A fast and reliable strategy to generate TALEN-mediated gene knockouts in the diatom *Phaeodactylum tricornutum*

M. Serif [1], B. Lepetit [1], K. Weißert, P.G. Kroth, C. Rio Bartulos *

Plant Ecophysiology, Fachbereich Biologie, Universität Konstanz, D-78457 Konstanz, Germany

A B S T R A C T

Reverse genetics techniques are powerful tools for studying gene functions. In the model diatom *Phaeodactylum tricornutum*, RNAi-mediated knockdown of genes still is the most commonly used reverse genetics technique. Due to the diploidic life cycle missing reproduction in lab cultures, many commonly used techniques to create knockout instead of knockdown lines are not applicable in *P. tricornutum*. These limitations can be overcome by using genome editing approaches like TALEN (Transcription activator-like effector nucleases), and/or CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats), allowing the introduction of targeted mutagenesis events. Both techniques have recently been adapted exemplarily for diatoms, however, no concise guidelines exist yet for routine utilization of these tools and the subsequent characterization of the mutants. We therefore have adapted a cost-effective TALEN generation system previously established for mammalian cells for the use in *P. tricornutum,* allowing the assembly of TALENs in about two weeks. We further provide protocols for: a) choosing a TALEN target site in order to avoid potentially ineffective and/or off-target prone TALEN constructs, b) efficient transformation of *P. tricornutum* with both TALEN constructs, utilizing two antibiotics resistance markers, c) effective screening of the transformants. In order to test our system we chose the blue-light dependent transcription factor Aureochrome 1a (*PtAureo1a*) as a target gene due to the known phenotype of previously characterized *P. tricornutum* RNAi knockdown strains. Our TALEN approach appears to be highly efficient: targeted mutation events were detected in 50% of all transformants obtained, whereas 21% of the transformants were found to be bi-allelic knockout lines. Furthermore, most TALEN transformed cell lines were found to be genetically homogeneous without the need for re-plating, which greatly facilitates the screening process.

1. Introduction

Diatoms are unicellular microalgae belonging to the Stramenopiles. They play an important role in global carbon fixation as well as for the nitrogen, phosphorous and silica cycles [1]. They are widespread in most aquatic habitats, where they need to cope with large variations of light quality and quantity [2,3]. As diatoms may contain larger amounts of lipids (up to 50% of dry weight), which even can be increased by genetic manipulation [4,5], they are suitable for the production of biodiesel and/or bioplastics [6–8]. The pennate *Phaeodactylum tricornutum* has become a model system for diatoms because of the availability of the genome sequence as well as genetic transformation techniques [9–14], allowing reverse genetics approaches. Because of their diplontic life cycle and the lack of of sexual reproduction in the lab, many methods for genetic manipulation like random mutagenesis or crossing are not available. Accordingly, knockdown via RNAi is

currently the most commonly used approach for genetic manipulations [15]. Recently, two new genetic tools for directed genome editing were developed, which allow induction of targeted DNA double-strand breaks to knock out genes irreversibly: the TALEN (Transcription activator-like effector nucleases) [16–18] and the CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats) systems [19–21]. Both approaches depend on nucleases that are guided to a specific DNA target sequence, and subsequently induce the formation of a DNA double strand break. The sequence-specific DNA binding of TALEN proteins is based on multiple 34 amino acid repeat modules, termed repeat variable di-residues (RVD), each binding specifically to one of the four nucleotides. All modules together (termed "targeting sequence") allow the recognition of a specific DNA sequence, so that in principle any DNA sequence of interest within the genome can be targeted [22,23]. The catalytic domain of the endonuclease *Fok*I, being fused to this targeting sequence, is only active as a dimer. Hence, for successful *Fok*I activity two TALEN proteins are required to bind in the right orientation and in close proximity onto the DNA double strand to induce double strand break formation. The mandatory binding of both TALENs strongly increases the targeting specificity [24]. The double strand break

* Corresponding author.
  E-mail address: riobartulos@gmail.com (C. Rio Bartulos).
[1] both authors contributed equally.

can be subsequently repaired by cellular mechanisms based either on homologous recombination (HR) or on non-homologous end joining (NHEJ). While NHEJ occurs during the whole cell cycle, HR is mainly restricted to the late S and G2 phase [25]. Both of these DNA repair mechanisms can be used to induce targeted mutations: HR can be exploited by introduction of foreign DNA with a strong homology to the DNA sequence surrounding the target site, which the cell uses as template to repair the double strand break [26]. If no HR template is used, the high error rate of NHEJ can be exploited to generate small random insertions/deletions until the target site is inaccessible for TALEN [27].

The successful application of TALEN and CRISPR/Cas9 in *P. tricornutum* has been published recently [28–30], however, a number of potential pitfalls have not yet been sufficiently addressed. Therefore, in this report, we describe how the TALEN approach can be optimized to obtain cell lines with targeted mutations at a high frequency and how to minimize the risk of potential off-target binding of TALENs. Additionally, we show that a thorough screening process is required for correctly distinguishing bi-allelic and mono-allelic knockout mutants, as well as to prove that a specific cell line is genetically homogeneous and does not contain different genotype lines. Although the CRISPR approach can be more easily adapted for a specific target site, we chose to improve the strategy of utilization of TALENs in diatoms because, based on research in other organisms [31], it offers the potential for reduced off-target effects. We developed a diatom-specific TALEN system by combining two previously published *P. tricornutum* transformation plasmids, pPha-T1 (Genebank ID: AF219942.1, [13]) and pPha-NR (Genebank ID: JN180663.1, [32]), and the TALEN plasmids developed by the Zhang lab for mammalian systems [33]. We adapted the Zhang protocol [33] for TALENs assembly and verification for diatoms, allowing the creation of TALENs in less than two weeks. We furthermore developed concise guidelines for all steps of this method, from target site design to screening mutated lines, which allows any molecular biology lab equipped with a particle gun or an electroporator to produce knockout mutants of *P. tricornutum*. In order to evaluate our TALEN system, we generated TALENs targeting the *PtAureo1a* gene, encoding the blue-light photoreceptor Aureochrome 1a (PtAUREO1a), because of the availability of a specific antiserum as well as the known phenotype of RNAi-silenced strains, such as lower chlorophyll *a* (Chl *a*) content per cell and a higher relative amount of xanthophyll cycle pigments [35].

2. Material & methods

2.1. Assembly of the TALEN plasmids

The plasmid kit used for building TALENs was a gift from Dr. Feng Zhang (Addgene, Cambridge, MA, USA; kit #1000000019) [33]. A *Hin*dIII restriction site was introduced upstream of the start codon of the TALEN plasmids using primer pair TALEN_HindIII + _for/rev (see Table SI). The *P. tricornutum* expression plasmids pPha-T1 (GenBank AF219942, [13]) and pPha-NR (GenBank JN180663, [32]) were modified using site-directed mutagenesis with the primer pairs PTV_BSAI1719SD_for/rev and PTV_BSAI2888SD_for/rev (see Table SI) to remove the two *Bsa*I restriction sites. Furthermore, the Zeocin resistance gene (*Sh ble*) from the pPha-NR vector was exchanged with the Nourseothricin resistance gene (*nat*) gene from the pNat vector [13], creating the pPha-NR-Nat vector. The different backbones of the TALEN expression plasmids from [33] (pTALEN_v2_NG (detects T), pTALEN_v2_NI (detects A), pTALEN_v2_NN (detects G or A), pTALEN_v2_HD (detects C)) were excised from the respective plasmids using a *Hin*dIII/*Sac*I double digest, and ligated into both pPha T1 and pPha-NR-Nat to create *P. tricornutum* specific TALEN plasmids pM9_fcpA_NG, pM9_fcpA_NI, pM9_fcpA_NN, pM9_fcpA_HD and pM9_NR_NG, pM9_NR_NI, pM9_NR_NN, pM9_NR_HD, respectively. The correct integration of the TALEN backbone was verified by Sanger sequencing (GATC, Konstanz, Germany). These plasmids specifically designed for producing TALENs

in *P. tricornutum* will be made available at Addgene (www.addgene. org).

We created TALEN pairs with a total of 20 repeat variable di-residues (RVD) each. Both the first RVD, which is always NG (targeting T), and the last RVD, which is only a half monomer, are already included in the respective plasmids. Based on these prerequisites, possible target sites for the *PtAureo1a* gene (JGI Protein ID 49116) were generated using the TAL Effector Nucleotide Targeter 2.0 [36] using a fixed repeat array length of 19 per TALEN (the first NG is not counted by the software), a spacer length of 15 to 22 bp between the two TALENs and the *P. tricornutum* Refseq ID (GCF_000150955.2) to predict possible off-target effects. Target sites were chosen according to the following parameters (summarized in Fig. 2): The target site should be an exon region in or upstream of the first functional domain, have no predicted off-targets and contain few NN repeat variable di-residues (RVDs), which lack specificity (targeting G/A) compared to the other RVDs. Lastly, cutting efficiency of good candidates was estimated using SAPTA (Scoring Algorithm for Predicting TALEN Activity; search type used: "Score individual TALEN pair(s), optimal spacer length is assumed in score calculation") [37].

