This is a repository copy of *Design of an effective sgRNA for CRISPR/Cas9 knock-ins and full mutant segregation in polyploid Synechocystis sp. PCC 6803*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/223141/

Version: Published Version

Article:

Article

Design of an Effective sgRNA for CRISPR/Cas9 Knock-Ins and Full Mutant Segregation in Polyploid *Synechocystis* sp. PCC 6803

Maria Isabel Nares-Rodriguez and Esther Karunakaran *

School of Chemical, Materials and Biological Engineering, University of Sheffield, Sheffield S1 3JD, UK; isabel.nares@gmail.com
* Correspondence: e.karunakaran@sheffield.ac.uk; Tel.: +44-(114)-2227166

Abstract: *Synechocystis* sp. PCC 6803 is a highly promising organism for the production of diverse recombinant compounds, including biofuels. However, conventional genetic engineering in *Synechocystis* presents challenges due to its highly polyploid genome, which not only results in low product yields but also compromises the reliability of recombinant strains for biomanufacturing applications. The CRISPR/Cas9 system, renowned for its precision, efficiency, and versatility across a wide range of chassis, offers significant potential to address the limitations posed by polyploid genomes. In this study, we developed and optimized an effective sgRNA for the targeted knock-in of nucleotide sequences of varying lengths into the neutral locus *slr*0168 of polyploid *Synechocystis* using CRISPR/Cas9. The gene encoding di-geranylgeranylglycerophospholipid reductase from *Sulfolobus acidocaldarius* and the methyl ketone operon from *Solanum habrochaites* were chosen as the exemplar nucleotide sequences for incorporation into the chromosome of *Synechocystis.* Our results demonstrate that the designed sgRNA effectively facilitated both knock-in events and that CRISPR/Cas9 enabled complete mutant segregation in a single round of selection and induction.

Keywords: CRISPR/Cas9; sgRNA design; full mutant segregation; *Synechocystis* sp. PCC 6803; single induction step

1. Introduction

A central objective of genetic engineering is to enable the expression of heterologous proteins in a host organism, facilitating the production of a desired target compound. For reliable and efficient industrial production, the genes encoding these heterologous proteins must be stably integrated and homogenic across all chromosomal copies in the host. This ensures their inheritance and maintenance through successive cell divisions, preventing mutation dilution and loss of mutants during manufacturing, thereby sustaining high yields of the target product [1–4].

The polyploid cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter referred to as *Synechocystis*) has emerged as a promising platform for the production of various recombinant products with applications across diverse industries, including biofuels and personal care [5–10]. However, traditional genetic engineering in *Synechocystis* is often labor-intensive, time-consuming, and costly, primarily due to the challenges posed by its highly polyploid genome. Strategies such as phosphate deprivation, multiple counterselections, and combined random mutagenesis have been employed to address these limitations [11–13]. These approaches leverage the naturally occurring mechanism of double homologous recombination (DHR) in *Synechocystis* [7,12,14,15], but achieving full

mutant segregation remains a significant challenge [3,11–13]. As a result, gene expression and the yields of target products are frequently suboptimal, limiting the feasibility of using *Synechocystis* as a host for industrial-scale production of heterologous products [1,4,16–18]. Emerging technologies, particularly CRISPR/Cas9, hold the potential to overcome these obstacles and offer a streamlined, effective solution for genetic engineering in *Synechocystis*.

The Type II CRISPR system (Clustered Regularly Interspaced Short Palindromic Repeats) and its associated nuclease Cas9 (CRISPR/Cas9) function as an adaptive immune mechanism in archaea and bacteria, repurposed for precise and efficient genome editing. The CRISPR/Cas9 system utilizes a single guide RNA (sgRNA), a Cas9 nuclease, and donor DNA (dDNA) to introduce targeted modifications into the host genome via homology-directed repair (HDR) triggered by double-strand breaks (DSBs) induced by Cas9 [19–23]. Previous gene editing strategies are based on DNA–protein interactions which are considered highly complex. These have not reached the level of effectiveness and precision deemed to CRISPR/Cas9 [23–26], which acts upon nucleotide sequence complementarity and is considered to be simpler; this endows CRISPR/Cas9 with its highly elevated rate of success [20,27,28]. Moreover, CRISPR/Cas9 is considered to be efficient, reliable and precise due to its chimeric and easily reprogrammable sgRNA, activity of Cas9 and the repair mechanism activated by DSB [29–31]. CRISPR/Cas9 is able to perform multiplexed insertions, inactivation, deletions and editing of genes in multiple loci in a simultaneous manner [20,32–34] and has successfully been implemented in a wide range of hosts, including humans [25,29,30,35–38]. Therefore, CRISPR/Cas9 may have the potential of overcoming bottlenecks such as the polyploid genome in the cyanobacterium *Synechocystis*.

