-type to identify hybridizing bands that are linked to the phenotype. If such a band is found, the correctly sized fragments are cloned into a library and then hybridized to a transposon probe to identify the transposon-containing clone. This clone can be sequenced to identify the mutated gene. This method was used to isolate the NIT2 gene, a transcription factor essential for nitrate assimilation (Camargo et al., 2007), which was disrupted by a Gulliver element (Schnell and Lefebvre, 1993).

Mutated loci can be identified with next-generation sequencing

With the advent of next-generation sequencing, whole mutant genomes can be sequenced to identify causal mutations (Dutcher et al., 2012; Lin et al., 2013a,b; Tulin and Cross, 2014). Individual mutants or bulked segregants can be sequenced to identify alterations in the genome that led to the observed phenotype (see Alford et al., 2013 supplemental figure for visualizations of bulked segregant mapping). Mutant mapping data can be useful for narrowing down the genomic regions that should be searched for mutations, especially if there are many SNPs between the reference genome and the mutant strain. With this technique, a causative mutation can be identified quickly and accurately (Dutcher et al., 2012; Lin et al., 2013b).

Whole-genome sequencing of several laboratory strains of Chlamydomonas has revealed 30 000–150 000 changes relative to the reference genome (Dutcher et al., 2012; Lin et al., 2013a). Thus, even when a mutant is mapped to just a few hundred kilobases, multiple nonsynonymous changes in the region can still make identification of the causative mutation challenging. The Dutcher laboratory has built a SNP/short indel library for 17 strains of Chlamydomonas (Lin et al., 2013a). They identified 2 547 089 polymorphisms within these strains and built a website that allows the Chlamydomonas community to quickly analyze whole-genome sequencing data from new mutant strains. With this tool, polymorphisms found in the background strain can be quickly dismissed, which facilitates the identification of unique (and likely causative) polymorphisms. As a proof of principle, they mapped the pf27 (paralyzed flagella) mutant using a bulked segregant approach by sequencing a pool of 14 meiotic progeny (Alford et al., 2013).

DNA CAN BE INSERTED INTO THE CHLAMYDOMONAS NUCLEAR GENOME

Multiple transformation methods exist to integrate DNA into the nuclear genome

Various techniques have been established to transform DNA into the Chlamydomonas nuclear genome. For an extensive review of the history of transformation in Chlamydomonas, see the Chlamydomonas sourcebook (Harris, 2009).

Biolistic transformation. The first successful method to stably incorporate exogenous DNA into the nuclear genome was biolistic transformation. This method utilizes a biolistic particle delivery system, commonly called a ‘gene gun’, which delivers metal microparticles coated with DNA at a high velocity into a cell. Biolistic transformation of Chlamydomonas was developed in the late 1980’s and was first used to deliver native genes that complemented auxotrophies (Debuchy et al., 1989; Kindle et al., 1989; Mayfield and Kindle, 1990). However, this method typically produces fewer transformants than other methods, and requires specialized equipment.

Glass bead transformation. This method requires no specialized equipment and is relatively simple. Cells to be transformed are put in a tube with glass beads that are agitated by vortexing (Kindle, 1990). A limitation to this method is that the recipient strain must lack a cell wall; thus, an enzymatic treatment of the cells with autolysin is required to remove the cell wall prior to transformation, unless the strain is already cell-wall-deficient (Kindle, 1990; Nelson and Lefebvre, 1995b; Shimogawara et al., 1998; Yamano et al., 2013).

Electroporation. Electroporation is also an efficient technique for delivery of DNA into the Chlamydomonas nucleus (Brown et al., 1991; Shimogawara et al., 1998; Yamano et al., 2013). Electroporation is among the most efficient methods for generating transformants, yielding up to 100 times more mutants for a given quantity of exogenous DNA than the glass bead method (Shimogawara et al., 1998).

Other techniques. Several other techniques exist for transformation of the nuclear genome. These include silicon carbide whiskers (Dunahay, 1993), biotransformation by Agrobacterium tumefaciens (Kumar et al., 2004; Prath-eesh et al., 2014), and nanoparticles (Kim et al., 2014).
Transforming DNA integrates at random sites into the genome

Insertion of DNA into the *Chlamydomonas* genome is rarely clean. Insertions are often accompanied by deletions of genomic sequence, deletions of the cassette, concatenation of cassettes, insertion of short DNA sequences between the cassette and the genome, and genomic rearrangements.

While the precise mechanism of insertion of DNA into the genome is unknown, a simple model explains all of the observed data (Zhang et al., 2014). In this model (Figure 2), electroporated transforming DNA is fragmented by a sequence-specific endonucleolytic activity (likely, an endonuclease). The transforming DNA is then ligated into the genome at the site of a double-stranded genomic DNA break by the non-homologous end joining repair pathway. Sometimes, multiple pieces of transforming DNA are concatenated together during this ligation process, resulting in tandem insertions of transforming DNA sequence.

A recent high-throughput study of thousands of insertion sites revealed several key characteristics of cassette fragmentation during transformation (Zhang et al., 2014). This study used next-generation sequencing to analyze

![Figure 2](image-url)

**Figure 2.** Transforming DNA integrates into the *Chlamydomonas* nuclear genome at double-stranded breaks and may be subject to endonucleolytic cleavage and ligation events. The high-throughput mapping of thousands of insertional mutants has allowed for the generation of a model of DNA integration into the *Chlamydomonas* nuclear genome. This model pertains to transformation cassettes that do not include homology regions intended for homologous recombination.

(a) Before transformation, the cassette is added to the extracellular medium that may also contain extracellular DNA from lysed *Chlamydomonas* cells. The extracellular DNA may include nuclear, chloroplastic, or mitochondrial DNA. (b) During the transformation procedure, both the cassette and extracellular DNA are subjected to fragmentation by site-specific endonucleases. (c–h) (c) Once inside of the nucleus, transforming DNA may be ligated into double-stranded breaks (DSB) within the genome. In some cases, whole cassettes (d) are ligated into the DSB; however, truncated versions or fragments of the cassette can also be inserted (e). Multiple cassettes (f, g) can also be ligated into a single location. Fragments of the extracellular *Chlamydomonas* DNA may also be ligated together with the cassette at a site of a DSB (h). In this case, the genomic DNA adjacent to the cassette is not from the location of the DSB, but from some other location in the genome. For these insertions, genomic sequence past this inserted genomic DNA fragment needs to be sequenced to correctly identify the site of insertion. Abbreviations: NHEJ, non-homologous end joining. Adapted from Zhang et al. (2014), Copyright American Society of Plant Biologists, www.plantcell.org.

© 2015 The Authors
junctions between cassette fragments that were concatenated during the insertion process. These data allowed the analysis of thousands of fragmented cassette ends inserted into the genome. The data argued against an exonuclease activity being responsible for most of the cassette fragmentation. Instead, the data argued strongly for much of the cassette fragmentation being due to digestion by an endonuclease with a preference for the consensus motif CA/TG, as well as variants of this sequence.

