

SCIENTIFIC REPORTS



OPEN

CRISPR/Cas9-induced knockout and knock-in mutations in *Chlamydomonas reinhardtii*

Received: 01 April 2016

Accepted: 25 May 2016

Published: 13 June 2016

Sung-Eun Shin^{1,*}, Jong-Min Lim^{2,*}, Hyun Gi Koh¹, Eun Kyung Kim³, Nam Kyu Kang¹, Seungjib Jeon¹, Sohee Kwon¹, Won-Sub Shin¹, Bongsoo Lee¹, Kwon Hwangbo^{2,4}, Jungeun Kim^{5,6}, Sung Hyeok Ye^{5,7}, Jae-Young Yun⁵, Hogyun Seo⁸, Hee-Mock Oh², Kyung-Jin Kim⁸, Jin-Soo Kim^{5,6}, Won-Joong Jeong², Yong Keun Chang^{1,3} & Byeong-ryool Jeong¹

Genome editing is crucial for genetic engineering of organisms for improved traits, particularly in microalgae due to the urgent necessity for the next generation biofuel production. The most advanced CRISPR/Cas9 system is simple, efficient and accurate in some organisms; however, it has proven extremely difficult in microalgae including the model alga *Chlamydomonas*. We solved this problem by delivering Cas9 ribonucleoproteins (RNPs) comprising the Cas9 protein and sgRNAs to avoid cytotoxicity and off-targeting associated with vector-driven expression of Cas9. We obtained CRISPR/Cas9-induced mutations at three loci including *MAA7*, *CpSRP43* and *ChlM*, and targeted mutagenic efficiency was improved up to 100 fold compared to the first report of transgenic Cas9-induced mutagenesis. Interestingly, we found that unrelated vectors used for the selection purpose were predominantly integrated at the Cas9 cut site, indicative of NHEJ-mediated knock-in events. As expected with Cas9 RNPs, no off-targeting was found in one of the mutagenic screens. In conclusion, we improved the knockout efficiency by using Cas9 RNPs, which opens great opportunities not only for biological research but also industrial applications in *Chlamydomonas* and other microalgae. Findings of the NHEJ-mediated knock-in events will allow applications of the CRISPR/Cas9 system in microalgae, including “safe harboring” techniques shown in other organisms.

Microalgae have emerged as the third generation biofuel feedstocks that are excellent renewable and sustainable energy sources for the future. Their use would allow us to replace the current fossil fuels that have been blamed for serious environmental problems, including pollution and global warming¹. Microalgae have many advantages over the previous generations of biofuel feedstocks: they have much higher photosynthetic productivity, require less land, may not compete with food sources, and can generate more value-added products². However, microalgal biofuels are currently impractical due to their high production costs. This can be partly solved by genetic improvements of microalgal strains to produce more biomass and/or lipids that can be converted to biofuels.

The green unicellular flagellate alga, *Chlamydomonas reinhardtii*, is a good model organism, with genetic, molecular and genomic resources available for both research fields of basic and applied sciences. Numerous techniques have been used to transform the nucleus, chloroplast and mitochondria of this microalga^{3–5}, allowing genetic improvements of traits for various purposes. These advanced resources and techniques have been applied to other microalgae for more practical purposes: for example, transformation techniques have been used

¹Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea. ²Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 34141, Republic of Korea. ³Advanced Biomass R&D Center (ABC), KAIST, Daejeon 34141, Republic of Korea. ⁴Department of Biological Science, Chungnam National University (CNU), Daejeon 34134, Republic of Korea. ⁵Center for Genome Engineering, Institute for Basic Science (IBS), Seoul 08826, Republic of Korea. ⁶Department of Chemistry, Seoul National University (SNU), Seoul 08826, Republic of Korea. ⁷Basic science, IBS school, Korea University of Science and Technology (UST), Seoul 08826, Republic of Korea. ⁸School of Life Sciences, KNU Creative BioResearch Group, Kyungpook National University (KNU), Daegu 41566, Republic of Korea. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to W.-J.J. (email: wonjoong@kribb.re.kr) or Y.K.C. (email: changyk@kaist.ac.kr) or B.-r.J. (email: bjeong@kaist.ac.kr)

to improve production of biomass and/or biofuels in industrial microalgae by overexpression of certain metabolic or regulatory genes^{6,7}. However, a remaining challenge in the genetic engineering of microalgae is the need for efficient methods to specifically knock-down or knockout unwanted genes.

Targeting specific genes for down-regulation or disruption has been well established in many organisms^{8–12}. In microalgae, genes can be down-regulated by RNA silencing techniques involving RNA interference (RNAi) or artificial microRNAs (amiRNA)^{13–17}. These techniques can be used to study essential genes, for which recessive mutations are lethal, particularly in haploid microalgae; however, these methods are often limited by low efficiency, non-specific targeting and silencing of the RNA constructs^{9,18}. Stable and permanent disruption of unwanted genes can be achieved by homologous recombination (HR), and by sequence-specific recombinant nucleases including zinc-finger nuclease (ZFN) and transcription activator-like effector nuclease (TALEN)^{11,12}. In some microalgae, only a few reports have described the specific disruption of genes via HR^{19,20} and ZFN- or TALEN-based knockout^{21–23}; however, these techniques are very hard to achieve in most microalgae.

Recently, a new system based on components of the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated protein 9 (Cas9), has become very popular in disrupting a gene due to its simplicity, improved specificity and higher versatility compared to ZFN and TALEN^{24,25}. The CRISPR/Cas9 system was initially identified as a bacterial adaptive immune system against invading plasmids and viral DNAs, where the system was found to cleave target DNAs guided by small RNAs^{24,26}. Since the introduction of CRISPR/Cas9 for heterologous genome editing²⁷, a plethora of studies have used CRISPR/Cas9 to knockout targeted genes²⁸. However, it has proven difficult to use this system for genome editing in microalgae. Only one such report exists for *Chlamydomonas*, and extremely low targeting efficiency was observed, possibly reflecting toxicity induced by vector-driven Cas9 expression²⁹.

In this study, we improved the targeting efficiency of CRISPR/Cas9 in *Chlamydomonas* by directly delivering the Cas9 protein and single-chain guide RNAs (sgRNAs), known as Cas9 ribonucleoprotein (RNP), which has been successfully tested in other organisms^{28,30,31}. We targeted three genes: the *MAA7* gene that encodes the beta subunit of tryptophan synthase (TSB); the antennal assembly gene *CpSRP43* that encodes chloroplast SRP43; and the chlorophyll biosynthetic gene *ChlM* that encodes Mg-protoporphyrin IX S-adenosyl methionine O-methyl transferase. Mutations of *MAA7* can be positively selected using 5-fluoroindole (5-FI)^{32,33}, while mutations of *cpsrp43* and *chlM* can be selected by a lighter colony color due to reduced antenna size and chlorophyll synthesis, respectively^{34,35}. For the latter two genes, we employed co-transformation of vectors that confer resistance to hygromycin due to low delivery efficiency in *Chlamydomonas*. Interestingly, the vectors used in our mutagenic screen were found to be integrated at the Cas9 cut site, indicating that the knock-in mutations were mediated by non-homologous end joining (NHEJ). In summary, we obtained many auxotrophic and visible colonies from the primary screen, and succeeded in obtaining sizable number of mutations targeted at the Cas9 cut sites.

Results

Design of sgRNAs and *in vitro* assessment of the Cas9 activity. To avoid the probable problems of toxicity²⁹ and off-targeting³⁰ that have been associated with vector-driven expression of the guide RNA and Cas9, we used synthetic sgRNAs and the Cas9 protein manufactured by ToolGen, Inc. (Republic of Korea)³⁰ instead of cloning their genes in a vector. The sgRNAs for the individual target genes were designed to maximize the generation of mutations, as described in the Methods section as per the manufacturer's guidelines (ToolGen, Inc., Republic of Korea). Target genes were chosen to facilitate easy selection of mutants. Mutants of the *MAA7* gene were identified via tryptophan auxotrophic selection using 5-FI^{32,33}, whereas mutations in *CpSRP43* and *ChlM* were identified by visual selection due to the reduced chlorophyll contents of these mutants^{34,35}.

The activities of the sgRNAs and Cas9 protein were assessed by *in vitro* reactions using individual genomic fragments containing the Cas9 cut site (Fig. 1). The primers used to amplify the relevant genomic fragments are summarized in Table S1. Figure 1 summarizes the genomic maps, the locations and sequences of the guide RNAs, the protospacer adjacent motifs (PAMs) and the Cas9 cut site for *MAA7*, *CpSRP43* and *ChlM*. We tested two sgRNAs for *MAA7* and found that the first (*MAA7*-1) failed to cleave, while the second (*MAA7*-2) succeeded in cleaving the genomic DNA target of 783 bp into two fragments of 488 bp and 295 bp (Fig. 1d left panel). The *CpSRP43* Cas9 RNP cleaved the 795 bp target DNA into 418 bp and 377 bp fragments, while the *ChlM* Cas9 RNP cleaved the targets of 1010 bp into 274 bp and 736 bp fragments (Fig. 1d right panels). Consistent with these *in vitro* results, our later *in vivo* mutagenic screens failed to isolate any targeted mutations for the *MAA7*-1 sgRNA (See below, and Table 1).

Small indel mutations found among the *MAA7* auxotrophic mutants. *In vivo* mutagenesis was carried out by electroporating *MAA7* Cas9 RNPs into *C. reinhardtii* CC-124. To optimize the delivery and mutagenic efficiencies, we performed experiments in which the amounts of sgRNA and Cas9 were varied while the mass ratio was kept constant at 4:3, as suggested by ToolGen, Inc. (Table 1). Selection with 5-FI yielded three and 21 tryptophan auxotrophic mutants from experiments performed using the *MAA7*-1 and *MAA7*-2 sgRNAs, respectively. The growth of some of these primary auxotrophic mutants on media containing tryptophan and 5-FI is shown in Fig. 2A. A *MAA7* RNAi mutant (*MAA7* RNAi #33) was used as a reference strain^{17,36}. Our mutant strains all showed comparable tolerances to 5-FI, but were consistently less tolerant than *MAA7* RNAi #33 (Fig. 2A). Genomic DNA was amplified using primers M2_F and M2-R (Table S1), and sequencing of the generated fragments revealed that zero and eight targeted mutations had been obtained using the *MAA7*-1 and *MAA7*-2 sgRNAs, respectively (Table 1, Fig. 2B). Therefore, as noted above, the *MAA7*-1 sgRNA failed to cleave the target DNA *in vitro* and produced no mutations *in vivo*. It is not clear why we obtained such different outcomes from the two target sites closely located on the same locus. However, our results suggest that it is important to test the compatibility of the sgRNA and Cas9 protein prior to undertaking a laborious mutagenic screen.