Several studies have explored the application of CRISPR/Cas technology in cyanobacteria. Early efforts utilized Cas9 from *Streptococcus pyogenes* in *Synechococcus* UTEX 2973 and *Synechococcus elongatus* PCC 7942, achieving transient expression. However, Cas9 toxicity hindered further genome editing [38,39]. Subsequently, studies adopted Cpf1 (Cas12a), enabling constitutive expression with reduced toxicity in cyanobacterial hosts such as *Synechococcus* UTEX 2973, *Anabaena* sp. PCC 7120 and *Synechocystis* sp. PCC 6803. These studies achieved efficiencies ranging from 20% to 90% for constructs up to 1950 bp, although multiple selection rounds were required to attain full segregation for chromosomal knock-ins, knock-outs and point mutations [40]. Later work revisited Cas9, engineering an inducible CRISPR/Cas9 system in *Synechocystis* to mitigate Cas9 toxicity via an anhydrotetracycline (aTc)-controlled promoter. This approach enabled genome integration and Cas9 expression but suffered from leaky expression, resulting in low transformation efficiency and limiting its applicability to small native plasmid deletions [41]. More recently, a study implemented a theophylline-inducible riboswitch to achieve stringent control over Cas9 for markerless insertions and deletions in the *slr*0168 neutral site and other integration loci in the *Synechocystis* genome. This method achieved efficiencies of 25% to 75% for fully segregated mutants with deletions and small insertions, including 100% efficiency for a 24 bp insertion after a single induction round. However, for large insertions, the results displayed a low efficiency of 8.3% of the knock-in of a 2240 bp construct, showing only a single fully edited colony per 12 screened transformants [42].

In this study, we engineered polyploid *Synechocystis* by optimizing the inducible expression of CRISPR/Cas9 under the control of an anhydrotetracycline (aTc)-regulated promoter to mitigate Cas9 toxicity. Additionally, we designed a novel single guide RNA (sgRNA) targeting the *slr*0168 neutral site for chromosomal insertions. Two successful chromosomal knock-ins were demonstrated at the *slr*0168 site: first, a 1551 bp markerless insertion containing the coding sequence for the heterologous enzyme digeranylgeranylglycerophospholipid reductase (GGR), and second, a 2500 bp insertion comprising the coding sequences for the heterologous enzymes MKS1 and MKS2 from

the methyl ketone synthesis operon, along with a chloramphenicol selection marker and control elements. Both transformants achieved full segregation after a single round of Cas9 induction, with transformation efficiencies of up to 70% and 80%, respectively. This study advances the state of the art by achieving full segregation of both markerless and marker-containing mutants for CRISPR/Cas9-mediated knock-ins of constructs larger than those previously reported, all within a single induction round.

2. Results

2.1. Selection of Target Site, Cas9 and Protospacer

The *slr*0168 gene was selected as the target site for incorporating genes encoding heterologous proteins (cargo) due to its identification as a highly effective neutral site in *Synechocystis* [43]. This selection minimizes the risk of unintended phenotypic changes in the host, ensuring stability and predictability during genetic engineering.

The Cas9 nuclease from *S. pyogenes* was employed in this study as it is the most widely utilized nuclease for CRISPR-mediated genome editing [44,45]. Cas9 recognizes the NGG protospacer adjacent motif (PAM) [46,47], which has been shown to exhibit the highest activity compared to other PAM sequences [48–50]. However, the specificity of CRISPR elements, particularly the sgRNA, is not guaranteed, as the protospacer region of the sgRNA may bind to unintended genomic sites, leading to off-target effects [51,52].