Genomic insertion sites appear to be random (Zhang et al., 2014). The same high-throughput study analyzed thousands of insertion sites in the genome. There were essentially no hotspots observed, and essentially no coldspots could be seen at the resolution available. This observation is consistent with a model in which transforming DNA is ligated into a site of a double-stranded DNA break, which would be expected to occur randomly throughout the genome, with little sequence specificity. Notably, a bias in gene features was observed, with fewer insertions into exons and 5’UTRs of genes than into the rest of the genome. This bias is likely downstream of the DNA integration: there could be lower recovery of such mutants due to the increased chance that insertions in these features would disrupt essential genes. Overall, these findings suggest that insertions in nearly all non-essential genes can be obtained efficiently through insertional mutagenesis.

Other DNA sequences are sometimes inserted between the cassette and the true flanking genomic sequence (Figure 2). Such sequences include short fragments (typically 20–500 bp) of genomic DNA (Zhang et al., 2014), as well as a short fragment of chloroplast genome sequence (in one of our unpublished mutants). The simplest explanation for these observations is that such insertions are caused by co-transformation of the cells with DNA from lysed cells, which could then get ligated between the cassette and the flanking genomic DNA at the insertion site.

DNA rearrangements are sometimes observed. In some reported mutants (Aksoy et al., 2013; Zhang et al., 2014), inversions of genomic sequence have been observed. In the context of the above model, these inversions could be due to two double-stranded breaks occurring simultaneously on one chromosome before or during the transformation. The genomic DNA fragment between the double-stranded breaks could flip before re-ligation, and the transforming DNA would be ligated into the genome during repair at one of the two break sites.

Different methods of transformation appear to lead to different distributions of deletions associated with the insertion of foreign DNA. Glass bead transformation appears to lead to bigger deletions than electroporation. In one study that compared the two methods (González-Ballesta et al., 2011), five out of 11 mutants generated by glass bead transformation carried deletions of 4–35 kb at the insertion site. In contrast, 23 out of 32 mutants generated by electroporation had no associated deletion, and the remaining mutants had at most a 32 bp deletion.

The observed phenotype is not always associated with the mapped insertion. When photosynthesis-deficient mutants generated by glass bead transformation were analyzed by crossing, the phenotype co-segregated with the resistance marker in nine out of 17 mutants (52%) (Dent et al., 2005). This is likely due to the presence of additional mutations in these strains, which are not linked to the insertion cassette but are causing the photosynthesis-deficient phenotype.

A variety of transformation markers exists for *Chlamydomonas* nuclear transformation

Transformation markers that have been used for nuclear transformation in *Chlamydomonas* include auxotrophic markers and antibiotic resistance markers (Table 1). Genes that have been used as auxotrophic markers include the ARG7, encoding argininosuccinate lyase (Debuchy et al., 1989), NIT1, encoding nitrate reductase (Kindle et al., 1989; Blakenship and Kindle, 1992; Tam and Lefebvre, 1993), NIC7, encoding quinolinate synthetase (Ferris, 1995), and THI10, encoding hydroxyethylthiazole kinase (Ferris, 1995). For the use of any of these markers, the strain to be transformed must be auxotrophic for the marker. The use of an auxotrophic mutant can complicate experiments due to the need to add an external metabolite. For example, nitrogen deprivation experiments conducted in the arg7 background are complicated by the fact that the supplemented arginine could be used as a nitrogen source (Work et al., 2010). In these cases antibiotic resistance markers may be more suitable.

The first nuclear antibiotic resistance marker available for use in *Chlamydomonas* was the CRY1 gene, which confers resistance to etemina via a dominant mutant allele of the ribosomal protein S14 (Nelson et al., 1994). Since then, various other marker genes have been used, including acetolactate synthase (resistance to sulfofomuron methyl) (Kovar et al., 2002), ble (resistance to zeocin) (Stevens et al., 1996; Lumberas et al., 1998), aphVIII (resistance to paromomycin) (Sizova et al., 1996, 2001), aph7 (resistance to hygromycin) (Berthold et al., 2002; Ladygin and Boutanaev, 2002), and aadA (resistance to spectinomycin) (Cerutti et al., 1997b; Meslet-Cladière and Vallon, 2011). Markers to confer resistance to herbicides have also been used and include acetolactate synthase (resistance to sulfofomuron methyl) (Kovar et al., 2002), protoporphyrinogen oxidase (resistance to diphenyl ether herbicides) (Randeloph-Anderson et al., 1996; Jiang et al., 2014), phytoene desaturase (resistance to norflurazon) (Jiang et al., 2014), and glyphosate aminotransferase (resistance to glyphosate) (Jiang et al., 2014). Several markers have also been developed as shuttle vectors that allow for selection in both *Chlamydomonas* and *Escherichia coli* and include...
Co-transformation can be used to insert DNA into Chlamydomonas

Multiple transgenes can be transformed into Chlamydomonas at one time, and successful incorporation of a second unselected gene has been demonstrated (Kindle, 1990). Applying this strategy can reduce the number of selectable markers needed. Co-transformed DNA may contain two or more genes that are completely independent of each other (e.g. on separate plasmids), or are together on a contiguous piece of DNA (e.g. on one plasmid). Rates of co-transformation efficiencies of independent, unselected transgenes with the selected marker are typically seen in the 10% range (Kindle, 1990; Periz and Keller, 1997; Berthold et al., 2002). Linking the unselected transgene and marker on the same plasmid can increase the frequency of expression of the unselected gene (Heitzer and Zschoernig, 2007) and in one study improved the co-transformation efficiency sevenfold (Periz and Keller, 1997). Co-transformation cassette systems for Chlamydomonas have been designed to ease the creation of vectors that contain multiple genes (Heitzer and Zschoernig, 2007).

Multiple strategies exist for mapping insertion sites

Plasmid rescue uses a bacterial resistance marker to allow recovery of flanking DNA (Tam and Lefebvre, 1993). In this approach, part of the transformation cassette integrated into the Chlamydomonas genome contains a bacterial resistance marker and replication origin. Genomic DNA from the mutant of interest is digested with a restriction enzyme whose recognition site is not present in the transformation cassette. This produces a fragment containing the bacterial resistance marker and replication origin, flanked on both sides by genomic DNA sequence. The DNA is then ligated, allowing the circularization of this fragment into a functional bacterial plasmid, which contains some flanking genomic DNA sequence. The DNA is transformed into bacteria, resistant bacteria are selected, and the Chlamydomonas genomic DNA sequence on the plasmid can be sequenced to identify the flanking sequences of the Chlamydomonas insert. This approach is useful for cases in which the flanking sequences may be difficult to directly amplify by PCR. However, for plasmid rescue to be successful, the origin of replication and the bacterial resistance marker must be intact. For cases in which they are disrupted or non-functional in E. coli, other strategies must be used.