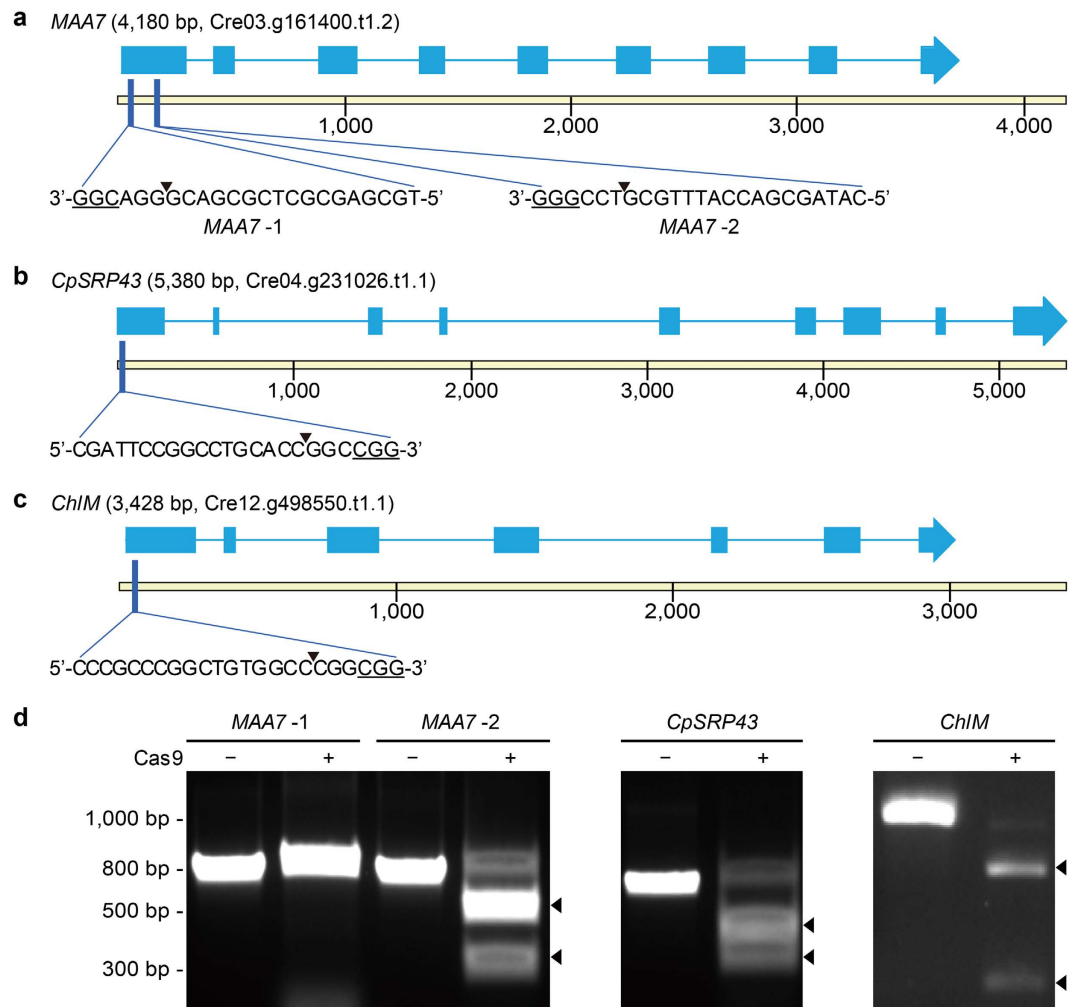


Figure 1. Map of CRISPR/Cas9 target sites in the *MAA7*, *CpSRP43* and *ChlM* loci, and *in vitro* cleavage of target DNAs by sgRNAs and the Cas9 protein. (a–c) Molecular maps of the *MAA7* (a), *CpSRP43* (b) and *ChlM* (c) loci with the CRISPR/Cas9 target sites indicated. CDS sequences are indicated as blue boxes, together with the sgRNA target sequences, the cut sites (arrowheads) and the associated PAM sites (5'-NGG-3', underlined). Target sites of *MAA7* were written in the 3'-to-5' orientation, since the negative strand was targeted by sgRNAs. (d) The activities of the Cas9 protein and individual sgRNAs were tested *in vitro*. The Cas9-degraded products of the amplified fragments are indicated by arrowheads, with degraded products of 488 and 295 bp, 418 and 377 bp, and 736 and 274 bp identified for *MAA7*, *CpSRP43* and *ChlM*, respectively.

Sequencing of the eight targeted mutations revealed an interesting pattern of small indel mutations that involved three base pairs or multiples. This yielded point mutations of only a few amino acids rather than the frameshift mutations commonly found in CRISPR/Cas9-induced mutations in other organisms. Our targeted mutations could be grouped into four types. The main type (found in mutants of *MAA7* 20-1, 40-1, 40-5, 40-12 and 40-20; mutant designation indicated amount sgRNA followed by a serial number.) was deletion of the CGC next to the Cas9 cut site producing a D54E point mutation and a deletion of A55. For convenience, we defined a conserved motif, RPDAN, flanking the Cas9 cut site (Fig. 2B); it was changed to RPEN in these mutants. The other three mutation types were alteration of the RPDAN motif to RPAN (*MAA7* 40-22), alteration of the RPDAN motif to RPDN (*MAA7* 40-6), and a 33 bp insertion at the Cas9 cut site (*MAA7* 20-10) that changed the RPDAN motif to RPDKWSLCYSESYAN via insertion of the 11 underlined amino acids.

To assess the significance of the RPDAN motif and validate our CRISPR/Cas9-induced mutagenic screen using 5-FI, we performed additional experiments including phylogenetic analyses of TSB homologs, sequencing of *MAA7*-associated genes, and complementation of *MAA7* mutants. Phylogenetic analyses of TSB homologs from various organisms representing the kingdoms of bacteria, plants and microalgae revealed that the RPDAN motif was a part of a conserved region found in all organisms (Fig. 2C; Fig. S1 for the full alignment; Fig. S2 for the phylogenetic tree). The R52 residue was conserved in plants and microalgae, P53 and D54 were invariable in all studied organisms, and A55 appeared to be conserved in bacteria and microalgae. Bacteria and microalgae that shared conservation of A55 appeared to share aspects of the entire sequence phylogeny, as they grouped together in the phylogenetic tree (Fig. S2). The RPDAN motif is clearly distant from the catalytic site of TSB³³. We

| Gene | Type | | Amounts of sgRNA (Cas9 protein) | | | | |
|---------|-----------|---|---------------------------------|--------------------------|-------------------------|-------------------------|----------------------|
| | | | Control | 10 μ g (7.5 μ g) | 20 μ g (15 μ g) | 40 μ g (30 μ g) | Total |
| MAA7-1 | Selection | Colonies | 0 | 1 | 2 | – ^a | 3 |
| | | Efficiency (cell ⁻¹) ^b | 0 | 1.1×10^{-8} | 6.7×10^{-8} | – | 3.3×10^{-8} |
| | Targeting | Colonies | 0 | 0 | 0 | – | 0 (0%) ^c |
| | | Efficiency (cell ⁻¹) | 0 | 0 | 0 | – | 0 |
| MAA7-2 | Selection | Colonies | 1 | 6 | 5 | 9 | 20 |
| | | Efficiency (cell ⁻¹) | 1.1×10^{-8} | 6.7×10^{-8} | 5.6×10^{-8} | 1.0×10^{-7} | 2.2×10^{-7} |
| | Targeting | Colonies | 0 | 0 | 2 | 6 | 8 (40%) |
| | | Efficiency (cell ⁻¹) | 0 | 0 | 2.2×10^{-8} | 6.7×10^{-8} | 8.9×10^{-8} |
| CpSRP43 | Selection | Colonies | 65 | 34 | 108 | 76 | 218 |
| | | Efficiency (cell ⁻¹) | 7.2×10^{-7} | 3.8×10^{-7} | 1.2×10^{-6} | 8.4×10^{-7} | 2.4×10^{-6} |
| | Targeting | Colonies | 0 | 3 | 0 | 0 | 3 (1.4%) |
| | | Efficiency (cell ⁻¹) | 0 | 3.3×10^{-8} | 0 | 0 | 3.3×10^{-8} |
| ChlM | Selection | Colonies | – | 6000 | – | – | 6000 |
| | | Efficiency (cell ⁻¹) | – | 3.0×10^{-5} | – | – | 3.0×10^{-5} |
| | Targeting | Colonies | – | 10 | – | – | 10 (0.17%) |
| | | Efficiency (cell ⁻¹) | – | 5.0×10^{-8} | – | – | 5.0×10^{-8} |

Table 1. Selection and targeting efficiencies in the CRISPR/Cas9 mutagenic screens. ^aNot tested. ^bThe efficiency was calculated by dividing the number of colonies with that of cells used for electroporation. ^cTargeting percentage was calculated by dividing the number of targeted colonies with that of selected colonies.

suggest that it may be involved in the interaction with tryptophan synthase alpha subunit (TSA), based on our structural modeling of TSA and TSB in bacteria (Fig. S3). When we compared the reported bacterial TSA-TSB structures with the amino acid sequence of *MAA7*, we found that the RPDAN motif was located proximal to the four-stranded β -sheet that is involved in the interaction between TSA and TSB (Fig. S3). Because the interface between the two subunits is crucial for the necessary consecutive enzyme reactions³⁷, we speculate that the mutation-based impairment of interface might reduce the enzyme activity of the protein. To exclude the possibility that this second locus might be involved in the suppression during the auxotrophic selection, we sequenced the genes encoding TSA and TSB in WT and mutants *MAA7* 20-1, 40-22, 40-6 and 20-10 (representing all four mutation classes). We did not identify any mutations other than those at the Cas9 cut site as shown in Fig. 2B. Alternatively, the RPDAN motif might be involved in other critical functions such as stability of the protein, which may explain the low expression of mutated TSB (Fig. S4c).