Cloning and insertion of the target sequences into our TALEN plasmids were performed as described in [33]. In short, a RVD monomer library was constructed via PCR from the plasmids pHD_v2, pNG_v2, pNI_v2 and pNN_v2 (Addgene Kit # 1000000019, [33]). The monomers were then digested and ligated in a Golden Gate type reaction into hexamers, amplified via PCR and purified. Last, the three hexamers and the plasmid containing the backbone were digested and ligated in a second Golden Gate reaction (a summary of the TALEN assembly procedure can be found in Fig. 1). The resulting plasmids were sequenced (GATC, Konstanz, Germany) to verify correct integration and order of the 20 targeting TALE monomers.

2.2. Cultivation of algae

The *P. tricornutum* (Bohlin) strain UTEX646 was obtained from the culture collection of algae of the University of Texas (UTEX, Austin, USA). *P. tricornutum* was grown axenically in liquid F/2 medium without added silica and 16.5‰ salt content or on solid F/2 media which contained additionally 1.2% (*w/v*) Bacto Agar (BD, Sparks, MD, USA). Cells in liquid F/2 medium were cultivated in a 16 h/8 h light/dark cycle in Erlenmeyer flasks under continuous shaking at 20 °C and an illumination of 35 µmol photons m^{-2} s^{-1} (Osram Lumilux L58 W/840, Munich, Germany). Plated cultures were cultivated under continuous illumination at 75 µmol photons m^{-2} s^{-1} (Osram Biolux L30W/965).

2.3. Nuclear transformation of P. tricornutum

Nuclear transformation of *P. tricornutum* was performed using a Bio-Rad Biolistic PDS-1000/He Particle Delivery System (Bio-Rad, Hercules, CA, USA) fitted with 900/1100/1350 psi rupture disks as described previously [13,38,39]. 10^8 cells per plate were bombarded with 1.25 µg of each plasmid. For selective cultivation of *P. tricornutum* transformants, 75 mg mL^{-1} Zeocin (Invitrogen, Carlsbad, CA, USA) and 150 mg mL^{-1} Nourseothricin (ClonNat, Werner Bioagents, Jena, Germany) were added to the solid F/2 media [13,38].

2.4. DNA isolation

Genomic DNA was isolated using the nexttec™ 1step DNA isolation from tissues & cells kit (Biozym, Hessisch Oldendorf, Germany) according to the manufacturer's instructions. A cell pellet corresponding to 10 ml of culture in mid exponential growth phase was used as starting material. Incubation at 56 °C and 1200 rpm was done either 6 h or overnight. Concentration of genomic DNA was measured by Nanodrop 2000 UV/VIS Spectrometer (Thermo Fisher, Schwerte, Germany).
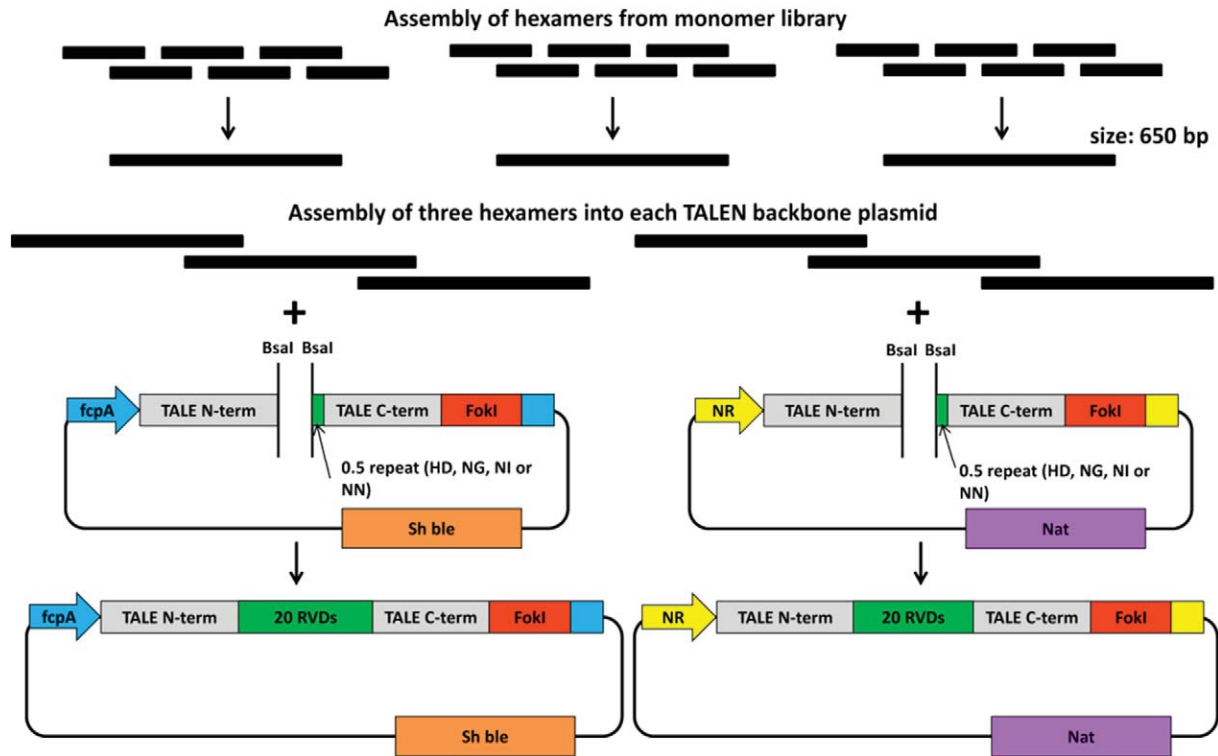
Fig. 1. Schematic overview of the TALEN assembly process (according to [33]) and the resulting plasmids. Six monomers are assembled into hexamers corresponding to the desired target site in a golden-gate type reaction from a library consisting of 72 monomers (18 positions of the TALEN target site and 4 different RVDs), allowing assembly of multiple fragments in the desired order in a single step. Three of these hexamers are then ligated in a second golden gate-type reaction into the plasmid containing the first RVD (always NG) and the last half RVD (HD, NG, NI or NN, depending on the target site). Correct insertion of the target sequence needs to be verified by colony PCR (insert size: 2.2 kbp). Clones indicated to be positive by colony PCR should be verified by restriction digest with *Afe*I (expected fragment lengths: 5 kbp, 2.2 kbp and 165 bp; exemplary shown in Fig. S1) as well as sequencing of the inserted fragment. fcpA: FcpA (= Lhcf1) promoter; *Fok*I: endonuclease; N-/C-term: N and C terminus, respectively; Nat: nourseothricin resistance cassette; NR: nitrate reductase promoter; RVD: repeat variable di-residue; Sh ble: Zeocin resistance cassette.

2.5. Allele-specific PCR

DNA sequences for the two alleles of the *PtAureo1a* gene (JGI ID: 49,116) were deduced from the whole genome shotgun sequencing (WGS) database (NCBI) by alignment of the individual sequence reads. Allele-specific primers for PCR were derived which include an allele-specific difference on the 3′ terminal base, thereby preventing polymerases without proofreading function from amplifying the respective other allele. PCR was performed using either Taq B polymerase (Biozym, Hessisch Oldenburg, Germany) or HiDi polymerase (myPOLS, Konstanz, Germany) according to the manufacturer's instructions. An extension time of 4 min and an annealing temperature of 52 °C (Taq B) or 55 °C (HiDi polymerase) were used to amplify the TALEN target site of both *PtAureo1a* alleles from isolated genomic DNA using primers Aureo1a_for and Aureo1a_rev (see Table S1). PCR products were separated on 1% agarose gels and PCR products were isolated using the Geneclean Turbo Kit (MP Biomedicals, Eschwege, Germany) according to the manufacturer's instructions. Purified DNA was analyzed by Sanger sequencing using primer Aureo1a_for (see Table SI) (GATC, Konstanz, Germany, or Source Bioscience, Berlin, Germany). If sequencing results indicated mixed populations due to small insertions and/or deletions in one or both alleles, the PCR product was sub-cloned using the pGEM-T system (Promega, Mannheim, Germany).

2.6. Southern blotting

Isolated genomic DNA was digested using each of the following restriction enzymes overnight according to the manufacturer's instructions: *Bam*HI, *Bsr*GI, *Hin*dIII (Thermo Fisher, Schwerte, Germany). Samples of 400 ng digested DNA were separated on 0.8% agarose gels.