### Table 1 List of commonly used markers for transformation of Chlamydomonas

<table>
<thead>
<tr>
<th>Selection marker gene</th>
<th>Selection</th>
<th>Background needed</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resistance to antibiotics</td>
<td>ble</td>
<td>Zeocin</td>
<td>WT</td>
<td>From <em>Streptomyces rimosus</em></td>
</tr>
<tr>
<td></td>
<td><em>aph</em>7</td>
<td>Hygromycin</td>
<td>WT</td>
<td><em>hindustanus</em></td>
</tr>
<tr>
<td></td>
<td><em>aph</em>VIII</td>
<td>Paromomycin</td>
<td>WT</td>
<td>From <em>Streptomyces hygroscopicus</em></td>
</tr>
<tr>
<td></td>
<td><em>aadA</em></td>
<td>Spectinomycin resistance</td>
<td>WT</td>
<td>From <em>Streptomyces rimosus</em></td>
</tr>
<tr>
<td>Resistance to herbicides and other drugs</td>
<td>ALS</td>
<td>Sulfometuron methyl (SMM)</td>
<td>WT</td>
<td>Mutant form of <em>Chlamydomonas</em> acetalactate synthase (ALS)</td>
</tr>
<tr>
<td></td>
<td>CRY1-1</td>
<td>Emetine</td>
<td>WT</td>
<td>Mutant form of cytosolic ribosomal protein S14</td>
</tr>
<tr>
<td></td>
<td><em>rs</em>-3</td>
<td>Diphenyl ether herbicides (such as oxyfluorfen)</td>
<td>WT</td>
<td>Mutant form of protoporphyrinogen oxidase</td>
</tr>
<tr>
<td></td>
<td>GAT</td>
<td>Glyphosate</td>
<td>WT</td>
<td>Glyphosate amino transferase</td>
</tr>
<tr>
<td></td>
<td>PDS1</td>
<td>Norflurazon</td>
<td>WT</td>
<td>Mutant form of phytoene desaturase</td>
</tr>
<tr>
<td>Complementing auxotrophies</td>
<td>ARG7</td>
<td>Arg media</td>
<td><em>arg7</em></td>
<td>Mutant in argininosuccinate lyase</td>
</tr>
<tr>
<td></td>
<td>NIT1</td>
<td>Ammonium media + nitrate media</td>
<td><em>nit1</em></td>
<td>Mutant in nitrate reductase</td>
</tr>
<tr>
<td></td>
<td>NIC7</td>
<td>Nicotinamide media</td>
<td><em>nic7</em></td>
<td>Mutant in quinolinate synthetase</td>
</tr>
<tr>
<td></td>
<td>THI10</td>
<td>Thiamine media</td>
<td><em>thi10</em></td>
<td>Mutant in hydroxyethylthiazole kinase</td>
</tr>
</tbody>
</table>

**ARG7** (Auchincloss et al., 1999) or *aadA* (Meslet-Cladière and Vallon, 2011).
Nested PCR strategies use a series of PCRs with several primers in the cassette to provide specificity in amplification of flanking sequences. In nested PCR, several rounds of PCR are performed. In each round, one primer anneals in the cassette and one degenerate primer anneals to the flanking genomic sequence. Increasingly stringent specificity for a PCR product containing the transformation cassette is generated by using a different cassette-binding PCR primer in each round. The primers are chosen in a sequence that moves out towards the end of the cassette, so that only products that contain the cassette sequence are amplified in all rounds. Several variants of nested PCR have been used in *Chlamydomonas*; each uses a different strategy for choosing genome-side primers and PCR cycling parameters. These variants include genome walking (Posewitz *et al.*, 2004; Meuser *et al.*, 2011), thermal asymmetric interlaced (TAIL) PCR (Dent *et al.*, 2005), restriction enzyme site-directed amplification (RESDA) PCR (González-Ballestero *et al.*, 2005), SiteFinding PCR (Li *et al.*, 2012b), and hairpin-PCR (Plecenikova *et al.*, 2014). For these methods to be successful, the PCR primer binding sites in the cassette must be intact. Other problems can arise if cassette concatemerization occurs in the insertion site, especially if any of the cassettes are inserted in a tandem head-to-head orientation, where a single cassette-binding PCR primer can amplify a cassette only sequence (Dent *et al.*, 2005).

A method involving 3′ rapid amplification of cDNA ends (RACE) has been demonstrated for identifying flanking sequences in *Chlamydomonas* (Meslet-Cladière and Vallon, 2012). This method relies on a transformation cassette containing a drug-resistance gene without a transcriptional terminator. In mutants, this cassette generates hybrid transcripts containing the drug-resistance sequence, which continue into the flanking genomic DNA until they are terminated, presumably upon reaching an endogenous transcription termination sequence. These transcripts can be isolated by reverse transcription, followed by nested PCR, revealing the sequence of the genomic DNA flanking the insertion.

ChlaMmeSeq (*Chlamydomonas Mmel-based insertion site sequencing*) is a method compatible with deep sequencing, which allows simultaneous identification of insertion sites in tens of thousands of pooled mutants (Zhang *et al.*, 2014). In this approach, a type II restriction enzyme site (*Mmel*) is included at each end of the transforming DNA cassette. To extract flanking sequences, genomic DNA is digested with *Mmel*, which cleaves in the flanking genomic DNA, 20 bp away from its recognition site. This generates a fragment containing the cassette and 20 bp of flanking genomic DNA. Adapters are ligated onto the *Mmel*-generated end, and PCR amplification between the cassette and adapter yields a product of uniform size that can be sequenced either by Sanger or by Illumina sequencing. This approach allows the mapping of tens of thousands of insertion sites present in a pool of mutants in approximately 2 weeks.

In some mutants, fragmentation of the cassette can make it difficult or impossible to use any of these methods to map the insertion (Zhang *et al.*, 2014). Upon fragmentation, the binding sites for primers can be eliminated from the cassette at the insertion site. In such cases, whole-genome sequencing is one possible alternative method for mapping the insertion.

**PCR can be used to isolate mutants in genes of interest from collections of random insertion mutants**

Collections of tens of thousands of random insertion mutants can be screened by PCR to isolate mutants in genes of interest (Krysan *et al.*, 1996; González-Ballestero *et al.*, 2011). Random insertion mutants are generated and propagated in 96-well plates. The mutants from each plate are pooled together, and these pools are pooled into super-pools representing 10 plates each. PCR primers tiling the gene(s) of interest are obtained, and used in combination with a primer that binds in the insertion cassette, to identify super-pools that contain a putative insertion, as evidenced by a PCR product spanning the end of the insertion cassette and a part of the gene of interest. The precise coordinates of the mutant containing the putative insertion are then identified by repeating the PCR on the plate pools, rows and columns.

**We are developing a collection of indexed mutants**

Our laboratory is presently using ChlaMmeSeq to generate a collection of thousands of mutants, where one or more insertion sites are known for each mutant. The collection is propagated on agar and cryogenically preserved. This collection should dramatically accelerate the pace of progress by eliminating the need for mutant generation and mapping in many future projects.