Complementation of the *MAA7* mutations. We further analyzed the functional significance of the RPDAN motif *in vivo* by performing complementation analyses. The TSB coding sequences (CDSs) were amplified from WT and the *MAA7* 20-1 mutant (containing RPEN instead of the RPDAN motif), cloned into the pCr202 vector³⁸ together with sequences encoding the FLAG tag, and transformed into *MAA7* RNAi #33 (Fig. S4a). The WT TSB CDS was also transformed into *MAA7* 20-1 to see if the defective TSB function of this mutant could be complemented by WT TSB. PCR was used to confirm transformation (Fig. S4b), and the expression of the TSB proteins tagged with the FLAG tag was assessed by Western blotting using an anti-FLAG antibody (Fig. S4c). All strains complemented with WT TSB (designated 20 W for the *MAA7* 20-1 mutant complemented with the WT TSB; RW for *MAA7* RNAi #33 transformed with the WT TSB) showed relatively high level of TSB expression. For unknown reasons, strains complemented with the mutagenized TSB CDS from *MAA7* 20-1 showed relatively lower expression levels.

Complementation was evaluated by analyzing the sensitivity to 5-FI (Fig. 3). All strains used in the complementation assay showed normal growth when cultured in TAP medium supplemented with tryptophan (Fig. 3A). In the presence of 5-FI, *MAA7* RNAi #33 transformed with WT TSB showed almost complete complementation when WT TSB expression was high (RW #19), but incomplete complementation when low WT TSB expression was low (RW #13). This dose dependence appears to be reasonable considering that higher level of WT TSB expression can counteract the RNA silencing in *MAA7* RNAi #33. WT TSB also completely complemented *MAA7* 20-1 as shown in 20 W #1 and 20 W #6 (Fig. 3B). This indicates that the defects found in *MAA7* 20-1 (Fig. 2A) arose solely from the RPEN mutation of the RPDAN motif, since no other mutation was found in the genes encoding TSA or TSB.

The results obtained from complementation with the TSB CDS containing point mutations in the RPDAN motif provide additional evidence for the functional importance of this motif. We transformed *MAA7* RNAi #33 with the TSB CDS from *MAA7* 20-1, which contained the RPEN mutation of the RPDAN motif (designated R20 #58 and R20 #61). Compared to the strains complemented with WT TSB (RW #13 and RW #19), these R20 strains completely failed to complement the RNAi mutation (Fig. 3B), even though the expression of TSB-FLAG in these strains was low (Fig. S4c). Nevertheless, this result suggests that the RPEN mutant protein lacks a significant portion of the WT TSB function. Notably, mutations in the RPDAN motif were weaker than the RNAi knockdown mutation (*MAA7* RNAi #33), as their tolerance to 5-FI was lower than that of *MAA7* RNAi #33 (Figs 2A and 3B).

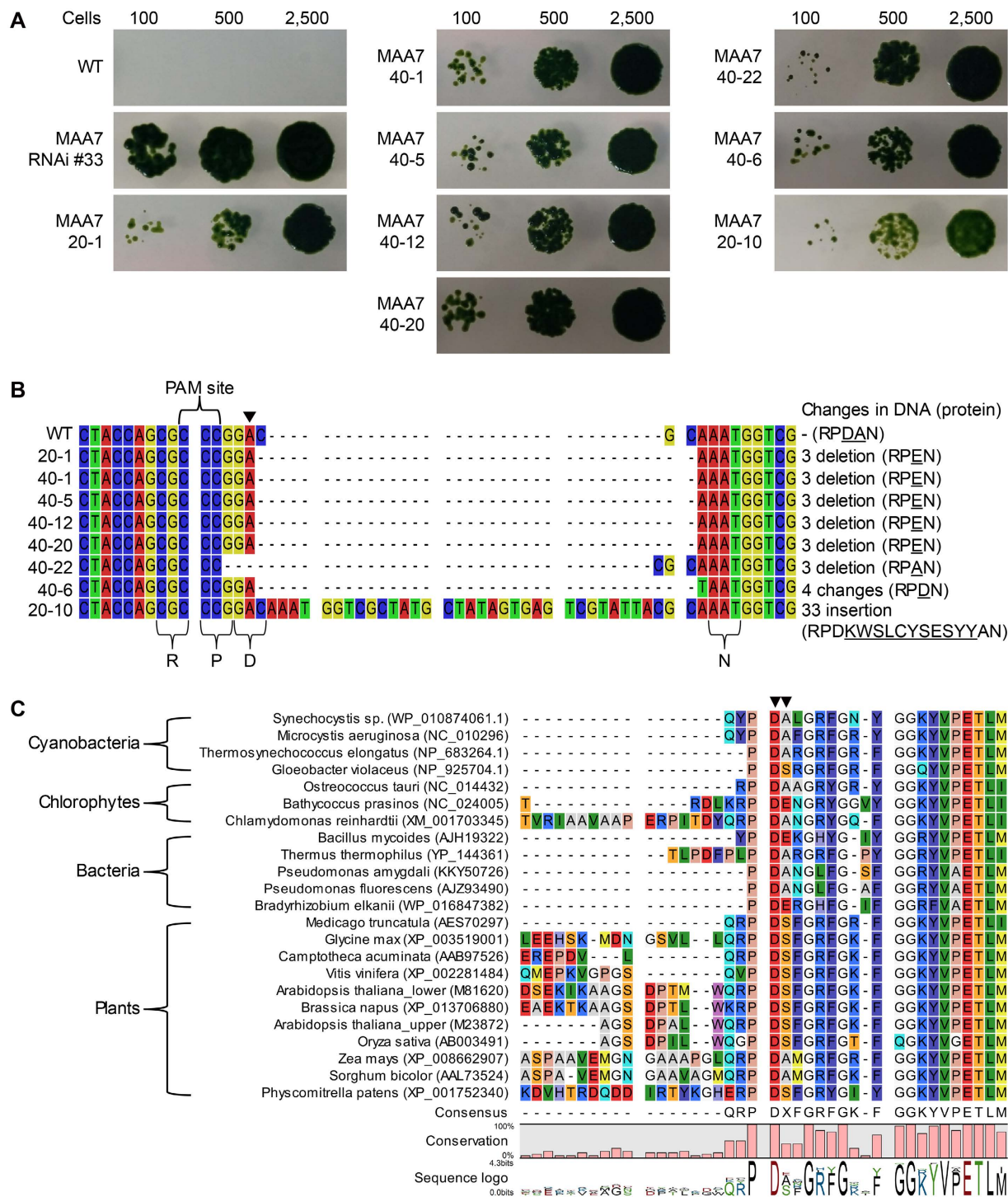


Figure 2. Auxotrophic selection of *MAA7* mutants and their sequence analyses. (A) The isolated *maa7* mutants were grown in the presence of 5-FI with tryptophan supplementation, and cell growth was examined by spot-test loading of 100, 500, and 2500 cells. The RNAi mutant, *MAA7* RNAi #33, was used as a reference strain. (B) Small indels are produced by CRISPR/Cas9 at the Cas9 cut site of the *MAA7* locus. Sequencing of genomic DNA fragments encompassing the cut site revealed eight small indels at the cut site (at the 3' side of the arrow head). The protein sequences of the target region (the RPDAN motif) are listed on the right side, and the altered sequences are underlined. (C) Alignments of the N-terminal side of the TSB sequences from various organisms representing three kingdoms. GenBank IDs are included to facilitate identification. Mutated sequences are marked with arrowheads.

NHEJ-mediated knock-in mutations at the CpSRP43 locus. We tested our CRISPR/Cas9 systems for two more genes that affect the chlorophyll contents allowing visual selection of their mutants. Considering the extremely low delivery efficiency of transgenic materials in *Chlamydomonas*, we co-transformed vectors carrying

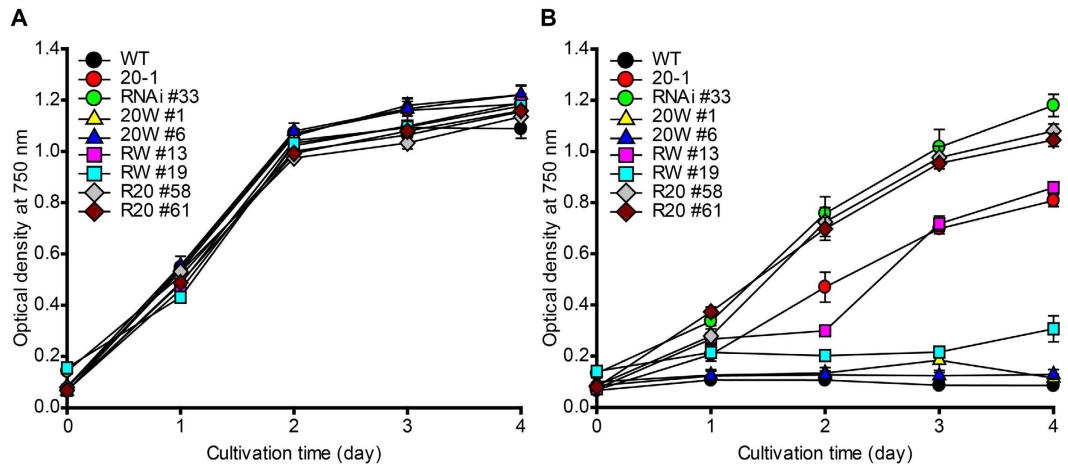


Figure 3. Complementation of *MAA7* mutants with WT and *MAA7* 20-1 TSB sequences. The *MAA7* 20-1 mutant was transformed with WT *MAA7* gene for complementation (generating strain 20W). RNAi #33 (a reference strain with a RNAi knockdown mutation of *MAA7*) was transformed with the WT (generating RW) or 20-1 (generating R20) *MAA7* genes. (A) Cell growth was analyzed in TAP medium supplemented with tryptophan. (B) The tolerance of each strain to 5-FI was analyzed by culturing cells in TAP medium containing both 5-FI and tryptophan. Error bars indicate standard deviations obtained from four independent experiments.

a selection maker. We were thus able to first select transformed colonies, and then identify mutants by visual selection of lighter green colonies. We targeted *CpSRP43* by delivering *CpSRP43*-specific sgRNA (Fig. 1b,d), Cas9 protein and *Xba*I-linearized pCr202 vector. Our initial screen yielded 283 colonies that showed resistance to hygromycin; from them, we selected three colonies that stably showed a lighter green coloration. *CpSRP43* is involved in antennal protein assembly and its mutations are well known. We obtained two deletion mutants, CC-4561 and CC-4562³⁴, from the *Chlamydomonas* Resource Center, and used these as reference strains. We also included two negative controls: *CpSRP43* 20-1 was from the same screen but showed a normal green colony color; *vec-2* was from a vector-only transformation with no sgRNAs or Cas9, and had a normal green colony color.