The agarose gel was incubated for 10 min in denaturation solution (0.5 M NaOH, 1 M NaCl), followed by 10 min incubation in neutralization solution (0.5 M Tris-HCl, 3 M NaCl, pH 7.5). A dry blot was then performed overnight: The gel was placed downside-up onto an acrylic glass plate and the positively charged nylon membrane (Roche, Mannheim, Germany; 11471240001), three layers of Whatman paper (3MM Chr, 3030917, VWR, Darmstadt, Germany) and absorbent paper were placed on top. The blotting setup was then weighed down. The 400 bp DIG-labeled probe was synthesized using the PCR DIG Probe Synthesis Kit (Roche, 11636090910) using an *PtAureo*1a-containing plasmid [35] as template and primers Aureo1a_probe_for/rev (see Table SI). Hybridization occurred overnight using DIG Easy Hyb (Roche, 11603558001) at 50 °C. Post-hybridization steps were performed using the DIG Block and Wash Buffer set (Roche, 11585762001), but with a shortened 68 °C washing step (2x10min). The Anti-DIG-AP antibody (Roche, 11093274910) was used at a 1:20.000 fold dilution; the alkaline phosphatase substrate used was CDP-Star (Roche, 12041677001). The blots were developed using X-ray films (Amersham Hyperfilm ECL, GE Healthcare, Munich, Germany) after 15 to 60 min incubations by a Konica SRX-201 Developer.

2.7. Protein isolation and immunoblotting

For protein extraction, cell pellets were resuspended in lysis buffer (4 M urea, 1.5 M thiourea, 1% SDS, 20 mM Tris pH 8) supplemented with protease inhibitor (Complete™ EDTA-free, Roche) according to the manufacturer's instructions. A spatula tip of 1 mm, 0.5 mm and 0.1 mm diameter beads was added and the cells were lysed in a Savant FastPrep FP120 bead mill (Thermo Scientific, Karlsruhe, Germany) six times for 20 s with cooling on ice for 1 min between each cycle. Cell

debris and residual beads were removed by two centrifugation steps at 18000 g and 4 °C for 30 min.

Proteins were separated in a 10% polyacrylamide gel by SDS-PAGE according to Laemmli et al. [40]. Each lane was loaded with protein extract of either wild type or mutant cell lines. After blotting, the nitrocellulose membrane (Amersham Protran 0.1 μm NC, GE Healthcare) was cut between 35 and 40 kDa and the top half (40–250 kDa) was used to detect PtAUREO1a, whereas the bottom half (0–35 kDa) was used to detect the D1 loading control. Immunoblots using a custom-made antiserum specific against PtAUREO1a (Agrisera AB, Vännas, Sweden) were performed as described in [35], whereas the D1-specific antiserum (AS05-084, Agrisera AB) was used according to the manufacturer's instructions. Blots were developed using an Odyssey FC Imaging System (Li-Cor, Bad Homburg, Germany).

2.8. Chlorophyll a determination

Chlorophyll *a* (Chl *a*) was isolated using 10% methanol and 90% acetone successively and its concentration was determined spectrophotometrically using the formula of Jeffrey and Humphrey [41].

2.9. Pigment extraction

Extraction of pigments and subsequent analyses via HPLC were done as described in [42]. Samples were analyzed on a calibrated Hitachi LaChrom Elite HPLC system equipped with a Nucleosil 120-5 C18 column (Macherey-Nagel, Düren, Germany).

2.10. Measurement of non-photochemical quenching (NPQ)

Cell suspensions in mid-exponential phase were adjusted to a Chl *a* content of 1 μg ml^{-1} and NPQ was measured with an AquaPen-C AP 100 (Photon Systems Instruments, Brno, Czech Republic) using light pulses with an intensity of 2100 μmol photons m^{-2} s^{-1} applied every 20 s to induce maximal fluorescence and 700 μmol photons m^{-2} s^{-1} of actinic light to induce NPQ.

2.11. Determination of cell size by microscopy

Exponentially growing cells were analyzed using an Olympus BX51 epifluorescence microscope equipped with a Zeiss AxioCam MRm digital camera system (Carl Zeiss Microscopy GmbH, Göttingen, Germany). Length and width of 50 cells were determined for each cell line and the cell volume was approximated as described in [43] using the formula $V = (\pi/12) * d^2 * h$ (d: diameter of the cell; h: length of the cell).

3. Results & discussion

3.1. Generation of the TALEN constructs

In order to perform a cost-effective and easy assembly of the TALEN targeting sequence, we chose a system developed previously for mammalian systems, allowing the complete assembly and sequence verification of individual TALEN plasmids within two weeks [33]. In contrast to previous publications on *P. tricornutum* [28,30] describing the expression of TALENs in a single transformation plasmid, we decided to clone the two TALEN backbones into two separate plasmids, one with a constitutive fcpA promoter and *sh ble* gene conferring resistance to Zeocin (pPha-T1), and one with an inducible nitrate reductase promoter [44] and a *nat* gene conferring resistance to Nourseothricin (pPha-NR-Nat). This design (shown schematically in Fig. 1) has several advantages: the expression of both TALENs from two plasmids, as compared to a single plasmid, reduces the size of the plasmids (from 13 to 14 to 7–8 kb), which facilitates cloning procedures and, based on results with other systems [45,46], may increase the transformation efficiency. Furthermore, introducing two plasmids at the same time usually results

in high rates of co-transformed *P. tricornutum* cells even without a second selection marker [14]. However, when using two different antibiotic resistances on the two plasmids encoding the individual TALENS, a higher selection stringency is achieved by screening for strains that have integrated both plasmids. Larger plasmids instead are more prone to random DNA double strand breaks induced by the tungsten particles used for the particle bombardment [47]. Additionally, the inducible promoter system allows switching off the expression the TALENs once favorable mutations have been demonstrated, which decreases the probability of off-target DNA modifications, and allows checking for lethal mutations by inducing the expression of the TALENs (by exchanging ammonia by nitrate) only after the initial round of antibiotics selection.

The recommended workflow for designing and generating TALEN constructs is presented in Fig. 2. In order to avoid allele-specific polymorphisms in the TALEN target sites, which could hinder correct binding, allele-specific gene sequences were deduced from the *P. tricornutum* whole genome shotgun sequences (WGS), thereby identifying 11 different allele-specific polymorphisms in the *PtAureo1a* gene for strain Pt1 (CCAP 1055/1). In the next step, a combination of two online tools was used to choose the best potential target sites. TALE-NT 2.0 [36] was used to predict target sites in the gene of interest and potential off-targets for each TALEN pair based on the *P. tricornutum* RefSeq sequence (GCF_000150955.2). A total of 191 potential TALEN pairs were suggested within the first 700 bp of the *PtAureo1a* gene, of which only 29 were targeting an exon region and had no predicted potential off-targets. These TALEN pairs were then sorted by the frequency of the less specific NN RVD, and the in vitro cutting efficiency was estimated by SAPTA (Scoring Algorithm for Predicting TALE(N) Activity) [37]. According to the SAPTA guidelines, composite scores above 30 are recommended for a high rate of gene modifications. Additionally, the scores of the individual TALENs should have similar values (e.g., 20 and 25 are better than 5 and 40, although resulting in a similar composite score). A SAPTA analysis of the TALEN pair chosen for *PtAureo1a* (left target site: TCCCTCCTTAAGGAAGAGAA; right target site: TCGCCCAAGTGCGA ACGAAT; spacer length: 19 bp) resulted in a composite score of 43.13 and scores for the individual TALENs of 25.44 and 20.15, respectively.

The nucleases of the TALEN pair cleave at position 679, which is part of the leucine zipper domain of PtAUREO1a (see Fig. 3, the TALENs are symbolized by a grey line and scissors symbols, the predicted cutting site by a dashed line). Thus, random mutations in this area could abolish DNA binding or may result in premature termination of protein translation. TALENs were assembled as described in Materials and Methods and verified by sequencing. We have applied the prediction tools to previously published TALEN target sites of *P. tricornutum* [28,30], which had been constructed before these tools became available, and found potential off-targets for most of them. While no direct evidence is available whether these potential off-targets are actually targeted by the TALENs in vivo, the search tool used here is much better suited than simple BLAST searches due to the complexity of the target sequence containing bi-specific RVDs and a gap of variable length.