**STRATEGIES EXIST FOR THE TARGETED MODIFICATION OF THE CHLAMYDOMONAS NUCLEAR GENOME**

RNA interference can be used to successfully modulate nuclear-gene expression in *Chlamydomonas*

*Chlamydomonas* contains an endogenous set of short RNAs (sRNAs), including small interfering RNAs (siRNAs) and microRNAs (miRNAs) (Molnár *et al.*, 2007; Zhao *et al.*, 2007), that are implicated in gene regulation, suppression of transposons, and transgene silencing (van Dijk *et al.*, 2006). The *Chlamydomonas* nuclear genome encodes the key enzymes for RNA silencing: Dicer and Argonaute (Cerutti and Casas-Mollano, 2006). These enzymes can process double-stranded RNA (dsRNA) into siRNAs, which serve as a guide in complexes that inhibit gene transcription or promote RNA degradation. This native ability for *Chlamydomonas* to process dsRNA has been used as a
reverse genetic tool for post-transcriptional gene silencing. The first approach taken in *Chlamydomonas* was to express the antisense sequence of a gene targeted for knockdown. Once transcribed, the antisense sequence can bind to the native miRNA, resulting in dsRNA. This method was used to target several genes (Schroda *et al.*, 1999; Fuhrmann *et al.*, 2001; Pfannenschmid *et al.*, 2003), but the knockdown phenotypes can be lost over time and the frequency of transformants where the target gene is knocked down can vary widely (0.3–50%) (Schroda, 2006).

RNA silencing constructs based on an inverted repeat design that contains cDNA sense and cDNA antisense sequences have been used for targeted knockdown of a variety of genes (Sineshchekov *et al.*, 2002; Rohr *et al.*, 2004; Soupene *et al.*, 2004). This design has been more successful than antisense only constructs (Schroda, 2006). Some inverted repeat constructs are made from genomic sense and cDNA antisense sequences. These use the 5’ end of the gene targeted for knockdown, containing the first few exons and introns, fused with antisense cDNA covering the exons included in the sense direction (Fuhrmann *et al.*, 2001). This inverted repeat structure can trigger RNA silencing with high efficiency and has been used to target a variety of genes (Fuhrmann *et al.*, 2001; Huang and Beck, 2003; Pan *et al.*, 2004).

Artificial miRNAs (amiRNAs) overcome some issues associated with long dsRNAs. Transgenes that express long dsRNA can themselves be targets of transcriptional silencing in *Chlamydomonas* (Rohr *et al.*, 2004). Also, because these long dsRNA give rise to many siRNAs, some of these may have off-target effects. An amiRNA is a modified version of an endogenous miRNA, which produces only a small single RNA that can be designed to avoid off-target effects. In *Chlamydomonas* the percentage of amiRNA transfectants that gave knockdown phenotypes ranged from 16 to 72% depending on the gene of interest (Molnár *et al.*, 2009; Zhao *et al.*, 2009).

Overall, RNAi silencing has been achieved for over 30 *Chlamydomonas* nuclear genes (Schroda, 2006). Despite limitations including incomplete target gene suppression, off-target effects, and loss of phenotype by silencing of the RNAi transgene, for now, RNAi remains the most reliable means for genetically targeting specific genes of interest.

**Targeting of transforming DNA by homologous recombination into the nuclear genome is possible, but rarely happens**

Homologous recombination (HR) has been demonstrated in *Chlamydomonas*, but its usefulness is presently limited by the low frequency at which it occurs. Homologous recombination between two co-transformed DNA fragments that overlap to form a gene has been demonstrated (Sodeinde and Kindle, 1993; Gumpel *et al.*, 1994). At a much lower frequency, mutated genes have been repaired by transformation of a fragment of a wild-type gene that recombines to repair and form a functional gene (Sodeinde and Kindle, 1993; Gumpel *et al.*, 1994; Zorin *et al.*, 2005, 2009). Additionally, wild-type genes have been mutated by HR with a mutated copy conferring a selection resistance (Randolph-Anderson *et al.*, 1998), and two specific wild-type genes have been disrupted by insertion of a resistance marker (Nelson and Lefebvre, 1995a; Zorin *et al.*, 2009). Most documented cases of HR in *Chlamydomonas* require direct selection strategies in which only cells that undergo HR are viable. Thus, HR is not practical for targeting most genes, which are not directly selectable.

Several strategies have been used to increase HR frequency in *Chlamydomonas*. These strategies include transforming single-stranded DNA, which is less likely to integrate via non-homologous end joining (NHEJ) (Zorin *et al.*, 2005); and transformation of synchronized cells at the start of mitosis (S/M phase), which is a time when HR is more likely to be active (Sizova *et al.*, 2013). HR rates can also be increased by inducing double-stranded breaks at the gene of interest, as has been done with zinc-finger nucleases (Sizova *et al.*, 2013). HR constructs can be designed with an antibiotic marker in frame with the target gene, relying on the target gene’s promoter to drive expression of the antibiotic marker (Zorin *et al.*, 2009). In this way, HR events with the target gene will lead to resistant transformants, while most NHEJ events will not result in resistant transformants, greatly increasing the ratio of HR to NHEJ events observed. This approach has been used successfully to disrupt the PHOT gene encoding the soluble sensory photoreceptor phototropin (Zorin *et al.*, 2009; Trippens *et al.*, 2012). This approach may not work well for knocking out membrane proteins, since the target gene needs to have an N-terminal section that is soluble so that the resulting fusion protein containing the selection marker is a soluble and active enzyme.

One yet-unexplored possible strategy to increase HR rates is to target Ku70/Ku80 for knockout. These proteins form a heterodimer that binds to double-stranded breaks and promotes NHEJ. In several organisms, it has been demonstrated that the frequency of homologous integration of exogenous DNA is greatly improved in mutants that lack Ku70 or Ku80 (Ninomiya *et al.*, 2004). The alga *Nannochloropsis*, which can perform high rates of HR (Kilian *et al.*, 2011), natively lacks the Ku80 gene (Jinkerson *et al.*, 2013). By analogy, it may be possible to isolate *Chlamydomonas* Ku70 or Ku80 mutants with increased homologous integration frequencies.

**Zinc-finger nucleases and transcription activator-like effectors (TALEs) can be used to edit and control the Chlamydomonas nuclear genome**

Zinc-finger nucleases (ZFNs) consist of a fusion of three to four zinc-finger DNA-binding modules to a FokI DNA
cleavage domain, forming an artificial restriction enzyme. A pair of these molecular scissors can be tailored to target specific DNA sequences, where dimerization of the FokI domain activates the nuclease, inducing site-specific DNA double-stranded breaks that enable genome editing (Kim et al., 1996; Townsend et al., 2009).