Our phenotypic analyses began with a visual assessment in which we loaded the same number of cells to each well of a multi-well plate, and compared their colors (Fig. 4A). Consistent with the colors observed for their colonies, the isolated candidates of *cpsrp43* mutants induced by CRISPR/Cas9 (*CpSRP43* 10-1, 10-2 and 10-3) were light green, similar to the *cpsrp43* deletion mutants CC-4561 and CC-4562. In contrast, the WT control, CC-124, and the other negative controls were dark green (Fig. 4A). We further analyzed the chlorophyll contents of these strains, and calculated the ratio of chlorophyll a to chlorophyll b, which is a simple criteria used for reduced antenna mutants^{39,40}. As expected from their lighter green color, the chlorophyll contents of the candidate mutants were low, and their ratios of chlorophyll a to chlorophyll b were higher than 8, similar to the deletion mutants (Fig. 4B). Critically, when Western blotting was performed with an anti-*CpSRP43* antibody, our CRISPR/Cas9-induced mutants and the deletion mutants all lacked the *CpSRP43* protein with the expected size of 43 kDa (Fig. 4C). These data indicate that *CpSRP43* sgRNA and Cas9 successfully induced mutations at the *CpSRP43* locus.

The molecular nature of the CRISPR/Cas9-induced mutants was analyzed with PCR using primers SRP43_2F and SRP43_2R (Table S1) with the expected size of 795 bp encompassing the Cas9 cut site. Initially, we failed to detect any amplification of the relatively small fragments with exception of a few fragments that were found to be nonspecific upon sequencing (Fig. 5A top panel). However, all of the transformants contained the vector sequence for the hygromycin resistance gene (Fig. 5B). Repeated PCR experiments using primer pairs spanning different areas of the locus yielded similar results. We thus sequenced their genomes, and found that the vector sequence was integrated at the Cas9 cut site with small insertions of a few nucleotides at either end (Fig. 5C). The insertion was further verified with PCR using different primer pairs (amplifying regions A, B and C) (Fig. 5B). Using optimized PCR conditions, we successfully amplified the whole *CpSRP43* locus in the mutants (Fig. 5A bottom panel). Assembly of all sequence information revealed that the vector sequence had been re-organized prior to integration as shown in Fig. 5C. All three mutants contained the same integrated sequence, suggesting that they arose from a single transformation event. We also characterized their integration in the genome by performing the Southern blot, and found that the integration was a single copy insertion with the same size of restriction fragments (Fig. S5A,B).

Off-targeting of the *CpSRP43* Cas9 RNP was assessed in our mutagenic screen by using Cas-OFFinder (<http://www.rgenome.net/cas-offinder/>)⁴¹. The guide sequence was searched on the *Chlamydomonas* reference genome sequence allowing 3 mismatches and 2 bulges, and a total of 404 off-target candidate sites could be identified. Among these, we were able to locate 340 sites on our WT (CC-124) genomic sequence that matched those of the *Chlamydomonas* reference sequence (CC-503)⁴. We surveyed these 340 sequences on the genomic sequences of *CpSRP43* 10-1, 10-2 and 10-3, and confirmed the presence of 333 sites. These sites contained no mutations,

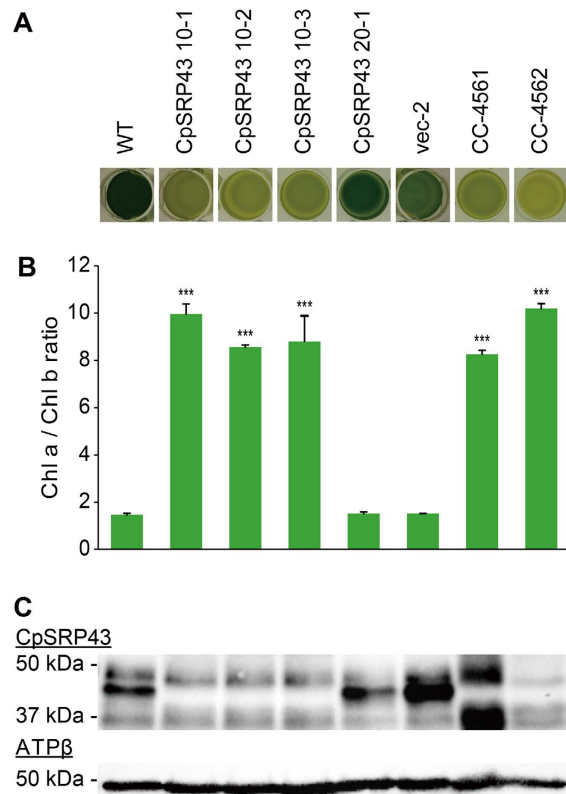


Figure 4. Phenotypic characterization of CpSRP43 mutants. Cells were co-transformed with sgRNA, the Cas9 protein and the selection vector, and three colonies showing a stable lighter green color were isolated and designated CpSRP43 10-1, 10-2 and 10-3. The negative controls included a colony from the same screen that carried the normal green color (designated CpSRP43 20-1), and a colony from an experiment in which cells were transformed with only the selection vector (designated vec-2). The positive controls included CC-4561 (mt+) and CC-4562 (mt-), which are deletion mutants of *CpSRP43* (also called *tla3-cpsrp43*). (A) The coloration of control colonies and those of CpSRP43 mutants induced by CRISPR/Cas9. Cells (2.50×10^7 cells/mL) were loaded to a 24-well plate in identical volumes of TAP medium. (B) Changes in the chlorophyll contents of CpSRP43 mutants. The contents of chlorophyll a and b were measured at 72 hours of cultivation, normalized by the cell density. The ratios of chlorophyll a to chlorophyll b were calculated and plotted. Error bars indicate standard deviations obtained from four independent experiments. Significant differences compared to the WT were determined by the Student's *t* test and are indicated by asterisk (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$) (C) Immunoblots were probed with antibodies against the target protein (CpSRP43) and the β -subunit of ATP synthase (ATP β). A band of the molecular mass expected for the former (43 kDa) is marked with an arrowhead.

suggesting that there were no, or rare, incidences of off-targeting with the CpSRP43 Cas9 RNP. This may be possible because the Cas9 RNP could have been degraded and/or diluted soon after electroporation.

NHEJ-mediated knock-in mutations at the *ChlM* locus. Finally, we used our optimized CRISPR/Cas9 delivery conditions to mutate another gene, *ChlM*, involved in chlorophyll biosynthesis³⁵. Vector pCr112 containing the hygromycin resistance marker was co-transformed for selection purpose. We obtained about 6000 initial transformants. From them, we isolated 10 clones that showed yellow or brown colony color, and designated ChlM-3, -4, -9, -10, -11, -17, -19, -21, -22, -24 (Fig. 6). Visual assessment of coloration (Fig. 6a) and measurement of chlorophyll contents (Fig. 6b) revealed that the mutant phenotype was associated with the expected chlorophyll deficiency³⁵. We performed semi-quantitative RT-PCR, and found no detectable expression of *ChlM* mRNA in the mutants (Fig. 6c). Southern blot analyses were performed to characterize genomic integration patterns after digestion of their genomic DNAs with *NcoI* that does not cut the vector (Fig. S5c,d). Most of the selected transformants contained a single copy of the vector except for ChlM-4 that contained two copies. Two pairs of clones (ChlM-10 and ChlM-17; ChlM-22 and ChlM-24) showed the same banding patterns, suggesting that they were duplicate clones representing the same integration events. We thus observed at least eight independent integration events, most of which were single-copy integration of the vector.

We further characterized the molecular nature of the CRISPR/Cas9-induced mutations at the *ChlM* locus (Fig. 7). The initial PCR for the *ChlM* locus spanning the Cas9 cut site failed to amplify genomic DNA from the isolated mutants. We tested combinations of primers (Table S1) located in the *ChlM* locus and the vector, and successfully determined the junction sequences for six clones including ChlM-3, -9, -10, -19, -22 and -24