The protospacer region of the sgRNA is complementary to the target genomic site and is designed to bind adjacent to the PAM sequence, facilitating cleavage at the desired target site [20]. Within the protospacer, the first 12 nucleotides proximal to the PAM form the seed region, which is critical for accurate cleavage. Mismatches in this seed region can significantly reduce the activity of the sgRNA/Cas9 complex and increase the likelihood of off-target effects. Conversely, mismatches in the non-seed region (the remaining protospacer) have less critical impact but can still contribute to off-target editing [48,49]. Hence, the fewer number of mismatches within the seed region would be less likely to have an impact on cleavage activity. Similarly, the fewer number of both overall mismatches and off-targets would increase the probabilities of effective target-binding from a protospacer, which is why careful evaluation is crucial in order to minimize the probable off-target activity of a given protospacer to increase its specificity and cleavage activity. To optimize specificity and minimize off-target activity, CasOT software [53] was used to scan the *slr*0168 gene and identify protospacer candidates associated with the designated PAM. Each candidate was evaluated for off-target effects, quantified as types of mismatches (e.g., A205, A206) and their respective off-target counts (shaded in blue) (Figure 1). In these codes, "A" signifies the selection of an NGG PAM, the first digit indicates mismatches in the seed region and the subsequent two digits represent mismatches in the non-seed region [40]. A total of 73 protospacer candidates were identified within *slr*0168, including protospacer 64r, which was selected for sgRNA design in this study. For comparative purposes, protospacers 33r, 45r and 287f were also discussed (Figure 1).

Figure 1. (**a**) Off-target analysis of the non-selected and selected (in red) protospacers showing the type and number of mismatches (codes: I.E. A206) and number of off-targets for each (shaded in blue) type of mismatch. (**b**) Description of the off-targets and mismatches (lowercase) of protospacer 64r within seed (green) and non-seed (orange) regions.

Protospacer 64r (Supplementary Table S1) is 20 nucleotides in length, aligning with the recommended range for effective binding [20,22]. It exhibited a total of three potential off-target sequences: one associated with the A205 mismatch type and two with the A206 mismatch type. For the A205 mismatch, the sequence features two mismatches in the seed region and five in the non-seed region. Similarly, the A206 mismatch type involves two mismatches in the seed region and six in the non-seed region. The A000 code represents the original target sequence of protospacer 64r and is classified neither as a mismatch nor an off-target. All identified off-target sequences for protospacer 64r are presented in Figure 1.

Compared to other candidates, protospacer 64r demonstrated superior specificity with the lowest number of off-targets and mismatches. It displayed only two mismatch types (A205 and A206), whereas other candidates, such as 287f, exhibited up to 11 mismatch types (A107 to A210). Moreover, protospacer 64r had just three off-target sequences (one for A205 and two for A206), in contrast to candidates like 33r, which showed up to 28 off-target sequences for a single mismatch type (A207). Additionally, protospacer 64r had the fewest mismatches in both the seed and non-seed regions, as well as the smallest number of off-target sequences. These characteristics make it least likely to bind to unintended genomic sites and highly likely to exhibit optimal activity, enabling efficient knock-ins at the desired editing site.

Protospacer 64r, when integrated with the Cas9 handle, must enable the proper formation of the handle's secondary structure, which is essential for assembling the sgRNA-Cas9 complex and inducing the double-strand break (DSB) at the target site [20,54]. The sequence of the protospacer plays a critical role in determining whether this secondary structure is correctly formed. The free energy associated with the interaction between the protospacer and the Cas9 handle significantly impacts the formation of the proper hairpin structure. Protospacer–handle sequences with a free energy value within 5% of the optimal folded structure are considered acceptable for effective functioning [54]. To ensure compatibility,