ZFN technology has been used in Chlamydomonas to modify a non-selectable gene of interest: the COP3 gene encoding light-activated ion channel channelrhodopsin-1 (Sizova et al., 2013). In order to enrich for ZFN mediated gene disruptions, Sizova et al. first generated a strain containing a non-functional aphVIII marker that was inactivated by the insertion of a short COP3 target sequence. ZFNs were designed to bind to this short COP3 gene sequence and then co-transformed with a partial-aphVIII DNA template into the non-functional aphVIII strain. If the ZFNs are functional, they induce a double-stranded break in the short COP3 gene sequence within the non-functional aphVIII marker, enabling the repair of aphVIII via HR with the partial-aphVIII DNA template. This strategy allowed for the optimization of the ZFNs in Chlamydomonas by using the number of paromomycin resistant colonies as an indicator of nuclease activity and recognition of the COP3 target sequence. In approximately 1% of the paromomycin resistant colonies, the ZFNs had also acted upon the endogenous COP3 gene, resulting in mutant alleles that had been repaired via the error prone NHEJ pathway. To promote targeted sequence alterations of the endogenous COP3 gene, a modified COP3 template that included a deletion or insertion and in-frame stop codons was used as template DNA for co-transformation experiments. This enabled recovery of mutants that had successfully undergone HR to replace the native COP3 gene sequence with the altered COP3 template (Sizova et al., 2013). For targeting other genes with ZFNs, a resource has been created, ZFNGenome, which has identified and provided quality scores for all potential ZFN target sites in the Chlamydomonas genome (Reyon et al., 2011). Over 330,000 target sites have been identified that cover approximately 93% of all gene models. Artificially designed TALEs have been used for gene-specific activation in Chlamydomonas. TALEs are secreted by the pathogenic bacteria Xanthomonas into host plant cells where they act as transcriptional activators (Boch et al., 2009). The main difference between zinc-finger domains and TAL-effector domains is their mechanism of DNA sequence recognition. Zinc-finger modules always recognize three base pairs, limiting the number of possible target sequences. In contrast, TALEs are composed of a variable number of four different DNA recognition domains, each of which specifically recognizes one of the four nucleotides. Therefore, 12-20 of these domains can be fused to generate a TALE that binds to any desired target DNA sequence.

Recently TALEs have been used to activate the expression of genes in Chlamydomonas (Gao et al., 2014, 2015). In one study, TALEs were designed to activate two genes, the periplasmic arylsulfatases ARS1 and ARS2, which are only expressed under sulfur deprived conditions. Robust ARS1 and ARS2 expression was observed under sulfur replete conditions and was closely correlated with TALE transcript abundance. Despite these genes being co-located next to each other and having similar promoter sequences, no off-target activation of non-targeted ARS genes was observed (Gao et al., 2014), which indicates that TALEs can be used in Chlamydomonas to activate specific genes in a targeted manner.

TALEs have also been adapted to include a nuclease domain (TALENs) to induce double-stranded breaks at precise locations in the nuclear genome in plants (Li et al., 2012a), but the use of this technology in Chlamydomonas has not been reported. Since TALE mediated gene activation has already been demonstrated (Gao et al., 2014), this gives credence to TALENs as possibly being an effective way to edit the nuclear genome in Chlamydomonas.

CRISPR/Cas9 could soon enable editing of the Chlamydomonas nuclear genome

The ability to precisely edit the genome of Chlamydomonas has been a desire of the research community since the first molecular manipulations in this organism took place. The recent discovery of CRISPR/Cas9 technology has the potential to provide this type of editing to the nuclear genome of Chlamydomonas. This technology is enabled by the recent discovery of key components of a bacterial/archaeal immune system that defends the host against phage attacks (Jinek et al., 2012; Wiedenheft et al., 2012). Differently from ZFNs and TALEs, the presently utilized Cas9-system uses a RNA-template, called single guide RNA (sgRNA), to mediate DNA-target recognition. The sgRNA forms a secondary structure that binds to the Cas9-nuclease resulting in the active holoenzyme. Target specificity of the guide RNA is mediated by the so-called ‘protoscaler’, an approximately 20 base pair sequence at the 5’-end of the sgRNA. This sequence stretch can be easily altered, targeting the Cas9-endonuclease to any desired genomic locus. Since it is unnecessary to create new enzymes for targeting different genes, multiple genes can be targeted in the same experiment using multiple different sgRNAs at the same time in animals cells (Shalem et al., 2014) and plants (Zhou et al., 2014).

Recently, the use of CRISPR/Cas9 to edit the nuclear genome of Chlamydomonas has been reported (Jiang et al., 2014). In this study, the two components of the CRISPR editing system (Cas9 and guide RNAs) were transiently expressed in the Chlamydomonas nuclear genome. Using a restriction digestion and PCR-based assay, the authors were able to detect CRISPR activity on the nuclear genome.
They also were able to select one mutant where the genomic FKB12 gene had a 2-bp deletion by apparent Cas9 activity resulting in resistance to rapamycin; however the efficiency was extremely low. Further development of a method for CRISPR/Cas9 genome editing in Chlamydomonas is needed to allow this technique to become more robust and widely applicable. Such advances should be possible given the variety of organisms in which this system has already been successfully applied (Shan et al., 2013; Nekrasov et al., 2014).

**EXPRESSION OF TRANSGENES IN THE CHLAMYDOMONAS NUCLEAR GENOME**

The expression of transgenes in the nuclear genome of *Chlamydomonas* has many basic and applied applications

Many heterologous transgenes have been expressed in *Chlamydomonas*, including many antibiotic resistance markers, reporter genes, and commercially relevant proteins, such as antigens for vaccines (Soria-Guerra et al., 2013).

Heterologous protein yields for transgenes expressed from the nuclear genome have been estimated to be up to 0.25% of total soluble protein (TSP) (Neupert et al., 2009; Rasala et al., 2012), which is 20 times lower than maximal yields for transgenes expressed in the chloroplast genome (Fuhrmann et al., 2004; Griesbeck et al., 2006; Manuell et al., 2007; Eichler-Stahlberg et al., 2009). Despite these low protein yields from nuclear-encoded proteins and the robust transformation and expression systems that exist for the *Chlamydomonas* chloroplast (Manuell et al., 2007) and mitochondrial genomes (Remacle et al., 2006), there are several advantages for nuclear genome-based transgene expression.

Expression from the nuclear genome is necessary for the industrial production of some heterologous proteins. Proteins expressed in the nuclear genome can include signal peptides (Table 2) that target the protein to the endoplasmic reticulum for secretion from *Chlamydomonas* for protein export

endoplasmic reticulum for secretion from the cell into the external media (Eichler-Stahlberg et al., 2009; Rasala et al., 2012; Lauersen et al., 2013a), whereas chloroplast produced proteins cannot be secreted and remain within the plastid. Nuclear-expressed proteins can also undergo specific post-translational modifications (PTM), like glycosylation, that are absent from proteins produced in the plastid (Mayfield et al., 2007). Many proteins of eukaryotic origin require some type of PTM, such as glycosylation of antibodies, and thus may be better suited for expression from the *Chlamydomonas* nuclear genome. Some of the PTMs documented to occur in *Chlamydomonas* include glycosylation (Mathieu-Rivet et al., 2014), acetylation (L'Hernault and Rosenbaum, 1985), S-thiolation (Michelet et al., 2008), and methylation (Decottignies et al., 1995). *Chlamydomonas* can even produce glycoproteins with mammalian-like sialylated N-glycans (Mamedov and Yusibov, 2011).

Furthermore, genes whose products are intended to modify or alter the nuclear genome, such as TALEs (Gao et al., 2014), must be integrated into the nuclear genome. Transgenes that alter transcript levels (e.g., RNAi or microRNAs) also need to be in the nucleus to be functional. Robust expression of transgenes in the *Chlamydomonas* nuclear genome is therefore essential for the functionality of many of the molecular tools for this organism; thus strategies are needed to allow robust nuclear-gene expression and to circumvent any difficulties associated with nuclear transgene expression.