3.2. Screening of the obtained transformants and statistical evaluation

A total of 24 *P. tricornutum* colonies were obtained after transformation. While several allele-specific differences were identified (based on the sequenced *P. tricornutum* strain Pt1 (CCAP 1055/1)), the two alleles could not be amplified separately from Pt4 (UTEX 646), the strain used in this study. As the availability of the genome sequence greatly facilitates both TALEN design and screening, the use of wild type strain Pt1 is recommended for creation of *P. tricornutum* knockout mutants, unless previous work may require the use of another strain. We could solve the problem by identifying a mixed trace peak (T/G) at position 81 of the *PtAureo1a* gene in the Pt4 wild type cells as well as in the mutants, which allowed distinguishing both alleles. Using PCR, several small deletions (e.g., clone 6: 29 bp deletion in allele 1) or insertions (e.g., clone
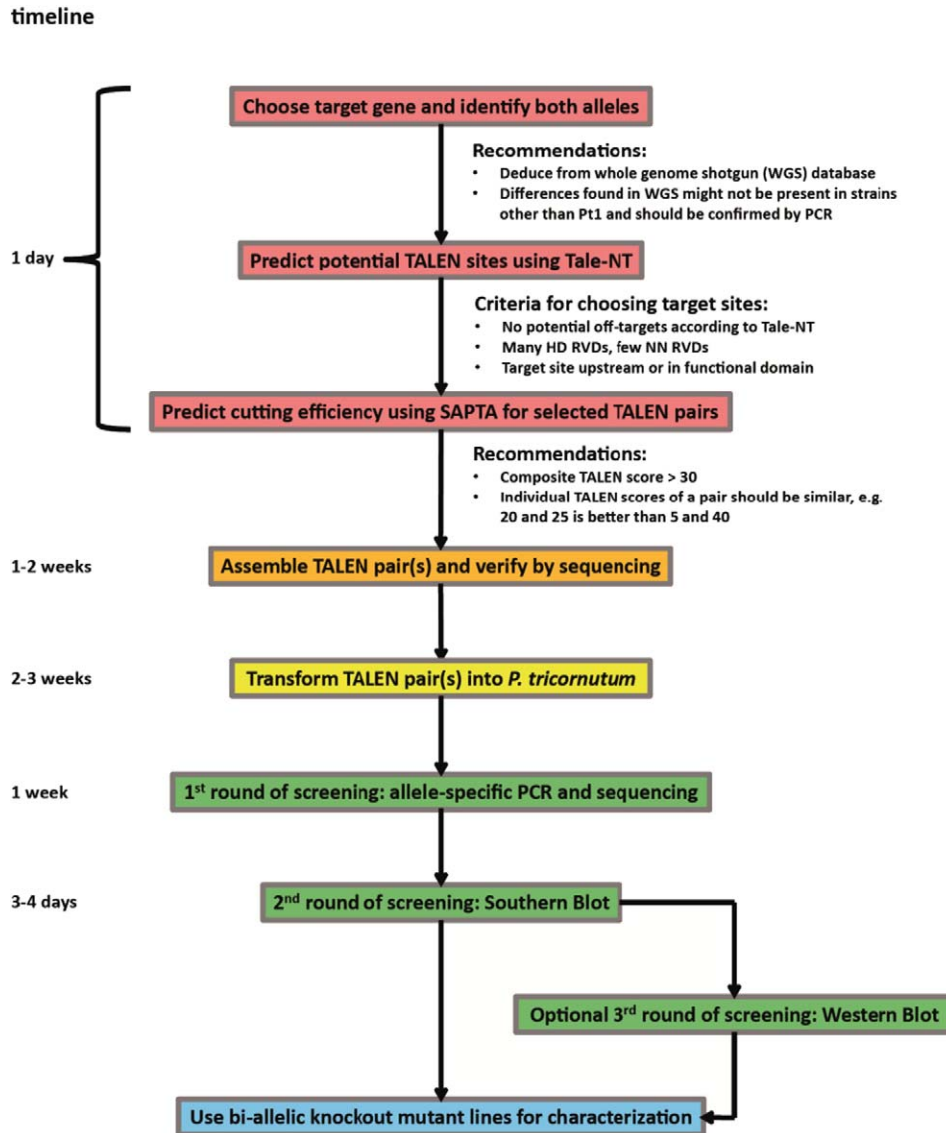
timeline

Fig. 2. Workflow and recommendations for design and assembly of TALEN constructs, as well as screening of the obtained transformants. An estimated timeframe is given for each step.

14: 143 bp insertion in allele 2) were detected in these transformants, which are summarized in Table 1. No PCR products could be obtained for cell lines 8 and 9, presumably due to large insertions or deletions at the target site in both alleles. Furthermore, in eight transformant cell lines only one allele could be amplified by PCR, indicating mutation events in the respective other alleles. In case of clone 11, sub-cloning of the PCR fragment was necessary to demonstrate small deletions in each allele (7 and 18 bp, respectively). In summary, 10 clones were shown by PCR and sequencing to have no targeted mutagenesis events, whereas

five clones (5, 6, 14, 16 and 21) were found to be at least mono-allelic mutants and clone 11 was found to be a bi-allelic mutant. As not all clones could be reliably screened by PCR, we introduced a second screening process by employing Southern Blots of genomic DNA. The restriction enzymes *Bam*HI, *Bsr*GI and *Hind*III were chosen for digestion of the genomic DNA prepared from each of the 24 mutants. For each of the chosen restriction enzymes, Southern Blots of genomic DNA of wild type strains showed a single band in the range of 3 to 5 kbp (see Fig. 4). In contrast, ten of the 24 transformants (1, 2, 3, 6, 8, 9, 13, 14,
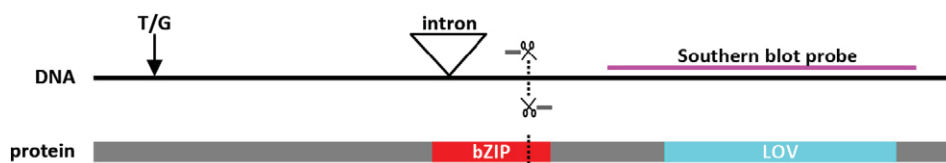

Fig. 3. Schematic drawing of *PtAureo1a* and its gene product PtAUREO1a including conserved domains (bZIP in red and LOV in blue). The TALEN recognition sites are indicated by a grey line and the *Fok*I endonuclease domains are symbolized by scissors. The predicted cutting site of the TALEN pair within the gene and its relative location within the gene product and its conserved domains is indicated by dashed lines. The allele-specific difference of *PtAureo1a* at position 81 (T/G) in strain Pt4 is indicated by an arrow, an 102 bp intron region by an inverted triangle shape and the binding site of the Southern blot probe by a magenta-colored line.

Table 1

Evaluation of the screening of the two alleles via PCR and Southern Blot. Detected mutagenic events are indicated in bold.

Clone	Allele 1 (PCR)	Allele 2 (PCR)	Southern Blot[a]	Conclusion
1	No change	**No amplification**	**1**	**Mono-allelic mutation**
2	No change	No change	1	**Mixed culture; mono-allelic mutation**
3	No change	**No amplification**	**1**	**Mono-allelic mutation**
4	No change	No change	0	Both alleles unchanged
5	No change	**2 bp deletion**	0	**Mono-allelic mutation**
6	**29 bp deletion**	**No amplification**	**1**	**Bi-allelic mutation**
7	No change	No change	0	Both alleles unchanged
8	**No amplification**	**No amplification**	**2**	**Bi-allelic mutation**
9	**No amplification**	**No amplification**	**2**	**Bi-allelic mutation**
10	No change	No change	0	Both alleles unchanged
11	**7 bp deletion**	**18 bp deletion**	0	**Bi-allelic mutation**
12	No change	No change	0	Both alleles unchanged
13	**No amplification**	No change	**1**	**Mono-allelic mutation**
14	No change	**143 bp insertion**	**1**	**Mono-allelic mutation**
15	No change	No change	0	Both alleles unchanged
16	**No amplification**	**3 bp insertion**	**1**	**Bi-allelic mutation**
17	No change	No change	0	Both alleles unchanged
18	**No amplification**	No change	**1**	**Mono-allelic mutation**
19	No change	No change	0	Both alleles unchanged
20	No change	No change	0	Both alleles unchanged
21	**692C → A**[b]	No change	0	**Mono-allelic mutation**
22	No change	**No amplification**	0	**Mono-allelic mutation?**
23	No change	**No amplification**	0	**Mono-allelic mutation?**
24	No change	No change	0	Both alleles unchanged

[a] Southern Blot results were grouped into three categories:
  0: indistinguishable from the wild type.
  1: one band is at the height of the wildtype band, the other is shifted due to large insertion or deletion events.
  2: there is no band at the height of the wildtype due to insertion or deletion events in both alleles.
[b] a point mutation in clone 21 at position 692 downstream of the translation start is indicated.

16 and 18) showed distinct shifts of up to a few kbp (marked with an arrow). Out of these ten, eight strains (1, 2, 3, 6, 13, 14, 16 and 18) retained one band at the size of the wild type fragment, indicating that these are either heterozygous mutants or that small insertion or deletion events occurred in the second allele, which are not visible in Southern Blots, but which can be detected by PCR and sequencing (strain 16, see above). While the shifted band of strain 14 seems to indicate a shorter fragment, the 143 bp insertion detected by PCR (see above) introduced a new HindIII site, thereby causing a shorter

fragment size. Two cell lines (8 and 9) did not show a band corresponding to the wild type fragment size.