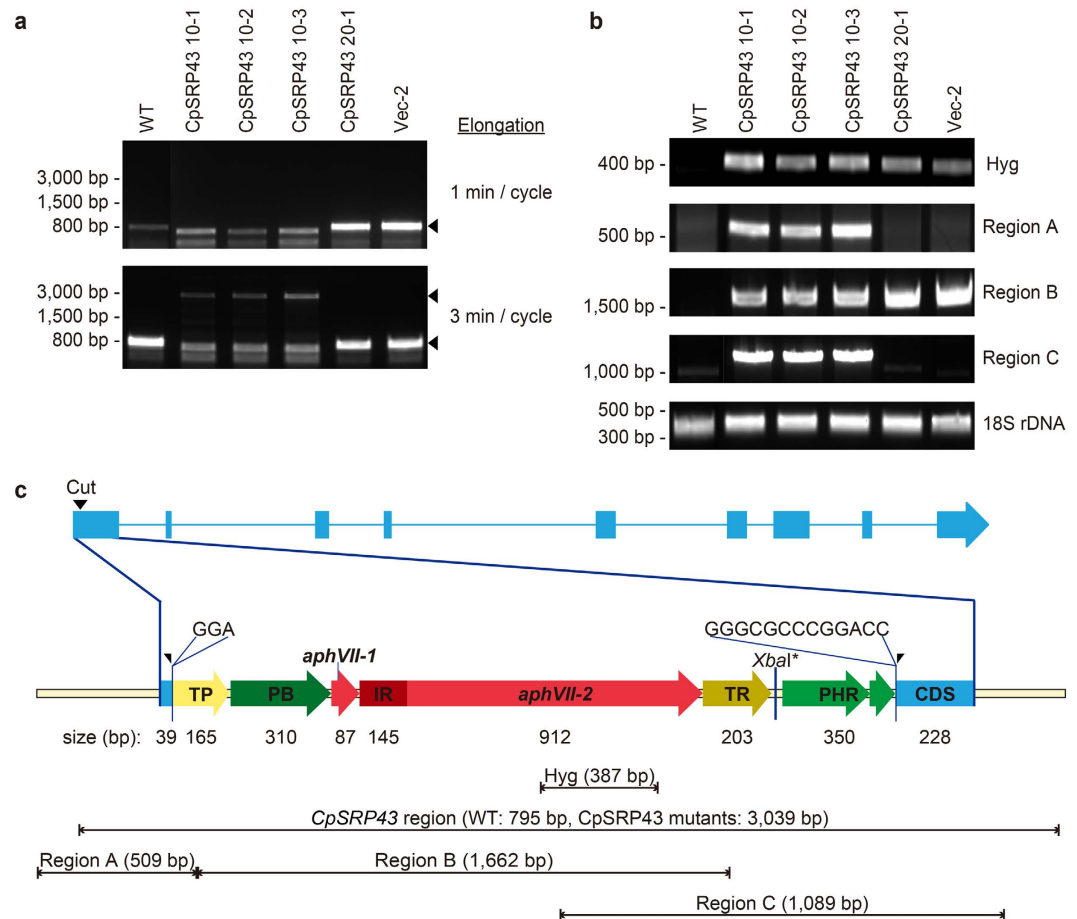


Figure 5. NHEJ-mediated knock-in of the co-transformed vector at the CRISPR/Cas9 cut site in the *CpSRP43* locus. **(a)** PCR amplification of the *CpSRP43* regions was performed using different elongation times. Amplification of fragments encompassing the cut site failed under a short elongation cycle (upper panel), but improved when we used a longer cycle (lower panel). **(b)** Various vector sequences were confirmed by PCR. The amplicon sizes of the fragments encompassing the hygromycin resistance region, region A, region B, region C, and the 18S rDNA were 387 bp, 509 bp, 1,662 bp, 1,089 bp and 380 bp, respectively. The specific primers used in this study are listed in Table S1. **(c)** Molecular map of the *CpSRP43* locus and the integrated vector sequences. An insertion of 3 bp (GGA) was identified next to the 5' side of the Cas9 cut site (left-half arrowhead). Considerable vector sequences (2,650 bp) were inserted at the cut site, followed by an additional 13 bp of unknown origin (GGGCGCCCGGACC) at the 3' side of the cleaved site (right-half arrowhead). All of the *CpSRP43* mutants (10-1, 10-2 and 10-3) contained the same sequences at the Cas9 cut site in the *CpSRP43* locus. Abbreviations: TP, PsaD terminator; PB, beta2-tub promoter; IR, RBCS2 intron; TR, RBCS2 terminator; and PHR, HSP70A-RBCS2 promoter.

(Fig. 7B–G). All of them showed integration of the vector at the Cas9 cut site, and were characterized by different rearrangements of vector sequences and different small indels next to the Cas9 cut site. ChLM-22 and ChLM-24 showed the same vector sequence, but contained different deletions at the junctions (Fig. 7E,G). Our results indicate that the 10 isolated mutants represent at least nine independent integration events with extensive rearrangements of the vector sequences. Interestingly, the six clones with identified junction sequences were examples of NHEJ-mediated knock-in of the vector, because the vector contained no sequences homologous to the *ChlM* locus. We failed to identify the integration junction sequences for the remaining four clones; however, ChLM-4, ChLM-11 and ChLM-17 showed the same bands hybridized to both probes of *ChlM* and *aphVII*, suggesting that they are also knock-ins of the vector sequence. Only ChLM-21 showed different bands for the *ChlM* and *aphVII* probes, suggesting that the integration occurred at a location outside of the *ChlM* locus. ChLM-21 showed a smaller band for the *ChlM* probe compared to that of WT, suggesting that it contained a sizable deletion of about 200 bps (Fig. S5c,d). The nature of ChLM-21 has not been analyzed further, even though we speculated that there was a deletion of about 200 bps in the *ChlM* locus based on the Southern blot (Fig. S5c,d). It should be noted that ChLM-4 contained two copies of vector sequences: one at the *ChlM* locus and the other at an unknown location, based on the Southern blot (Fig. S5d). These mutations could be categorized into three classes as summarized in Fig. S5: simple knock-ins (ChLM-3, -9, -10, -11, -17, -19, -22, -24), knock-in plus integration at another location (ChLM-4), and deletion plus integration at other location (ChLM-21).

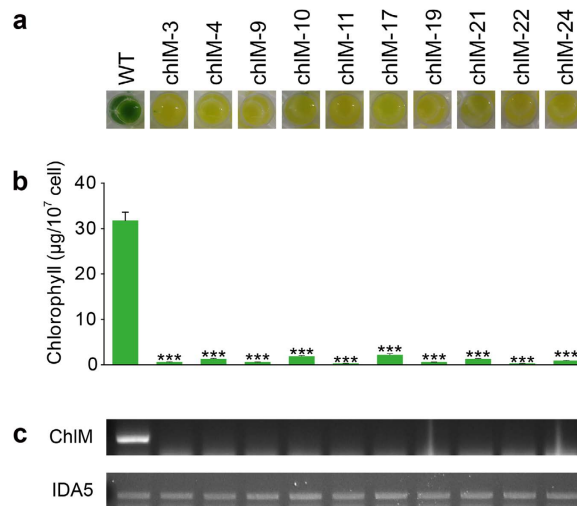


Figure 6. Phenotypic characterization of light green mutants induced by CRISPR/Cas9 at the *ChlM* locus. (a) Coloration of 10 *ChlM* mutants (2.0×10^8 cells/mL) loaded to a 24-well plate in liquid TAP medium. (b) Chlorophyll contents of the *ChlM* mutants. Data are expressed as \pm SD ($n = 4$ replicates). Significant differences compared to the WT control were determined by the Student's *t* test and are indicated by asterisks ($*P < 0.05$, $**P < 0.01$, $***P < 0.001$). (c) Semi-quantitative RT-PCR analysis was used to detect *ChlM* transcripts from the mutant lines. *IDA5* was detected as a loading control.

Discussion

We successfully generated a sizable number of CRISPR/Cas9-induced mutations at the *MAA7*, *CpSRP43* and *ChlM* loci by delivering Cas9 RNP. This is a significant improvement over the first report of a CRISPR/Cas9-induced mutation in *Chlamydomonas*²⁹. This improvement may be attributed to the delivery of the Cas9 RNP, as employed successfully in other organisms^{28,30,31}. Vector-driven expression of Cas9 may be toxic to *Chlamydomonas*, probably due to the continuous expression²⁹. Delivery of purified Cas9 RNP may have other advantages including technical simplicity and fewer occurrences of off-target events^{30,42}. From the auxotrophic selection of *MAA7* mutations, we obtained eight small indel mutations that were specifically induced by the *MAA7-2* sgRNA and the Cas9 protein. Occurrence of these targeted mutations was highest when cells were delivered with the highest amount of Cas9 and *MAA7-2* sgRNA (Table 1), even though we have not determined optimum conditions that might have to be determined empirically. However, none of the CRISPR/Cas9-induced mutations were frame-shifts; we observed only small indels of three base pairs or multiples. This odd mutation pattern defies a clear explanation, but there are reports of non-random occurrence of mutations induced by CRISPR/Cas9, including the insertion of an A or in-frame deletions^{43–45}. The error-prone double strand break (DSB) repair of DNA must be involved in generation of small indel mutations⁴⁶, and the mechanism of DSB repair in *Chlamydomonas* might have played a role in the unique pattern of mutations obtained in our mutagenic screen. We also found that the indels generated at the *ChlM* locus (not counting the vector insertion) tended to be in-frame except for *ChlM-24*: *ChlM-3* and *ChlM-22* showed no change; *ChlM-9* and *ChlM-19* contained in-frame indels. Alternatively, complete loss-of-function mutations could not survive the selection conditions employed in our auxotrophic screens. Even though tryptophan was supplemented in the media, uptake of tryptophan in *Chlamydomonas* is not known, since no true loss-of-function mutations of TSB have been characterized^{17,33}.

It is interesting to note that the integration of the vector sequences at the Cas9 cut sites of *CpSRP43* and *ChlM* was mediated by NHEJ, since the co-transformed vectors did not contain any sequence from the target genes or the neighboring loci. Knock-in can be achieved by homology-driven recombination (HDR) wherein the vector sequence contains wings of target-flanking sequences, or by NHEJ that does not require any cloning of the wings^{24,47–51}. It is not clear why we obtained predominantly knock-in events. It is possible that DSBs induced by Cas9 facilitated integration of available free ends of co-transformed vectors at the cut site. Our present finding opens up a range of possible applications of NHEJ-mediated knock-in, including insertion of a gene at a specific location without requiring the cloning of homologous sequences.

Conclusively, delivery of Cas9 RNPs improved mutagenic efficiency and specificity of the CRISPR/Cas9 system at the *MAA7*, *CpSRP43* and *ChlM* loci in *Chlamydomonas*. The efficiency of obtaining targeted mutations was up to 100 fold improvement compared to the first report of CRISPR/Cas9-induced mutagenesis in *Chlamydomonas*²⁹. The direct delivery of Cas9 RNPs may have other advantages including minimized off-targeting events, less toxicity of Cas9, and no laborious cloning work compared to the transgenic techniques³⁰. In our mutagenic screen targeting the *CpSRP43* locus, sequencing the whole genomes of *cpsrp43* mutants revealed no incidences of off-targeting by using Cas-OFFinder⁴¹. We also found predominant NHEJ-mediated knock-in events when co-transformed with unrelated vectors used for the selection purpose. New concepts of the targeted transgene integration for stable expression, such as the “safe harbor” technique and other gene manipulations, are already underway in some organisms^{47,52}. This work can be applied not only to biological research in