**Several strategies exist to improve transgene expression in the nuclear genome**

Specific mutant backgrounds can be used to overcome silencing of transgenes by epigenetic mechanisms. Transgene inactivation occurs at both the transcriptional and post-transcriptional levels in *Chlamydomonas* (Cerutti et al., 1997a; Wu-Scharf et al., 2000; Jeong et al., 2002). Transcriptional gene silencing can occur if a transgene is inserted multiple times into the nuclear genome, resulting in subsequent DNA methylation (Cerutti et al., 1997a; Babinger et al., 2001) or histone modification (Casas-Mollano et al., 2007). Single-gene copies of transgenes can also be silenced. It was found that the *aadA* gene from *E. coli*, conferring spectinomycin resistance, stably integrates into the nuclear genome, but in approximately half of all transformants its expression is reversibly suppressed (Cerutti et al., 1997a). This gene silencing was found to be a result of transcriptional inactivation and did not correlate with methylation of the integrated DNA. Multiple intermediate levels of repression were observed in genetically identical cells. To identify the mechanism of silencing, an insertional mutagenesis screen was conducted in a background with a silenced *aadA* resistance marker, and mutants were selected for *aadA* reactivation (Wu-Scharf et al., 2000; 2014).

<table>
<thead>
<tr>
<th>Secretion signals used</th>
<th>Used to secrete</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARS1</td>
<td>Xylanase</td>
<td>Rasala et al. (2012)</td>
<td>Eichler-Stahlberg et al. (2009)</td>
</tr>
<tr>
<td>ARS2</td>
<td>Luciferase</td>
<td>Eichler-Stahlberg et al. (2009)</td>
<td></td>
</tr>
<tr>
<td>ARS2</td>
<td>Erythropoietin</td>
<td>Form <em>Lolium perenne</em> from <em>Gaussia princeps</em></td>
<td>Eichler-Stahlberg et al. (2009)</td>
</tr>
<tr>
<td>CAH1</td>
<td>Luciferase</td>
<td>Lauer sen et al. (2013a)</td>
<td></td>
</tr>
<tr>
<td>CAH1</td>
<td>Ice-binding protein</td>
<td>Lauer sen et al. (2013b)</td>
<td></td>
</tr>
<tr>
<td>Luciferase</td>
<td>Luciferase</td>
<td>Ruecker et al. (2008)</td>
<td></td>
</tr>
</tbody>
</table>

© 2015 The Authors
Jeong et al., 2002). Several important genes were identified that play a role in transgene silencing. Two proteins were found to be involved in transcriptional gene silencing, MUT9 and MUT11, which likely function in the formation of chromatin structure (Jeong et al., 2002; Casas-Mollano et al., 2008) and as a global transcriptional repressor (Zhang et al., 2002), respectively. Another protein identified, MUT6, is likely involved in post-transcriptional gene silencing and encodes a DEAH-box RNA helicase. This protein is necessary to silence transgenes by aiding in degradation of certain aberrant RNAs, which can be formed when transcripts of transgenes are improperly processed (Wu-Scharf et al., 2000). Although these mutants have repressed transgene silencing, their use as a transgene expression system may be limited by the fact that they are impaired in DNA double-stranded break repair and have decreased regulation of transposons (Jeong et al., 2002).

Several other mutant lines have been created with enhanced transgene expression. Neupert et al. (2009) developed a genetic screen for Chlamydomonas that utilized a selectable marker gene (CRY1-1) whose expression level is directly proportional to the phenotypic resistance to the selecting agent. With this system, they could increase the selection pressure and select for mutants with high transgene expression. Utilizing this selection strategy, two UV mutants were isolated that showed increased transgene expression (Neupert et al., 2009). To further demonstrate improved transgene expression they transformed these mutants, UVM4 and UVM11, with green fluorescent protein (GFP) and YFP constructs and showed increased fluorescence over non-selected strains. Interestingly, a higher percentage of the UVM4 and UVM11 transformants incorporated the full-length transgene cassette into the genome than wild-type strains.

Several follow-up studies have utilized the UVM mutants for increased transgene expression. Many of the uses have been for tagging native Chlamydomonas proteins with a fluorescent protein to determine subcellular localization (Karcher et al., 2009; Komsic-Buchmann et al., 2012; Zäuner et al., 2013) and also to express artificial miRNA (Tenenboim et al., 2014). One study used UVM4 as a background for the production of a secreted luciferase protein. The UVM4 mutant produced greater than threefold more luciferase than a wild-type strain using the same expression cassette. Additionally, the UVM4 background produced more luciferase luminescence positive transformants and the bioluminescence signals were generally more robust than wild-type (Lauersen et al., 2013a).

Another study utilized UVM mutants for the expression of three genes in the biosynthetic pathway of the terpene squalene (DXS, DXR, SQS). While the authors only saw a slight increase in co-transformation rates, compared to the fivefold increase originally reported (Neupert et al., 2009), no gene silencing was observed during the study (Kong et al., 2014). Transgene protein levels were increased in the UVM mutants over wild-type, but protein levels were variable between individual UVM transformants, which could be explained by variability in transcription levels due to positional effects of transgene genomic location. While the UVM strains have already been shown useful for transgene expression, further rounds of mutagenesis or improvements in these or other strains are needed to achieve Chlamydomonas strains with truly robust transgene expression.

Codon optimization to match the GC content of the Chlamydomonas nuclear genome may improve transgene expression. Few studies have systematically investigated the effect of codon optimization on transgene expression levels (Heitzer et al., 2007). However, circumstantial evidence suggests that codon optimization may improve transgene expression. The first attempts to transform the nuclear genome utilized yeast and bacterial genes, but these attempts were not very successful (Harris, 2009). The next attempts utilized native genes in auxotrophic backgrounds, which worked more successfully (Debuchy et al., 1998; Kindle et al., 1989). Early efforts to express foreign genes were most successful when organisms with a similar codon usage as Chlamydomonas where used as a source of genetic material (Sizova et al., 1996; Stevens et al., 1996). Chlamydomonas has a G+C nucleotide content of 68% in coding sequences (Merchant et al., 2000), and when bacterial genes, such as aadA from Escherichia coli, were used, they expressed poorly and could become silenced over time (Carutti et al., 1997a). Today, genes can be synthesized or recoded with codons that more closely match those found in Chlamydomonas and this may have contributed to the successful expression of many proteins (Fuhrmann et al., 1999; Ruecker et al., 2008; Shao and Bock, 2008).

Optimized native promoter sequences and regulatory elements help transgene expression. The most successful promoters to date have been native promoters or chimeras of native promoters. Widely used promoters include PsAD, RBCS2, and the HSP70A-RBCS2 fusion (Stevens et al., 1996; Schroda et al., 2000; Fischer and Rochaix, 2001). The use of foreign promoters, such as CaMV 35S, have been documented, but do not perform as well as endogenous promoters (Day et al., 1990; Blankenship and Kindle, 1992).