An overview of the screening results for all mutants is given in Table 1. We conclude that 50% of all mutants had at least mono-allelic mutations induced by TALEN, whereas clones 6, 8, 9, 11 and 16 are bi-allelic mutant cell lines (20.83% of all mutants obtained). As only the combination of sequencing of PCR-amplified target genes and Southern Blots using a target gene specific probe allowed us to identify all mutated strains, we highly recommend screening TALEN-transformants using
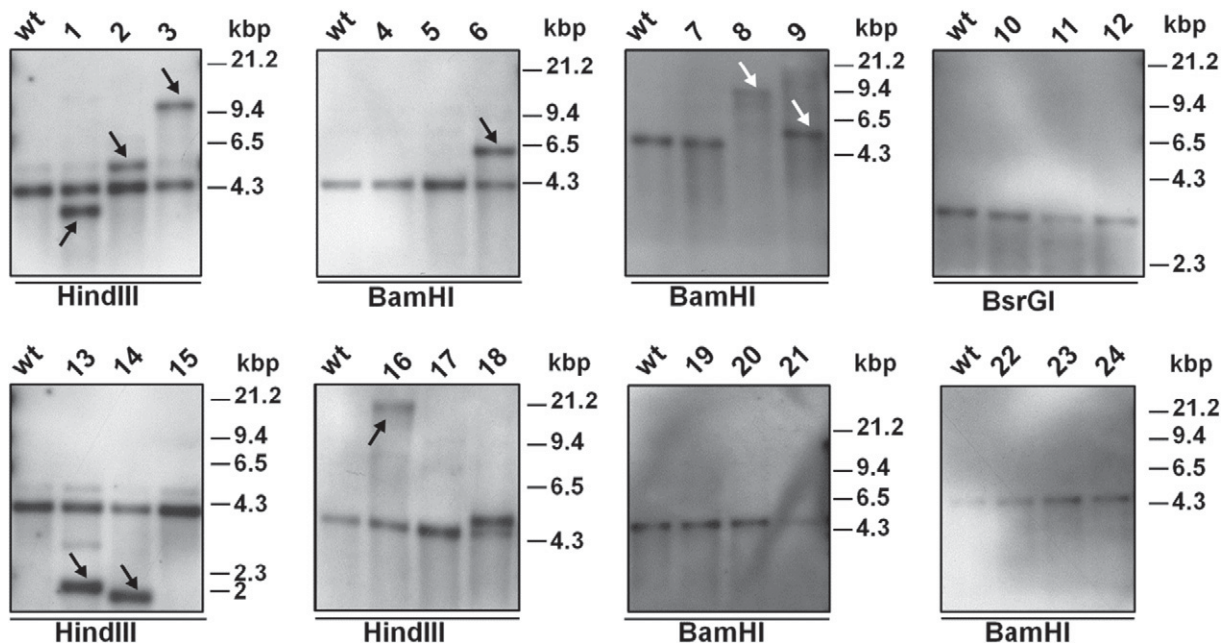

Fig. 4. Southern Blots using digested genomic DNA of either wild type or transformed cell lines. Shifts of the DNA fragments can be caused either by insertion or deletion events. Shifted DNA bands due to insertions or deletions are marked with arrows.

both methods whenever possible. The workflow for screening the obtained mutants is presented in Fig. 2.

3.3. Homogeneity of the generated mutants

As the average time required by the TALEN proteins to conduct a successful mutation in *P. tricornutum* is unknown, we have tested the homogeneity of the obtained cell lines with regard to double mutations. One mono-allelic knockout line (14) and four bi-allelic knockout lines (6, 8, 9 and 11) were spread onto individual agar plates, and three colonies of each line were re-isolated from single cell colonies. One re-isolated colony of cell line 9 did not grow properly in liquid medium, and thus was withdrawn from further analysis. Southern Blots and allele-specific PCRs were repeated in order to prove that the re-isolated clones were genetically identical regarding the mutated target site. Sequencing of PCR products showed no differences between re-isolated clones derived from cell lines 6, 11 and 14 and the sequences obtained prior to re-isolation. No PCR products could be obtained for re-isolated colonies of knockout lines 8 and 9 due to large insertion events. Southern Blots also showed identical band patterns for each mutant derived from the same clone, indicating homogeneity of the cell lines (see Fig. 5).

As a final demonstration that both alleles had been effectively knocked-out, we studied the respective protein levels of the four bi-allelic lines and the mono-allelic knockout line by Western Blots using PtAUREO1a-specific antiserum [35]. An antiserum labeling the plastid-encoded D1 protein of photosystem II was used as a loading control. Western Blots using extracts of wild type cells or from the mono-allelic knockout mutant showed a single band, which corresponds to the molecular weight of 41.5 kDa of PtAUREO1a (see Fig. 6). Knockout mutants derived from cell lines 6, 8 and 9 did not show any detectable bands at the expected size. All strains derived from cell line 11, however, showed a slightly truncated PtAUREO1a protein matching to the 18 bp deletion (corresponding to a 6 amino acid deletion) observed by PCR in one of the alleles. Thus, clones 6–1, 8–1, 9–1 and 11–1 were chosen to characterize the phenotype of PtAUREO1a knockout mutants.

Interestingly, the mono-allelic knockout line 14 no longer showed a reduced level of PtAUREO1a protein, as it had done prior to re-isolation (data not shown), although the TALEN induced mutations were still present, thus excluding allele repair/recombination (compare Fig. 4 and Fig. 5). We therefore conclude that *PtAureo1a* is a haplosufficient gene, i.e., one functional gene copy is sufficient to produce wild type-like protein levels when given some time for adaptation. Instability of knockdown strains has already been observed for RNAi-based silencing lines in *P. tricornutum* [48] including the previously studied PtAUREO1a RNAi knockdown strain (data not shown). Thus, if knockdown cell lines are required, a variation of the amount of the targeted protein may
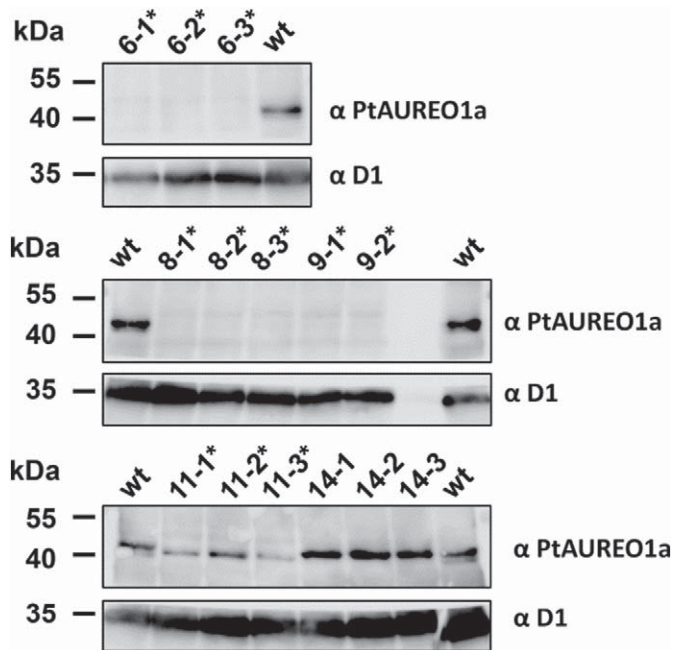

Fig. 6. Western Blot of wild type and two mono-allelic as well as four bi-allelic cell lines (marked with a *) after re-isolation using the PtAUREO1a antiserum. The expected molecular weight of PtAUREO is 41.5 kDa. A D1-specific antiserum was used as a loading control.

occur over time and the TALEN-induced mono-allelic knockout strains should be tested regularly.

The generation of homogeneous bi-allelic knockout mutants without re-isolation is in strong contrast to the previously published diatom TALEN systems, which did not yield any homogeneous clones without replating [28,30]. Homogeneous bi-allelic knockout mutations require the TALEN proteins to induce the mutagenesis events in both alleles prior to the first cell division on selective media. We routinely replated transformed cells after 24 h, hence this time span seems to be sufficient for the TALENs to operate. The reason why previously published systems [28,30] did not obtain genetically homogeneous transformants, despite replating the cells onto selection media after 48 h, is unclear. One explanation could be the carefully chosen target site based on the newly available prediction techniques, which avoid off-targets, and thus minimizes binding events to unintended target sites, and allows for selection of TALEN pairs with a high cutting efficiency. Additionally, it was shown that in human cells DNA repair via HR can take up to 7 h to complete and only occurs in late S and G2 phase, whereas NHEJ works much faster [25,49]. Thus, when using an HR-based system as in [30],
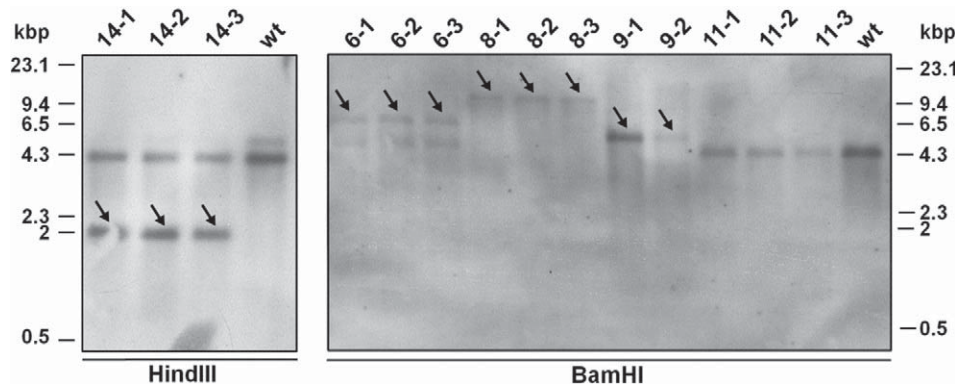

Fig. 5. Southern Blot using digested genomic DNA of either wild type or transformed cell lines. Shifts of the DNA fragments can be caused either by insertion or deletion events. Shifted DNA bands due to insertions and deletions are marked with arrows.