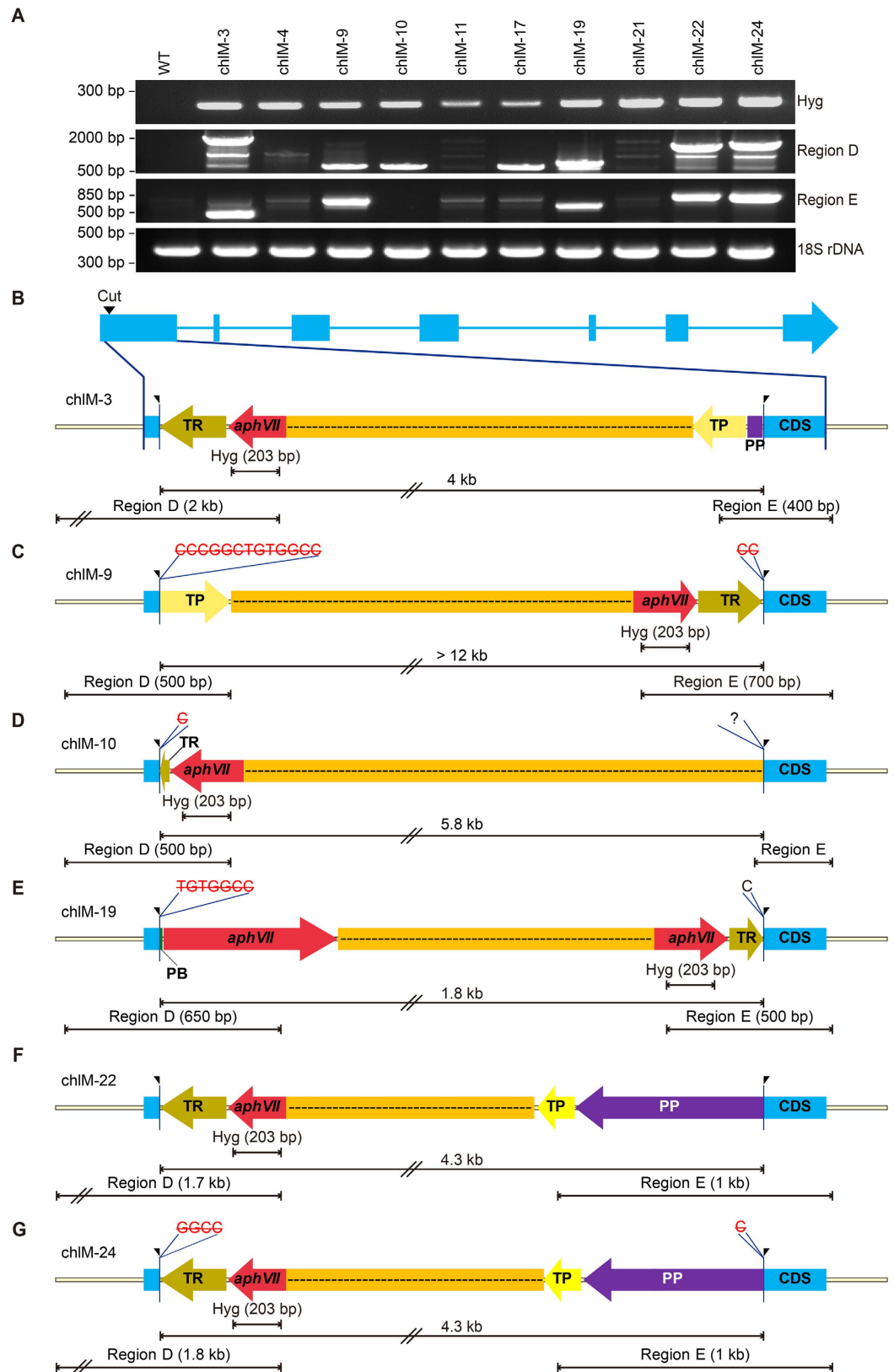


Figure 7. NHEJ-mediated knock-in mutants of the co-transformed vector at the CRISPR/Cas9 cut site on the *ChlM* locus. (A) Multiple PCR fragments encompassing the vector sequences. (B–G) Schematic molecular maps of the knock-in events in each ChIM mutant. The sizes of the inserted vector and PCR amplification are shown. Red and black letters indicated the deleted and inserted sequences, respectively. Abbreviations: TR, RBCS2 terminator; TP, PsaD terminator; PP, PsaD promoter; and PB, beta2-tub promoter.

Chlamydomonas, but also to industrial applications of genome editing in other microalgae for large scale production of biofuels and other biomaterials.

Methods

Target genes and sgRNA design. We chose target genes based on our ability to identify mutations via auxotrophic or visual means. The *MAA7* gene (GenBank accession: XM_001703345) in *Chlamydomonas* encodes the beta subunit of tryptophan synthase (TSB; XP_001703397), and its mutations can be positively selected using 5-FI^{32,33}. Mutations in the other two tested genes could be selected based on the lighter colony color arising from reduced chlorophyll synthesis^{34,35}. The *CpSRP43* gene (GenBank accession: XM_001703652) and *ChlM* gene (GenBank accession: XM_001702328) encodes chloroplast SRP43 (XP_001703704) and Mg-protoporphyrin IX S-adenosyl methionine O-methyl transferase (XP_001702380), respectively. For the latter two genes, we co-transformed cells with a hygromycin-resistance vector and used antibiotic selection to improve the delivery efficiency in *Chlamydomonas*.

The sgRNAs were designed by ToolGen, Inc (Republic of Korea) based on the following inclusion criteria: (1) the target site is located within 50% of coding sequence (CDS) from the N terminus; (2) the out-of-frame score is >60; and (3) the mismatch is 2 or greater. The generated sgRNAs included the following guide RNA sequences. For *MAA7*, *MAA7*-RG1 (5'-UCGAGCGCUCGCGACGGGA-3') and *MAA7*-RG2 (5'-CAUAGCGA CCAUUUGCGUCC-3'); for *CpSRP43*, *CpSRP43*-RG1 (5'-CGAUUCCGGCCUGCACCGGC-3'); and for *ChlM*, *ChlM*-RG1 (5'-CCCGCCCGCUGUGGCCCGG-3'). Their locations along each locus are shown in Fig. 1a–c. All sgRNAs included the common sequence containing the crRNA, tetraloop (underlined) and tracrRNA: GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAA GUGGCACCGAGUCGGUGCUUUUUUU²⁴. The Cas9 protein contained a nuclear localization signal at the C-terminus, of which sequence is PKKKRKV.

Microalgal strains and culture conditions. *Chlamydomonas reinhardtii* CC-124 was used as the wild type, and was obtained from the Chlamydomonas Resource Center in the University of Minnesota (<http://www.chlamycollection.org/>). An RNAi knockdown mutant of *MAA7* (RNAi #33) was kindly provided by Heriberto Cerutti of Nebraska University-Lincoln¹⁷. Mutant strains of *CpSRP43* CC-4561 (*tla3-Δcpsrp43* cw+ mt+) and CC-4562 (*tla3-Δcpsrp43* cw+ mt-) were obtained from the Chlamydomonas Resource Center, and were used as positive controls for our CRISPR/Cas9-induced mutants³⁴. All strains were maintained on Tris-acetate-phosphate (TAP) agar plates [2.42 g/L Tris, 0.375 g/L NH₄Cl, 0.1 g/L MgSO₄ 7H₂O, 0.05 g/L CaCl₂ 2H₂O, 0.0108 g/L K₂HPO₄, 0.0054 g/L KH₂PO₄, 1 mL/L glacial acetic acid, and 1 mL/L Hutner's trace elements (50 g/L Na₂EDTA 2H₂O, 22 g/L ZnSO₄ 7H₂O, 11.4 g/L H₃BO₃, 5.06 g/L MnCl₂ 4H₂O, 1.61 g/L CoCl₂ 6H₂O, 1.57 g/L CuSO₄ 5H₂O, 1.10 g/L (NH₄)₆Mo₇O₂₄ 7H₂O, and 4.99 g/L FeSO₄ 7H₂O) at 25 °C under continuous illumination of 50 μmol m⁻² s⁻¹.

In vitro digestion of target DNAs with the Cas9 protein. The activities of the sgRNAs and the Cas9 protein were verified *in vitro* prior to their use *in vivo*. The DNA targets for this assay were produced by PCR using the Instagene Matrix (Bio-Rad, USA). Briefly, cells were harvested by centrifugation and genomic DNAs were extracted according to the instructions provided with the Instagene Matrix. Substrate DNAs encompassing the Cas9 cut site were amplified using 2X EF-Taq PCR Pre-mix (Solgent, Republic of Korea). The utilized primers were as follows: for *MAA7*, M2_F and M2_R; for *CpSRP43*, SRP43_2F and SRP43_2R; and for *ChlM*, CrChlM_F1 and CrChlM_R3 (Table S1). The PCR conditions consisted of 35 cycles of 95 °C for 20 sec, 52–55 °C for 40–60 sec and 72 °C for 1 min. The elongation time was extended to 3 min for the longer amplicon of the *CpSRP43* knock-in mutants. The PCR products were purified with a QIAquick Gel Extraction Kit (Qiagen, the Netherlands). For *in vitro* assays, 500 ng of sgRNA and 600 ng of Cas9 protein were pre-mixed at 37 °C for 5 min. For experiments, 100 ng of PCR amplicons and 3.1 NEB buffer (NEB, USA) were added, and the mixtures were incubated at 37 °C for 2 hours. RNaseA (Bioneer, Republic of Korea) was added, and the mixture was incubated at 37 °C for 20 min. 6× STOP solution (0.5 M EDTA, 80% glycerol, and 10% SDS) was added, and the mixture was incubated 37 °C for 20 min. The final products were separated by 0.8% agarose gel electrophoresis.

Transformation and selection of *C. reinhardtii*. For transformation of *C. reinhardtii*, we performed electroporation using a modification of a previously reported protocol⁵³. Cells were cultured under continuous illumination (200 μmol m⁻² s⁻¹) for 2–3 days to a cell density of 3 × 10⁶ cells/ml. The cells were then concentrated to 3 × 10⁸ cells/ml in MAX Efficiency Transformation Reagent for Algae (Thermo Fisher Scientific, USA), and electroporation was conducted using an ECM 830 Square Wave Electroporation System (Bio-Rad, USA) and cuvettes with a 0.2-cm gap (BTX, USA). Prior to electroporation, various amounts of Cas9 and sgRNAs were incubated together at 37 °C for 30 min (Table 1), and cells (300 μL) were placed in the cuvette and cooled on ice for 5 min. To enable visual selection of *CpSRP43* and *ChlM* knockout mutants, 1 μg of pCr202 (for *CpSRP43*) or pCr112 (for *ChlM*) were linearized with *Xba*I and added to the mixture, thereby conferring resistance to hygromycin³⁸. The electroporation voltage and pulse interval were 250 V and 15 ms, respectively. Electroporated cells were recovered for 16 hours in TAP medium containing 40 mM of sucrose, with agitation (120 rpm) under continuous low light. Mutant cells were selected on TAP agar plates containing 1.5 mM of L-tryptophan and 25 μM of 5-FI (Sigma-Aldrich, USA) for tryptophan auxotrophic selection, or on TAP agar plates with 15 mg/L of hygromycin B (Thermo Fisher Scientific, USA) for co-transformation.