Native promoter sequences (such as that of HSP70A) can contain enhancer elements, which decrease the probability of transgene silencing and contribute to high transcription levels (Schroda et al., 2002). Modifying transgene sequences with the inclusion of an intron from a native gene can also improve transgene expression in Chlamydomonas. Introns in Chlamydomonas can contain enhancers (Kang and Mitchell, 1998;
Lumbreras et al., 1998) and it is also thought that the inclusion of introns helps in transcript maturation and export from the nucleus. Some of the strongest evidence for the improvement of transgene expression by native introns comes from resistance markers in which the addition of introns increases transformation efficiencies and tolerance to higher antibiotic titers (Lumbreras et al., 1998; Berthold et al., 2002). The addition of the RBCS2 intron 1 into the ble coding region can increase the number of transformants recovered by 10-fold, and inserting the intron twice increases transformation efficiency by up to 19-fold (Lumbreras et al., 1998). This increase in transformation efficiency corresponds to an increase in ble gene expression. The addition of the RBCS2 introns into other transgenes has also been shown to help expression, such as in luciferase (Eichler-Stahlberg et al., 2009).

**Transgene expression can be increased by coupling with resistance to a selecting agent.** Rasala et al. (2012) developed a system that links transgene expression to that of the selection marker for bleomycin antibiotic resistance, BleR. The BleR gene product neutralizes the DNA-damaging effects of bleomycin (or zeocin) by forming a 1:1 complex with the antibiotic, and is thus needed in high concentrations to protect transformed cells. In the case of the system developed by Rasala et al. (2012), the BleR gene is fused to the heterologous gene of interest via the foot-and-mouth-disease-virus (FMDV) 2A linker peptide. When this fusion protein is translated, the last two amino acids of the FMDV 2A linker fail to form a peptide bond and results in the translation of two discrete proteins, the resistance protein and the heterologous protein. Thus, the resistance to bleomycin is directly coupled to the translation of the protein of interest, allowing selection for increased production of the protein of interest. This system led to an approximately 100-fold increase in the production of a fungal xylanase over a typical expression cassette (Rasala et al., 2012). This same system was used to successfully express six fluorescent proteins (Rasala et al., 2013). A modification of this system was used to drive expression of three transgenes at once, by including a second 2A peptide from equine rhinitis A virus and a third protein coding sequence (Rasala et al., 2014). This multi-cistron transgene allowed for expression of one marker gene and two fluorescent proteins in one cassette, and in practice resulted in high expression levels of all three proteins.

**Several strategies could decrease fragmentation of the transforming DNA.** If an endonuclease is acting on transforming DNA as suggested by recent high-throughput analyses (Zhang et al., 2014), one might expect that smaller constructs would be fragmented less frequently than big constructs, as larger constructs will tend to contain more restriction sites by chance. Removal of the CA/TG recognition motifs and related variants from transforming DNA may also reduce fragmentation. Additionally, it may be possible to isolate new recipient strains that would lack the putative endonuclease, thus also having decreased cassette fragmentation. Such modifications could enhance transformation efficiencies both for insertional mutagenesis and for expression of heterologous genes in *Chlamydomonas*.

**Fluorescent proteins and bioluminescent reporter genes can be used in *Chlamydomonas***

The use of fluorescent proteins (Figure 3 and Table 3) in *Chlamydomonas* has been limited by poor transgene expression from the nuclear genome and high levels of autofluorescence. The first successful fluorescent protein expressed in *Chlamydomonas* was a codon-optimized version of GFP from the jellyfish, *Aequorea victoria* (Fuhrmann et al., 1999). This protein was successfully expressed when fused to the ble antibiotic resistance gene (Fuhrmann et al., 1999; Rasala et al., 2012). This GFP has also been fused to endogenous genes. Fusions to flagella proteins have been the most successful (Schoppmeier et al., 2005; Diener, 2009), due in part to the lack of fluorescent pigments within the flagella. Fusions of other proteins and GFP have been less successful. To visualize the plasma membrane protein Rhesus 1-GFP fusion, a white carotenoid biosynthesis mutant was needed (Yoshihara et al., 2008).

A YFP gene optimized for expression in mammalian cells has been successfully visualized and expressed in *Chlamydomonas*, though it was only visible in mutants with increased transgene expression (UVM4 and UVM11) and not visible in wild-type cells (Neupert et al., 2009). A cyan fluorescent protein (CFP) gene that was co-optimized for expression in *Chlamydomonas* and *Pyropia* sp. was successfully visualized in UVM11 (Lim et al., 2013).

Recently, several other fluorescent proteins have been codon-optimized for *Chlamydomonas* and evaluated for functionality. Rasala et al. (2013) synthesized genes for blue (mTagBFP), cyan (mCerulean), yellow (Venus), orange (tdTomato), and red (mCherry) fluorescent proteins. Using the ble-2A expression strategy (Rasala et al., 2012), all of the fluorescent proteins were visible above background fluorescence. The brightest fluorescence was observed with the tdTomato protein, which unfortunately is twice the size of the other fluorescent proteins. mCherry was the second brightest and was considered to have the best overall performance of all the fluorescent proteins evaluated. Venus also performed well, whereas mtagBFP and mCerulean both had high background fluorescence. Interestingly, the *Chlamydomonas* GFP (Fuhrmann et al., 1999) was the worst overall performer in the study, in part to a high background of green autofluorescence and key amino acid changes in the protein (Rasala et al., 2013). The poor
performance from the most widely used fluorescent protein in the literature may explain some of the dogma of difficulties associated with fluorescent protein use in Chlamydomonas; however, this should be eliminated with the advent of these more robust fluorescent protein systems.

The bioluminescent protein luciferase has also been used as a reporter gene in Chlamydomonas. Luciferases from *Renilla reniformis* and *Gaussia princeps* have been codon optimized for expression in Chlamydomonas and have been used to monitor gene expression in the nuclear genome (Fuhrmann et al., 2004; Croft et al., 2007; Ruecker et al., 2008; Shao and Bock, 2008). On direct comparison, the version of luciferase from *G. princeps* had several fold higher bioluminescence activity than the *R. reniformis* enzyme (Ruecker et al., 2008; Shao and Bock, 2008). Luciferase has been successfully secreted from Chlamydomonas (Eichler-Stahlberg et al., 2009; Lauersen et al., 2013a) and could have applications as a reporter for gene expression or as an environmental sensor.

Epitope tags can also be successfully added to Chlamydomonas proteins and are useful for immunoprecipitation, immunolocalization, and protein purification (Figure 3 and Table 4) (Kozminski et al., 1993; Sugiuara et al., 1998; Petersen and Small, 2001; Silflow et al., 2001; Lechtreck et al., 2009; Sizova et al., 2013; Gao et al., 2014).