Table 2

Analysis of pigment composition and non-photochemical quenching (NPQ) capacity of wild type (wt) and the four bi-allelic PtAUREO1a knockout strains (6, 8, 9 and 11). Chlorophyll a (Chl a) was determined photometrically (wild type $n = 5$, bi-allelic mutants $n = 3$), whereas the other pigments were assayed by HPLC ($n = 2$). Pigments were isolated from strains being in mid exponential growth phase and having comparable Chl a per culture-volume ratios. NPQ was measured by PAM fluorometry ($n = 3$). Statistical significant differences compared to the wild type were calculated using Student's t-test (*: $p < 0.01$; **:$p < 0.001$). XC pool: xanthophyll cycle pool (diadinoxanthin + diatoxanthin).

	wt	6	8	9	11
Chl a [pg cell^{-1}]	0.54 ± 0.02	0.32 ± 0.04**	0.33 ± 0.04**	0.34 ± 0.07*	0.33 ± 0.02**
XC pool [mmol (mol Chl a)$^{-1}$]	65.54 ± 0.06	94.90 ± 1.33**	87.36 ± 0.31**	90.40 ± 1.58**	82.33 ± 0.44**
Chl c [mmol (mol Chl a)$^{-1}$]	143.33 ± 0.14	130.54 ± 1.27**	130.79 ± 1.88**	119.45 ± 0.55**	138.70 ± 0.53**
Fucoxanthin [mmol (mol Chl a)$^{-1}$]	566.24 ± 0.56	539.31 ± 0.38**	538.98 ± 2.89*	517.63 ± 1.88**	555.38 ± 0.10*
β-carotene [mmol (mol Chl a)$^{-1}$]	53.95 ± 0.18	57.00 ± 0.18*	54.66 ± 0.14	54.25 ± 0.66	53.97 ± 0.01
Non-photochemical quenching	0.476 ± 0.026	0.306 ± 0.005**	0.261 ± 0.001**	0.299 ± 0.001**	0.456 ± 0.017

more time prior to selection might be required in order to directly obtain genetically homogeneous mutants. Nonetheless, we highly recommend re-isolation of knockout strains after the screening process, which requires much less effort than re-isolation of clones prior to the screening process as described in previous studies, in order to avoid the potential risk of mixed cultures.

While CRISPR/Cas9 can be adapted to a specific target site much faster, the use of a TALEN based knockout system has other pronounced benefits. The targeting region of the guideRNA of CRISPR is only 20 bp in length and has been shown to tolerate several mismatched base pairs [50–52]. TALEN on the other hand is much more robust due to the longer targeting site of two individual TALEN monomers separated by a 12-21 bp spacer sequence to allow double strand break formation [53]. Additionally, TALEN seems less prone to mismatched base pairing indicated by a reduced off-target frequency relative to targeted mutagenesis events compared to CRISPR using a randomized DNA library [54,55], as has been summarized in [56].

3.4. Phenotypic characterization of PtAUREO1a knockout mutants

Previously characterized PtAUREO1a knockdown strains generated via RNAi showed a significantly reduced chlorophyll a (Chl a) content per cell compared to the wild type strain under medium light conditions (Qphar 30; 30 µmol photons m^{-2} s^{-1} absorbed photosynthetic radiation) [35]. Hence, we determined the Chl a content per cell and performed a pigment composition analysis via HPLC (see Table 2). Similar to the respective silencing strains, the four bi-allelic PtAUREO1a knockout mutants showed a reduction in Chl a content of about 30% compared to the wild type ($p < 0.005$, Student's t-test) and an increase of the xanthophyll cycle (XC, diadinoxanthin + diatoxanthin) pool size by >30% in clones 6, 8 and 9 and by 25% in clone 11 ($p < 0.001$). Additionally, the ratio of Chl c and fucoxanthin relative to Chl a was found to be decreased by 10–17% and 5–10%, respectively ($p < 0.001$).

Interestingly, the NPQ capacity was reduced by about 40% in the three knockout lines 6, 8 and 9 despite the increased amount of XC pigments, whereas cell line 11 expressing a truncated PtAUREO1a showed wild type NPQ levels (See Table 2; Student's t-test, $p < 0.001$). NPQ in diatoms depends not only on XC pigments, but also on the expression of the different Lhcx proteins [57,58]. Hence, in the PtAUREO1a knockout mutants the amount of Lhcx proteins may be reduced. The reduced NPQ in the PtAUREO1a knockouts is in line with previous results, showing that blue light induces higher NPQ values in the wild type than red light [59], and hence stresses the importance of blue light photoreceptors for generating NPQ capacity. However, in PtAUREO1a silencing strains, an increased NPQ capacity had been observed, which was unexpected and the opposite effect of what we could observe in the knockout lines [35]. These results highlight that in certain cases stable knockout lines may not only amplify phenotypic effects compared to silencing lines, but may also provide completely different results. Poor correlation between knockdown lines and knockout lines has already been observed in several other model organisms [60–62].

Despite the involvement of PtAUREO1a in the cell cycle [63], no changes of growth rates in the bi-allelic knockout lines were found at the described cultivation conditions, indicating that a loss of PtAUREO1a can be compensated by the cell. However, we could detect a difference in the cell morphology of P. tricornutum. Length and width of 50 cells per cell line were measured microscopically and the cell volume was estimated by assuming P. tricornutum cell shape can be regarded as two joined cones as described in [43] (cell width, length and volume of each measured cell is shown in Table S2, an example for each cell line is shown in Fig. S2). While cell width was found to be highly similar in all strains, cell length differed significantly between wild type and knockout cell lines (see Fig. 7A/B). Mono-allelic knockout line 14 as well as clone 11 showed a 10% decrease in cell length, whereas bi-allelic knockout lines 6, 8 and 9 showed a 25% decrease in cell length (One-way ANOVA followed by Tukey's HSD, $p < 0.01$). When
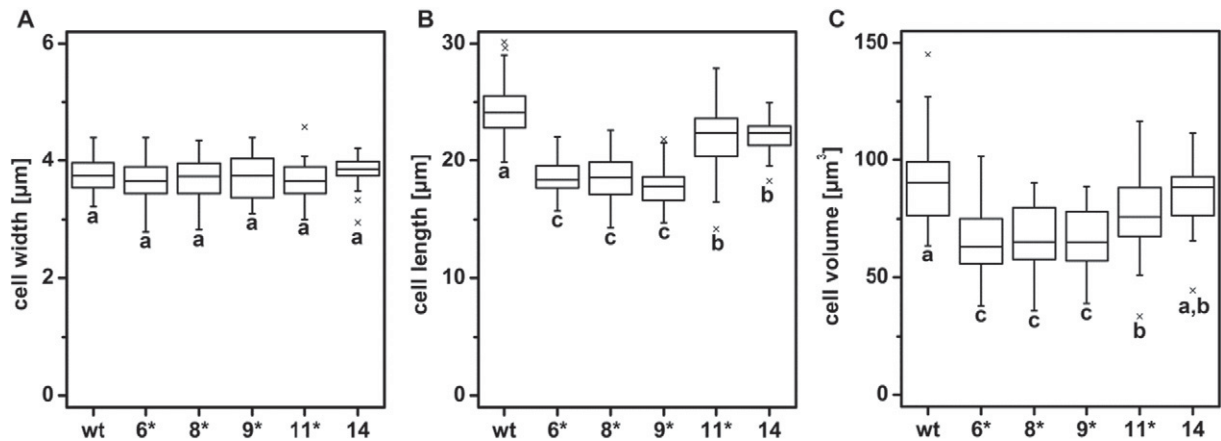

Fig. 7. 50 cells of the wildtype and each mutant strain (6, 8, 9, 11, 14) were examined microscopically and cell width (A) and length (B) was measured. Cell volume was approximated as two cone shapes (C). The bi-allelic knockout mutants are marked with a *. Whiskers represent the outermost data point within the 1.5-fold of the interquartile range; outliers are represented by an x. Different letters represent statistical significant differences in cell width, length or volume between strains (One-way ANOVA, followed by Tukey's HSD, $p < 0.01$).

approximating the volume of fusiform *P. tricornutum*, only the bi-allelic knockout lines showed significant decreases in cell volume compared to wild type cells. While cell volume was reduced by >20% in cell lines 6, 8 and 9, the reduction in cell line 11 was only 13% (see Fig. 7C, $p < 0.01$). Such a reduced cell length and volume had not been observed for the previously characterized PtAUREO1a silencing strains [35], confirming that knockout approaches may yield much more obvious phenotypes than RNAi-based silencing. We also have to note that, with respect to XC-pool content and cell volume, the three bi-allelic knockout lines 6, 8 and 9 showed a very similar phenotype compared to the wild type, while the bi-allelic knockout line 11 expressing a truncated PtAUREO1a protein showed a weaker phenotype (see Table 3 and Fig. 7). Hence, the truncated PtAUREO1a protein in strain 11 seems to still confer some activity.