Complementation of CRISPR/Cas9-induced *MAA7* mutants. To confirm the functional significance of the isolated *MAA7* mutations, we cloned the *MAA7* CDS from CC-124 and *MAA7* 20-1 and performed complementation assays on *MAA7* 20-1 and the RNAi mutant, *MAA7* RNAi #33¹⁷. Cells were exponentially

grown in TAP with shaking at 120 rpm under continuous light of $120 \mu\text{mol m}^{-2} \text{s}^{-1}$, and total RNAs were isolated using the RNeasy Plant Mini kit (Qiagen, the Netherlands) and the RNase-Free DNase Set (Qiagen, the Netherlands) according to the manufacturer's instructions. Any remaining DNAs were removed with a DNA-free DNase kit (Ambion, USA), and cDNAs were produced using Superscript III Reverse Transcriptase and oligo (dT)₂₀ primers (both from Invitrogen, USA). Individual CDSs were amplified using the Phusion High-Fidelity DNA Polymerase (NEB, USA) and primers I202TSB_3F and I202TSB_3R, and the obtained fragments were cloned into the pCr202 vector³⁸ using a Gibson Assembly kit (NEB, USA) according to the manufacturer's instructions. CDS expression was controlled by the HSP70A and RBCS2 promoter and terminated by the *psaD* terminator³⁴. We added a translation initiation context (AACA) in front of the start codon⁵⁵ and a FLAG-tag sequence (GACTACAAGGACGACGACGACAAG) next to the CDS⁵⁶. Cloned plasmids were delivered to *MAA7* 20-1 and *MAA7* RNAi #33 by electroporation, as described in the previous section. After recovery, transformed cells were spread on TAP agar plates containing 1.5 mM L-tryptophan and 15 mg/L hygromycin B. After 1 week, hygromycin-resistant colonies were subjected to PCR verification of the following: transformant status, as assessed by the presence of the *aphVII* gene (using primers *aph72_L* and *aph72_R*); the presence of the transgenic *MAA7* gene, as assessed using TSB flag_4F and TSB flag_4R; and proper amplification of the 18S rDNA (using primers SR6 and SR9) as a positive control. Expression of the transgenic *MAA7* protein was confirmed by Western blotting using a rabbit anti-FLAG-tag antibody (diluted 1:1000; Cell Signaling Technology, USA). Clones of Western-positive cells were subjected to complementation assays by growth in 24-well plates (SPL, Republic of Korea) under continuous light ($120 \mu\text{mol m}^{-2} \text{s}^{-1}$) with shaking at 150 rpm for 4 days in the presence or absence of 25 μM 5-FI. Cell growth was monitored using the OD_{750 nm}, which was assessed using a SpectraMax M2e Microplate reader (Molecular Devices, USA).

Chlorophyll fluorescence. Chlorophyll contents were measured using methanol extraction according to published protocols⁵⁷. Briefly, 5 mL of cells were centrifuged at $5,035 \times g$ for 10 min and washed twice with distilled water twice. Pellets were resuspended in the same volume of methanol and incubated in the dark at 4 °C for 1 hour. After centrifugation at $5,035 \times g$ for 10 min, the optical density of supernatant was measured at 652 and 665 nm using a UV-1800 UV-VIS Spectrophotometer (Shimadzu, Japan). The concentrations of chlorophyll a and b were calculated as follow:

$$\begin{aligned} \text{Chl a } [\mu\text{g/mL}] &= -8.0962 \times \text{OD}_{652 \text{ nm}} + 16.5169 \times \text{OD}_{665 \text{ nm}} \\ \text{Chl b } [\mu\text{g/mL}] &= 27.4405 \times \text{OD}_{652 \text{ nm}} - 12.1688 \times \text{OD}_{665 \text{ nm}} \end{aligned}$$

The obtained chlorophyll contents were normalized to the cell density, which was measured using a Cellometer AutoT4™ (Nexcelom Bioscience, USA).

Genomic DNA extraction and whole genome sequencing. Genomic DNA was extracted according to published methods^{58,59}. Briefly, cells were grown to the stationary phase, harvested, washed with 50 mM EDTA, and resuspended in 150 μL distilled water and 300 μL SDS-EB. DNA was extracted with a mixture of phenol and chloroform (1:1), with repeated applications of the mixture (500 μL) made until the interphase became clear. The DNA was washed with 500 μL chloroform and precipitated with two volumes of ethanol. The obtained DNA pellet was washed with 70% ethanol and air-dried, and the DNA was resuspended in 40 μL of TE buffer. Whole-genome sequencing was performed using an Illumina HiSeq 2000 (Illumina, USA) according to the provided protocol. Genomic DNA sequences were read with a paired-end sequencing platform (Sequencing, Republic of Korea). The SolexaQA package software (version 1.13) was used to trim the short read sequences for better quality. More specifically, the DynamicTrim module was used to eliminate poor-quality bases (Phred score < 20) and the LengthSort module was used to exclude short read lengths < 25 bp. The trimmed reads were assembled using the SOAPdenovo version 2.04 (<http://soap.genomics.org.cn//soapdenovo.html>) under default options⁶⁰. The assembled contigs were aligned with the *C. reinhardtii* reference genome v5.5 (<http://www.phytozome.net>) and with the sequence of the vector pCr202 for *CpSRP43* mutagenic experiments. BLASTN searches were performed using cutoffs of $1e-10$ and 90% identity.

Phylogenetic analysis. Phylogenetic analyses of TSB sequences were performed using the CLC Genomics Workbench version 7.0 (CLC bio, USA). Twenty-three TSB amino acid sequences from organisms representing all five kingdoms were obtained from GenBank (Fig. 2C). For the alignment parameters, the gap-open cost and gap-extension cost were set at 10.0 and 1.0, respectively. The full sequence of the TSB protein, including the Cas9 cut site, is presented in Fig. S1. A phylogenetic tree of the aligned sequences was constructed using the neighbor-joining method, and the Jukes-Cantor model was used to measure protein distance.

Western blot analysis. To confirm our knockout of the *CpSRP43* protein, Western blotting was performed. *C. reinhardtii* cells were grown to the exponential phase, and the cell density was normalized to 10^7 cells. The cells were harvested, washed once with distilled water, and collected by centrifugation at 13,000 rpm for 10 min. The cell pellets were resuspended with $1.5 \times$ Laemmli buffer (62.5 mM Tris-HCl pH 7.6, 7% sodium dodecyl sulfate, 25% glycerol, 5% β -mercaptoethanol, and 0.02% bromophenol blue) and incubated at 100 °C for 5 min. The samples were centrifuged at 13,000 rpm for 5 min, and the supernatants were loaded to Mini-PROTEAN® TGX Stain-Free™ Gel of Any KD (Bio-Rad, USA) and run in $1 \times$ TGX buffer. The resolved proteins were transferred to PVDF membranes using the Trans-Blot® Turbo™ System (Bio-Rad, USA). The membranes were blocked with $1 \times$ PBS containing 5% skim milk and 0.1% Tween 20 with agitation. The *CpSRP43* protein was detected using polyclonal antibodies raised against the antigenic sequence, CERRFKIRWSDGYPTSWEPEE³⁴. The β -subunit of ATP synthase (ATP β) was detected as a loading control. Then, the blots were incubated with horseradish peroxidase (HRP)-conjugated anti-rabbit secondary antibodies (Cell Signaling Technology, USA) for 1 hour, and the

results were visualized by chemiluminescent detection using the Clarity Western ECL substrate and a ChemiDoc system (Bio-Rad, USA).

Southern blot analysis. For Southern blot analysis of the CpSRP43 and ChlM mutants, 10 µg of genomic DNA was digested with *Pst* I/*Nco* I and *Nco* I, respectively, separated by 0.8% agarose gel electrophoresis, and then blotted onto Hybond-N⁺ nylon membranes (Amersham Biosciences, Sweden). 0.3-kb PCR fragments corresponding to the *aphVII*, *CpSRP43* and *ChlM* genes were used as probes. The primers used to generate each probe are listed in Table S1. ³²P-labeled probes were produced using the RediprimeTM II Random Labeling System (Amersham Biosciences, Sweden) and hybridization was performed following the manufacturer's instructions. The probe-bound membrane was washed and exposed to a phosphor-imaging plate (IP) for 3 days, and the obtained signals were detected using a Bio-Imaging Analyzer BAS-1800II (Fuji, Japan).

Semi-quantitative RT-PCR analyses. Total RNA was extracted from WT and ChlM mutants using the TRIzol reagent (Invitrogen, USA) according to the manufacturer's instructions. Reverse transcription was carried out using 2 µg total RNA, 50 µM oligo (dT)₁₈, 200 U M-MLV reverse transcriptase (Promega, USA), 500 µM of each dNTP, and 20 U ribonuclease inhibitor (Thermo Fisher Scientific, USA). For semi-quantitative RT-PCR, cDNAs were amplified with 1.5 U Ex Taq DNA polymerase (TaKaRa, Japan), 100 µM of each dNTP and 10 pmol of each gene-specific primer, using a T1 thermal cycler (Biometra GmbH, Germany). Amplification was performed with 25–40 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 30 s. The primers used to amplify the ChlM and IDA5 cDNAs are listed in Table S1.