The Chlamydomonas Center is an invaluable resource

The Chlamydomonas Resource Center at the University of Minnesota (http://chlamycollection.org/) is a community and NSF-supported organization supplying a wide variety of *Chlamydomonas* strains and mutants, DNA constructs (containing cloned Chlamydomonas genes, widely used promoters, selectable marker genes), and other resources of significant value to the worldwide Chlamydomonas research community. A companion web site, the Chlamydomonas Connection (http://www.chlamy.org/), is a gateway providing access to genomic, genetic, and bibliographic information on Chlamydomonas and related algae.

**THE DEVELOPMENT OF MOLECULAR TECHNIQUES AND RESOURCES WILL CONTINUE TO PROPEL CHLAMYDOMONAS RESEARCH**

Despite all of the advantages of using *Chlamydomonas* as a model system and the extensive amount of molecular tools that have been adapted for its use, several dogmas exist about using this organism. These include the ideas that *Chlamydomonas* is unable to express transgenes effectively, that the high G+C content of the nuclear genome makes molecular biology in Chlamydomonas difficult, and that determining the location of DNA insertions in the nuclear genome can be challenging.

In many cases, newly developed molecular tools are invalidating these long held assumptions. A variety of techniques exists to aid in nuclear transgene expression, including the use of mutant backgrounds in which epigenetic silencing is suppressed, expression systems that can produce high levels of transgene expression [such as the ble-2A (Rasala et al., 2012)], and modular vector toolkits that have been designed specifically for Chlamydomonas (Lauersen et al., 2015). Techniques for high-throughput mapping of genomic insertions have also been developed,
which can accelerate mutant identification (Zhang et al., 2014).

With the development of new molecular tools for *Chlamydomonas*, increasingly challenging biological questions can be answered with this organism. Advances in tagging proteins with various fluorescent markers has aided with protein localization studies. Systems for targeted disruption of genes have been demonstrated and proven effective in knocking out a few genes, but their true potential still has not been realized. The development of a whole-genome knockout library will revolutionize mutant screening and will be a powerful resource for reverse genetic studies.

The field is poised to benefit from the introduction of several new techniques and the optimization of existing techniques. The continued development of genome editing technologies such as CRISPR/Cas9 and TALENs will likely be as beneficial in *Chlamydomonas* as it has been in other organisms for making targeted modifications to the genome (Jiang et al., 2013). Finally, synthetic chromosomes have been used in other organisms for the expression of multiple genes or entire metabolic pathways (Annaluru et al., 2014). Such a technology could be a very powerful synthetic biology tool for *Chlamydomonas*, and could assist in developing this organism for advanced biotechnological uses. A synthetic chloroplast genome has already been developed for use in *Chlamydomonas* (O’Neill et al., 2012), and an artificial nuclear chromosome would be a welcome complement to this technology.

Past innovations have been inspired by work in other systems, but *Chlamydomonas* could serve as a model system to enable the development of easy to use and broadly applicable tools for organisms that are harder to work with. This includes development of novel insertion mutant mapping technologies and strategies to overcome transgene silencing. Overall, the molecular tools that have been developed and are continually being developed for *Chlamydomonas* will keep generating research advances in this and other organisms, while directly answering key biological questions.

### ACKNOWLEDGEMENTS

We thank Susan Dutcher, Elizabeth Freeman, Andre Greiner, Peter Hegemann, Xiaobo Li, Luke Mackinder, Stephen Mayfield,

### Table 3 Reporter proteins successfully used in *Chlamydomonas*

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>mTagBFP</td>
<td>Codon optimized, expressed with ble-2A system</td>
<td>Rasala et al. (2013)</td>
</tr>
<tr>
<td>CFP</td>
<td>Codon optimized, expressed in UVM11</td>
<td>Lim et al. (2013)</td>
</tr>
<tr>
<td>mCerulean</td>
<td>Codon optimized, expressed with ble-2A system; Codon optimized</td>
<td>Rasala et al. (2013) and Lauersen et al. (2015)</td>
</tr>
<tr>
<td>GFP</td>
<td>Codon optimized</td>
<td>Fuhrmann et al. (1999)</td>
</tr>
<tr>
<td>Clover</td>
<td>Codon optimized</td>
<td>Lauersen et al. (2015)</td>
</tr>
<tr>
<td>Venus</td>
<td>Codon optimized, expressed with ble-2A system; Codon optimized</td>
<td>Rasala et al. (2013) and Lauersen et al. (2015)</td>
</tr>
<tr>
<td>YFP</td>
<td>Not optimized, expressed in UVM4 and UVM11</td>
<td>Neupert et al. (2009)</td>
</tr>
<tr>
<td>tdTomato</td>
<td>Codon optimized, expressed with ble-2A system</td>
<td>Rasala et al. (2013)</td>
</tr>
<tr>
<td>mRuby2</td>
<td>Codon optimized</td>
<td>Lauersen et al. (2015)</td>
</tr>
<tr>
<td>mCherry</td>
<td>Codon optimized, expressed with ble-2A system</td>
<td>Rasala et al. (2013)</td>
</tr>
<tr>
<td>Luciferase</td>
<td>From <em>Renilla reniformis</em>; Codon optimized; Secreted using ARS2 export sequence</td>
<td>Fuhrmann et al. (2004) – original and Eichler-Stahlberg et al. (2009)</td>
</tr>
<tr>
<td>Luciferase</td>
<td>From <em>Gaussia princeps</em>; secreted using CAH1 export sequence</td>
<td>Ruecker et al. (2008), Shao and Bock (2008) and Lauersen et al. (2013a)</td>
</tr>
<tr>
<td>Gus</td>
<td>β-Glucuronidase</td>
<td>Pratheesh et al. (2014)</td>
</tr>
<tr>
<td>Ars</td>
<td>Arylsulfatase</td>
<td>Davies et al. (1992) and Specht et al. (2015)</td>
</tr>
</tbody>
</table>

### Table 4 Epitope tags used in *Chlamydomonas*

<table>
<thead>
<tr>
<th>Tag</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLAG tag</td>
<td></td>
<td>Karcher et al. (2009), Rasala et al. (2012), Sizova et al. (2013) and Gao et al. (2014)</td>
</tr>
<tr>
<td>Influenza virus hemagglutinin (HA)</td>
<td>1× HA (Kozinski et al., 1993); 3× HA (Silflow et al., 2001)</td>
<td>Kozinski et al. (1993), Silflow et al. (2001) and Lechtreck et al. (2009)</td>
</tr>
<tr>
<td>Polyhistidine (His)</td>
<td>6× His (Gulis et al., 2008)</td>
<td>Sugiura et al. (1998) and Gulis et al. (2008)</td>
</tr>
<tr>
<td>Metal affinity tag (MAT)</td>
<td></td>
<td>Rasala et al. (2012)</td>
</tr>
<tr>
<td>Myc</td>
<td></td>
<td>Petersen and Small (2001)</td>
</tr>
</tbody>
</table>

© 2015 The Authors

Christine Roche, Irina Sizova, and Ru Zhang for useful feedback on this manuscript. This review was supported by the Carnegie Institution for Science, a Simons Fellowship of the Life Science Research Foundation to REJ, and NSF grants MCB-1146621 and IOS-1359682 to MCJ.

REFERENCES


Molecular techniques in Chlamydomonas