4. Conclusion

An efficient TALEN system has been established for the diatom *P. tricornutum*, which leads to a high frequency of targeted mutation events and yields mainly genetically homogeneous cell lines. Additionally, all important steps, from target site design, TALEN construction using an easy-to-follow and publically available modular construction system, up to screening of the mutants, have been addressed to avoid the potential pitfalls for knockout generation in *P. tricornutum*, like potential off-targets and inefficient TALEN proteins. The TALEN system presented here may expand the application of reverse genetics approaches for creation of *P. tricornutum* knockout mutants to a broader scientific community.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.algal.2017.02.005.

References

[1] A. Yool, T. Tyrrell, Role of diatoms in regulating the ocean's silicon cycle, Glob. Biogeochem. Cycles 17 (2003) 1103–1121.

[2] H.L. MacIntyre, T.M. Kana, R.J. Geider, The effect of water motion on short-term rates of photosynthesis by marine phytoplankton, Trends Plant Sci. 5 (2000) 12–17.

[3] M. Ragni, M.R. D'Alcalà, Light as an information carrier underwater, J. Plankton Res. 26 (2004) 433–443.

[4] O. Levitan, J. Dinamarca, E. Zelzion, D.S. Lun, L.T. Guerra, M.K. Kim, J. Kim, B.A. Van Mooy, D. Bhattacharya, P.G. Falkowski, Remodeling of intermediate metabolism in the diatom *Phaeodactylum tricornutum* under nitrogen stress, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 412–417.

[5] E.M. Trentacoste, R.P. Shrestha, S.R. Smith, C. Gle, A.C. Hartmann, M. Hildebrand, W.H. Gerwick, Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 19748–19753.

[6] Y. Chisti, Biodiesel from microalgae, Biotechnol. Adv. 25 (2007) 294–306.

[7] T.M. Mata, A.A. Martins, N.S. Caetano, Microalgae for biodiesel production and other applications: a review, Renew. Sust. Energ. Rev. 14 (2010) 217–232.

[8] P. Roesle, F. Stempfle, S.K. Hess, J. Zimmerer, C.R. Bartulos, B. Lepetit, A. Eckert, P.G. Kroth, S. Mecking, Synthetic polyester from algae oil, Angew. Chem. Int. Ed. 53 (2014) 6800–6804.

[9] C. Bowler, A.E. Allen, J.H. Badger, J. Grimwood, K. Jabbari, A. Kuo, U. Maheswari, C. Martens, F. Maumus, R.P. Otillar, E. Rayko, A. Salamov, K. Vandepoele, B. Beszteri, A. Gruber, M. Heijde, M. Katinka, T. Mock, K. Valentin, F. Verret, J.A. Berges, C. Brownlee, J.P. Cadoret, A. Chiovitti, C.J. Choi, S. Coesel, A. De Martino, J.C. Detter, C. Durkin, A. Falciatore, J. Fournet, M. Haruta, M.J. Huysman, B.D. Jenkins, K. Jiroutova, R.E. Jorgensen, Y. Joubert, A. Kaplan, N. Kroger, P.G. Kroth, J. La Roche, E. Lindquist, M. Lommer, V. Martin-Jezequel, P.J. Lopez, S. Lucas, M. Mangogna, K. McGinnis, L.K. Medlin, A. Montsant, M.P. Oudot-Le Secq, C. Napoli, M. Obornik, M.S. Parker, J.L. Petit, B.M. Porcel, N. Poulsen, M. Robison, L. Rychlewski, T.A. Rynearson, J. Schmutz, H. Shapiro, M. Siaut, M. Stanley, M.R. Sussman, A.R. Taylor, A. Vardi, P. von Dassow, W. Vyverman, A. Willis, L.S. Wyrwicz, D.S. Rokhsar, J. Weissenbach, E.V. Armbrust, B.R. Green, Y. Van de Peer, I.V. Grigoriev, The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes, Nature 456 (2008) 239–244.

[10] B.J. Karas, R.E. Diner, S.C. Lefebvre, J. McQuaid, A.P. Phillips, C.M. Noddings, J.K. Brunson, R.E. Valas, T.J. Deerinck, J. Jablanovic, J.T. Gillard, K. Beeri, M.H. Ellisman, J.I. Glass, C.A. Hutchison 3rd, H.O. Smith, J.C. Venter, A.E. Allen, C.L. Dupont, P.D. Weyman, Designer diatom episomes delivered by bacterial conjugation, Nat. Commun. 6 (2015) 6925.

[11] M. Miyahara, M. Aoi, N. Inoue-Kashino, Y. Kashino, K. Ifuku, Highly efficient transformation of the diatom *Phaeodactylum tricornutum* by multi-pulse electroporation, Biosci. Biotechnol. Biochem. 77 (2013) 874–876.

[12] Y.F. Niu, Z.K. Yang, M.H. Zhang, C.C. Zhu, W.D. Yang, J.S. Liu, H.Y. Li, Transformation of diatom *Phaeodactylum tricornutum* by electroporation and establishment of inducible selection marker, BioTechniques 52 (2012) 1–3.

[13] L.A. Zaslavskaia, J.C. Lippmeier, P.G. Kroth, A.R. Grossman, K.E. Apt, Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes, J. Phycol. 36 (2000) 379–386.

[14] C.Y. Zhang, H.H. Hu, High-efficiency nuclear transformation of the diatom *Phaeodactylum tricornutum* by electroporation, Mar. Genomics 16 (2014) 63–66.

[15] V. De Riso, R. Raniello, F. Maumus, A. Rogato, C. Bowler, A. Falciatore, Gene silencing in the marine diatom *Phaeodactylum tricornutum*, Nucleic Acids Res. 37 (2009) e96.

[16] M. Christian, T. Cermak, E.L. Doyle, C. Schmidt, F. Zhang, A. Hummel, A.J. Bogdanove, D.F. Voytas, Targeting DNA double-strand breaks with TAL effector nucleases, Genetics 186 (2010) 757–761.

[17] T. Li, S. Huang, X. Zhao, D.A. Wright, S. Carpenter, M.H. Spalding, D.P. Weeks, B. Yang, Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes, Nucleic Acids Res. 39 (2011) 6315–6325.

[18] A.J. Wood, T.-W. Lo, B. Zeitler, C.S. Pickle, E.J. Ralston, A.H. Lee, R. Amora, J.C. Miller, E. Leung, X. Meng, L. Zhang, E.J. Rebar, P.D. Gregory, F.D. Urnov, B.J. Meyer, Targeted genome editing across species using ZFNs and TALENs, Science 333 (2011) 307.

[19] L. Cong, F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P.D. Hsu, X. Wu, W. Jiang, L.A. Marraffini, F. Zhang, Multiplex genome engineering using CRISPR/Cas systems, Science 339 (2013) 819–823.

[20] J.E. DiCarlo, J.E. Norville, P. Mali, X. Rios, J. Aach, G.M. Church, Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems, Nucleic Acids Res. 41 (2013) 4336–4343.

[21] W.Y. Hwang, Y. Fu, D. Reyon, M.L. Maeder, S.Q. Tsai, J.D. Sander, R.T. Peterson, J.R.J. Yeh, J.K. Joung, Efficient genome editing in zebrafish using a CRISPR-Cas system, Nat. Biotechnol. 31 (2013) 227–229.

[22] M.J. Moscou, A.J. Bogdanove, A simple cipher governs DNA recognition by TAL effectors, Science 326 (2009) 1501.

[23] S. Schornack, A. Meyer, P. Romer, T. Jordan, T. Lahaye, Gene-for-gene-mediated recognition of nuclear-targeted AvrBs3-like bacterial effector proteins, J. Plant Physiol. 163 (2006) 256–272.

[24] J. Bitinaite, D.A. Wah, A.K. Aggarwal, I. Schildkraut, FokI dimerization is required for DNA cleavage, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 10570–10575.

[25] M.R. Lieber, The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway, Annu. Rev. Biochem. 79 (2010) 181–211.

[26] Y. Zu, X. Tong, Z. Wang, D. Liu, R. Pan, Z. Li, Y. Hu, Z. Luo, P. Huang, Q. Wu, Z. Zhu, B. Zhang, S. Lin, TALEN-mediated precise genome modification by homologous recombination in zebrafish, Nat. Methods 10 (2013) 329–331.

[27] C. Gong, P. Bongiorno, A. Martins, N.C. Stephanou, H. Zhu, S. Shuman, M.S. Glickman, Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C, Nat. Struct. Mol. Biol. 12 (2005) 304–312.

[28] F. Daboussi, S. Leduc, A. Marechal, G. Dubois, V. Guyot, C. Perez-Michaut, A. Amato, A. Falciatore, A. Juillerat, M. Beurdeley, D.F. Voytas, L. Cavarec, P. Duchateau, Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology, Nat. Commun. 5 (2014) 3831.