References

- Ho, S. H., Ye, X., Hasunuma, T., Chang, J. S. & Kondo, A. Perspectives on engineering strategies for improving biofuel production from microalgae - A critical review. *Biotechnol Adv* **32**, 1448–1459 (2014).
- Wijffels, R. H. & Barbosa, M. J. An outlook on microalgal biofuels. *Science* **329**, 796–799 (2010).
- Harris, E. H. *Chlamydomonas* as a model organism. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 363–406 (2001).
- Merchant, S. S. *et al.* The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**, 245–250 (2007).
- Mussgnug, J. H. Genetic tools and techniques for *Chlamydomonas reinhardtii*. *Appl. Microbiol. Biotechnol.* **99**, 5407–5418 (2015).
- Kang, N. K. *et al.* Effects of overexpression of a bHLH transcription factor on biomass and lipid production in *Nannochloropsis salina*. *Biotechnol Biofuels* **8**, 200 (2015).
- Yao, Y. *et al.* Glycerol and neutral lipid production in the oleaginous marine diatom *Phaeodactylum tricornerutum* promoted by overexpression of glycerol-3-phosphate dehydrogenase. *Biotechnol Biofuels* **7**, (2014).
- Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
- Gaj, T., Gersbach, C. A. & Barbas, C. F. 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (2013).
- Molnar, A. *et al.* Highly specific gene silencing by artificial microRNAs in the unicellular alga *Chlamydomonas reinhardtii*. *Plant J.* **58**, 165–174 (2009).
- Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* **161**, 1169–1175 (2002).
- Mussolino, C. *et al.* A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* **39**, 9283–9293 (2011).
- Ly, H. *et al.* Transcriptome analysis of *Chlamydomonas reinhardtii* during the process of lipid accumulation. *Genomics* **101**, 229–237 (2013).
- Zhu, B. H. *et al.* Silencing UDP-glucose pyrophosphorylase gene in *Phaeodactylum tricornerutum* affects carbon allocation. *N Biotechnol* **33**, 237–244 (2016).
- De Riso, V. *et al.* Gene silencing in the marine diatom *Phaeodactylum tricornerutum*. *Nucleic Acids Res.* **37**, e96 (2009).
- Yoon, K., Han, D., Li, Y., Sommerfeld, M. & Hu, Q. Phospholipid:diacylglycerol acyltransferase is a multifunctional enzyme involved in membrane lipid turnover and degradation while synthesizing triacylglycerol in the unicellular green microalga *Chlamydomonas reinhardtii*. *Plant Cell* **24**, 3708–3724 (2012).
- Rohr, J., Sarkar, N., Balenger, S., Jeong, B. R. & Cerutti, H. Tandem inverted repeat system for selection of effective transgenic RNAi strains in *Chlamydomonas*. *Plant J* **40**, 611–621 (2004).
- Boutros, M. & Ahringer, J. The art and design of genetic screens: RNA interference. *Nat Rev Genet* **9**, 554–566 (2008).
- Kilian, O., Benemann, C. S., Niyogi, K. K. & Vick, B. High-efficiency homologous recombination in the oil-producing alga *Nannochloropsis* sp. *P Natl Acad Sci USA* **108**, 21265–21269 (2011).
- Nelson, J. A. & Lefebvre, P. A. Targeted disruption of the *NIT8* gene in *Chlamydomonas reinhardtii*. *Mol. Cell. Biol.* **15**, 5762–5769 (1995).
- Daboussi, F. *et al.* Genome engineering empowers the diatom *Phaeodactylum tricornerutum* for biotechnology. *Nat Commun* **5**, 3831 (2014).
- Weyman, P. D. *et al.* Inactivation of *Phaeodactylum tricornerutum* urease gene using transcription activator-like effector nuclease-based targeted mutagenesis. *Plant Biotechnol J* **13**, 460–470 (2015).
- Sizova, I., Greiner, A., Awasthi, M., Kateriya, S. & Hegemann, P. Nuclear gene targeting in *Chlamydomonas* using engineered zinc-finger nucleases. *Plant J* **73**, 873–882 (2013).
- Kim, H. & Kim, J. S. A guide to genome engineering with programmable nucleases. *Nat Rev Genet* **15**, 321–334 (2014).
- Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
- Mojica, F. J. M., Juez, G. & Rodriguezvalera, F. Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified *Pst* I sites. *Mol Microbiol* **9**, 613–621 (1993).
- Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- Cho, S. W., Lee, J., Carroll, D., Kim, J. S. & Lee, J. Heritable gene knockout in *Caenorhabditis elegans* by direct injection of Cas9-sgRNA ribonucleoproteins. *Genetics* **195**, 1177–1180 (2013).
- Jiang, W., Bruggeman, A. J., Horken, K. M., Plucinak, T. M. & Weeks, D. P. Successful transient expression of Cas9 and single guide RNA genes in *Chlamydomonas reinhardtii*. *Eukaryot Cell* **13**, 1465–1469 (2014).
- Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J. S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**, 1012–1019 (2014).

31. Sung, Y. H. *et al.* Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. *Genome Res* **24**, 125–131 (2014).
32. Dutcher, S. K., Galloway, R. E., Barclay, W. R. & Poortinga, G. Tryptophan analog resistance mutations in *Chlamydomonas reinhardtii*. *Genetics* **131**, 593–607 (1992).
33. Palombella, A. L. & Dutcher, S. K. Identification of the gene encoding the tryptophan synthase beta-subunit from *Chlamydomonas reinhardtii*. *Plant Physiol.* **117**, 455–464 (1998).
34. Kirst, H., Garcia-Cerdan, J. G., Zurbriggen, A., Ruehle, T. & Melis, A. Truncated photosystem chlorophyll antenna size in the green microalga *Chlamydomonas reinhardtii* upon deletion of the *TLA3-CpSRP43* gene. *Plant Physiol.* **160**, 2251–2260 (2012).
35. Meinecke, L. *et al.* Chlorophyll-deficient mutants of *Chlamydomonas reinhardtii* that accumulate magnesium protoporphyrin IX. *Plant Mol Biol* **72**, 643–658 (2010).
36. Ahn, J. W., Yin, C. J., Liu, J. R. & Jeong, W. J. Cucumber mosaic virus 2b protein inhibits RNA silencing pathways in green alga *Chlamydomonas reinhardtii*. *Plant Cell Rep.* **29**, 967–975 (2010).
37. Last, R. L., Bissinger, P. H., Mahoney, D. J., Radwanski, E. R. & Fink, G. R. Tryptophan mutants in *Arabidopsis*: the consequences of duplicated tryptophan synthase beta genes. *Plant Cell* **3**, 345–358 (1991).
38. Lim, J. M. *et al.* Threonine 286 of fatty acid desaturase 7 is essential for omega-3 fatty acid desaturation in the green microalga *Chlamydomonas reinhardtii*. *Front Microbiol* **6**, 66 (2015).
39. Polle, J. E., Benemann, J. R., Tanaka, A. & Melis, A. Photosynthetic apparatus organization and function in the wild type and a chlorophyll *b*-less mutant of *Chlamydomonas reinhardtii*. Dependence on carbon source. *Planta* **211**, 335–344 (2000).
40. Polle, J. E. W., Kanakagiri, S., Jin, E., Masuda, T. & Melis, A. Truncated chlorophyll antenna size of the photosystems - a practical method to improve microalgal productivity and hydrogen production in mass culture. *Int J Hydrogen Energy* **27**, 1257–1264 (2002).
41. Bae, S., Park, J. & Kim, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).
42. Liang, X. *et al.* Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J Biotechnol* **208**, 44–53 (2015).
43. Malina, A. *et al.* Repurposing CRISPR/Cas9 for in situ functional assays. *Genes Dev.* **27**, 2602–2614 (2013).
44. Shen, B. *et al.* Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods* **11**, 399–402 (2014).
45. Koike-Yusa, H., Li, Y., Tan, E. P., Velasco-Herrera Mdel, C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol* **32**, 267–273 (2014).
46. Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211 (2010).
47. Wu, Y. *et al.* TALE nickase mediates high efficient targeted transgene integration at the human multi-copy ribosomal DNA locus. *Biochem. Biophys. Res. Commun.* **446**, 261–266 (2014).
48. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
49. Wang, Y. *et al.* Simultaneous editing of three homoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotechnol* **32**, 947–951 (2014).
50. Auer, T. O., Duroure, K., De Cian, A., Concorde, J. P. & Del Bene, F. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res.* **24**, 142–153 (2014).
51. He, X. *et al.* Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res* (2016).
52. Tuggle, C. K. & Waters, W. R. Tuberculosis-resistant transgenic cattle. *P Natl Acad Sci USA* **112**, 3854–3855 (2015).
53. Yamano, T., Iguchi, H. & Fukuzawa, H. Rapid transformation of *Chlamydomonas reinhardtii* without cell-wall removal. *J Biosci Bioeng* **115**, 691–694 (2013).
54. Schroda, M., Blocker, D. & Beck, C. F. The *HSP70A* promoter as a tool for the improved expression of transgenes in *Chlamydomonas*. *Plant J* **21**, 121–131 (2000).
55. Ikeda, K. & Miyasaka, H. Compilation of mRNA sequences surrounding the AUG translation initiation codon in the green alga *Chlamydomonas reinhardtii*. *Biosci. Biotechnol. Biochem.* **62**, 2457–2459 (1998).
56. Hopp, T. P. *et al.* A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification. *Bio-Technol* **6**, 1204–1210 (1988).
57. Ritchie, R. J. Consistent sets of spectrophotometric chlorophyll equations for acetone, methanol and ethanol solvents. *Photosynth Res* **89**, 27–41 (2006).
58. Cerutti, H., Johnson, A. M., Gillham, N. W. & Boynton, J. E. A eubacterial gene conferring spectinomycin resistance on *Chlamydomonas reinhardtii*: integration into the nuclear genome and gene expression. *Genetics* **145**, 97–110 (1997).
59. Jeong, B.-r., Wu-Scharf, D., Zhang, C. & Cerutti, H. Suppressors of transcriptional transgenic silencing in *Chlamydomonas* are sensitive to DNA-damaging agents and reactivate transposable elements. *Proc. Natl. Acad. Sci. USA.* **99**, 1076–1081 (2002).
60. Luo, R. *et al.* SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience* **1**, 18 (2012).

Acknowledgements

We honor and acknowledge deceased Professor Ji-Won Yang for his inspiration on this project and establishment of the Advanced Biomass R&D Center (ABC). We thank Dr. Heriberto Cerutti for providing the *MAA7* RNAi #33 strain. This work was supported by the Advanced Biomass R&D Center (ABC) of Global Frontier Project funded by the Ministry of Science, ICT, and Future Planning (ABC-2011-0031350, ABC-2010-0029728 and ABC-2011-0031343) and by a grant from Institute for Basic Science (IBS-R021-D1).

Author Contributions

H.-M.O., J.-S.K., W.-J.J., Y.K.C. and B.-r.J. designed the research; J.K. designed the sgRNAs of target genes; S.-E.S., J.-M.L., S.H.Y. and J.-Y.Y. contributed to the *in vitro* experiments; S.-E.S. performed the mutagenesis and phenotypic characterization of *MAA7* and *CpSRP43*; J.-M.L. and K.H. performed the mutagenesis and phenotypic characterization of *ChlM*; S.-E.S., H.G.K., E.K.K., N.K.K., S.J., S.K., W.-S.S. and B.L. participated in the complementation experiment of *MAA7*; H.S. and K.-J.K. analyzed the structure modeling of the TSB protein; S.-E.S., J.-M.L., W.-J.J. and B.-r.J. wrote the manuscript; H.-M.O., J.-S.K., K.-J.K. and Y.K.C. revised the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Shin, S.-E. *et al.* CRISPR/Cas9-induced knockout and knock-in mutations in *Chlamydomonas reinhardtii*. *Sci. Rep.* **6**, 27810; doi: 10.1038/srep27810 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>