[29] M. Nymark, A.K. Sharma, T. Sparstad, A.M. Bones, P. Winge, A CRISPR/Cas9 system adapted for gene editing in marine algae, Sci. Report. 6 (2016) 24951.

[30] P.D. Weyman, K. Beeri, S.C. Lefebvre, J. Rivera, J.K. McCarthy, A.L. Heuberger, G. Peers, A.E. Allen, C.L. Dupont, Inactivation of *Phaeodactylum tricornutum* urease gene using transcription activator-like effector nuclease-based targeted mutagenesis, Plant Biotechnol. J. 13 (2015) 460–470.

[31] X. Wang, Y. Wang, X. Wu, J. Wang, Y. Wang, Z. Qiu, T. Chang, H. Huang, R.J. Lin, J.K. Yee, Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors, Nat. Biotechnol. 33 (2015) 175–178.

[32] S. Stork, D. Moog, J.M. Przyborski, I. Wilhelmi, S. Zauner, U.G. Maier, Distribution of the SELMA translocon in secondary plastids of red algal origin and predicted uncoupling of ubiquitin-dependent translocation from degradation, Eukaryot Cell 11 (2012) 1472–1481.

[33] N.E. Sanjana, L. Cong, Y. Zhou, M.M. Cunniff, G. Feng, F. Zhang, A transcription activator-like effector toolbox for genome engineering, Nat. Protoc. 7 (2012) 171–192.

[35] B. Schellenberger Costa, M. Sachse, A. Jungandreas, C.R. Bartulos, A. Gruber, T. Jakob, P.G. Kroth, C. Wilhelm, Aureochrome 1a Is involved in the photoacclimation of the diatom, PLoS One (2013).

[36] E.L. Doyle, N.J. Booher, D.S. Standage, D.F. Voytas, V.P. Brendel, J.K. Vandyk, A.J. Bogdanove, TAL effector-nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction, Nucleic Acids Res. 40 (2012) W117–W122.

[37] Y. Lin, E.J. Fine, Z. Zheng, C.J. Antico, R.A. Voit, M.H. Porteus, T.J. Cradick, G. Bao, SAPTA: a new design tool for improving TALE nuclease activity, Nucleic Acids Res. 42 (2014), e47.

[38] K.E. Apt, P.G. Kroth-Pancic, A.R. Grossman, Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*, Mol. Gen. Genet. 252 (1996) 572–579.

[39] P.G. Kroth, Genetic transformation: a tool to study protein targeting in diatoms, Methods Mol. Biol. 390 (2007) 257–267.

[40] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.

[41] S.W. Jeffrey, G.F. Humphrey, New spectrophotometric equations for determining chlorophylls a, B, C1 and C2 in higher-plants, algae and natural phytoplankton, Biochem. Physiol. Pflanz. 167 (1975) 191–194.

[42] T. Jakob, R. Goss, C. Wilhelm, Activation of diadinoxanthin de-epoxidase due to a chlororespiratory proton gradient in the dark in the diatom *Phaeodactylum tricornutum*, Plant Biol. 1 (1999) 76–82.

[43] J. Rottberger, A. Gruber, P.G. Kroth, Analysing size variation during light-starvation response of nutritionally diverse chrysophytes with a Coulter counter, Algol. Stud. 141 (2013) 37–51.

[44] L. Chu, D. Ewe, C. Rio Bartulos, P.G. Kroth, A. Gruber, Rapid induction of GFP expression by the nitrate reductase promoter in the diatom *Phaeodactylum tricornutum*, PeerJ 4 (2016), e2344.

[45] D. Hanahan, Studies on transformation of *Escherichia coli* with plasmids, J. Mol. Biol. 166 (1983) 557–580.

[46] M. Ohse, K. Takahashi, Y. Kadowaki, H. Kusaoke, Effects of plasmid DNA sizes and several other factors on transformation of Bacillus subtilis ISW1214 with plasmid DNA by electroporation, Biosci. Biotechnol. Biochem. 59 (1995) 1433–1437.

[47] C. Krysiak, B. Mazus, J. Buchowicz, Generation of DNA double-strand breaks and inhibition of somatic embryogenesis by tungsten microparticles in wheat, Plant Cell Tissue Organ Cult. 58 (1999) 163–170.

[48] J. Lavaud, A.C. Materna, S. Sturm, S. Vugrinec, P.G. Kroth, Silencing of the violaxanthin de-epoxidase gene in the diatom *Phaeodactylum tricornutum* reduces diatoxanthin synthesis and non-photochemical quenching, PLoS One 7 (2012), e36806.

[49] Z. Mao, M. Bozzella, A. Seluanov, V. Gorbunova, Comparison of nonhomologous end joining and homologous recombination in human cells, DNA Repair (Amst) 7 (2008) 1765–1771.

[50] Y. Fu, J.A. Foden, C. Khayter, M.L. Maeder, D. Reyon, J.K. Joung, J.D. Sander, High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells, Nat. Biotechnol. 31 (2013) 822–826.

[51] P.D. Hsu, D.A. Scott, J.A. Weinstein, F.A. Ran, S. Konermann, V. Agarwala, Y. Li, E.J. Fine, X. Wu, O. Shalem, T.J. Cradick, L.A. Marraffini, G. Bao, F. Zhang, DNA targeting specificity of RNA-guided Cas9 nucleases, Nat. Biotechnol. 31 (2013) 827–832.

[52] C. Kuscu, S. Arslan, R. Singh, J. Thorpe, M. Adli, Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease, Nat. Biotechnol. 32 (2014) 677–683.

[53] J.C. Miller, S. Tan, G. Qiao, K.A. Barlow, J. Wang, D.F. Xia, X. Meng, D.E. Paschon, E. Leung, S.J. Hinkley, G.P. Dulay, K.L. Hua, I. Ankoudinova, G.J. Cost, F.D. Urnov, H.S. Zhang, M.C. Holmes, L. Zhang, P.D. Gregory, E.J. Rebar, A TALE nuclease architecture for efficient genome editing, Nat. Biotechnol. 29 (2011) 143–148.

[54] J.P. Guilinger, V. Pattanayak, D. Reyon, S.Q. Tsai, J.D. Sander, J.K. Joung, D.R. Liu, Broad specificity profiling of TALENs results in engineered nucleases with improved DNA cleavage specificity, Nat. Methods 11 (2014) 429–435.

[55] V. Pattanayak, S. Lin, J.P. Guilinger, E. Ma, J.A. Doudna, D.R. Liu, High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity, Nat. Biotechnol. 31 (2013) 839–843.

[56] S. Stella, G. Montoya, The genome editing revolution: a CRISPR-Cas TALE off-target story, Inside Cell 1 (2016) 7–16.

[57] B. Bailleul, A. Rogato, A. de Martino, S. Coesel, P. Cardol, C. Bowler, A. Falciatore, G. Finazzi, An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 18214–18219.

[58] B. Lepetit, G. Gelin, M. Lepetit, S. Sturm, S. Vugrinec, A. Rogato, P.G. Kroth, A. Falciatore, J. Lavaud, The diatom *Phaeodactylum tricornutum* adjusts nonphotochemical fluorescence quenching capacity in response to dynamic light via fine-tuned Lhcx and xanthophyll cycle pigment synthesis, New Phytol. (2016), http://dx.doi.org/10.1111/nph.14337.

[59] B. Schellenberger Costa, A. Jungandreas, T. Jakob, W. Weisheit, M. Mittag, C. Wilhelm, Blue light is essential for high light acclimation and photoprotection in the diatom *Phaeodactylum tricornutum*, J. Exp. Bot. 64 (2013) 483–493.

[60] A. Rossi, Z. Kontarakis, C. Gerri, H. Nolte, S. Holper, M. Kruger, D.Y. Stainier, Genetic compensation induced by deleterious mutations but not gene knockdowns, Nature 524 (2015) 230–233.

[61] F.O. Kok, M. Shin, C.W. Ni, A. Gupta, A.S. Grosse, A. van Impel, B.C. Kirchmaier, J. Peterson-Maduro, G. Kourkoulis, I. Male, D.F. DeSantis, S. Sheppard-Tindell, L. Ebarasi, C. Betsholtz, S. Schulte-Merker, S.A. Wolfe, N.D. Lawson, Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish, Dev. Cell 32 (2015) 97–108.

[62] Y. Gao, Y. Zhang, D. Zhang, X. Dai, M. Estelle, Y. Zhao, Auxin binding protein 1 (ABP1) is not required for either auxin signaling or Arabidopsis development, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 2275–2280.

[63] M.J. Huysman, A.E. Fortunato, M. Matthijs, B.S. Costa, R. Vanderhaeghen, H. Van den Daele, M. Sachse, D. Inze, C. Bowler, P.G. Kroth, C. Wilhelm, A. Falciatore, W. Vyverman, L. De Veylder, AUREOCHROME1a-mediated induction of the diatom-specific cyclin dsCYC2 controls the onset of cell division in diatoms (*Phaeodactylum tricornutum*), Plant Cell 25 (2013) 215–228.