the secondary structure of protospacer 64r alone was analyzed to confirm its ability to form a proper handle loop without the presence of rigid structures within its sequence (Figure 2b). The frequency of the minimum free energy (MFE) structure for the protospacer–handle sequence was calculated at 76.03%, which is within 5% of the optimal Cas9 handle folded structure (74.71%). These results suggest that an effective sgRNA-Cas9 complex is likely to form intracellularly.
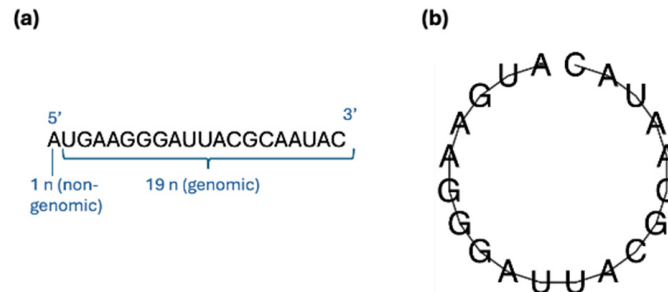

Figure 2. (**a**) Sequence of the protospacer 64r and (**b**) secondary structure analysis of 64r protospacer alone, showing the absence of rigid and secondary structures within itself that might hamper the formation of a proper Cas9 handle.

Considering this, along with the other selection criteria previously discussed, the designed sgRNA (Figure 3d; full sequences provided in Supplementary Table S1) demonstrates the highest probability of facilitating efficient genome editing at the *slr*0168 target site.
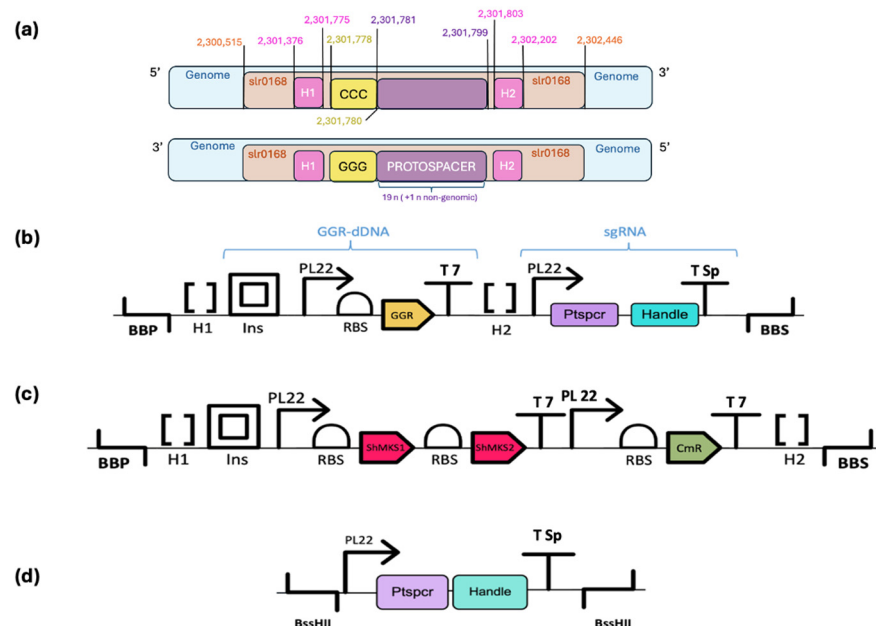

Figure 3. (**a**) Genetic context of protospacer 64r of the sgRNA, showing elements upstream and downstream of the protospacer in the genome of *Synechocystis* WT; (**b**) schematic of the sgRNA_GGR-dDNA, including flanking BioBrick prefix and suffix (BBP and BBS), upstream and downstream homology regions (H1 and H2), an insulator (ins), PL22 promoter, selected RBS, heterologous *GGR* gene, T7 Terminator followed by the sgRNA, using the same PL22 promoter but a *S. pyogenes* terminator (T Sp); (**c**) schematic showing the ShMKS-dDNA, with similar elements to (**b**) but two heterologous genes (*ShMKS1* and *ShMKS2*), encompassing the operon and a chloramphenicol resistance marker (CmR); (**d**) schematic showing the designed sgRNA with BssHII overhangs, PL22 promoter, 64r protospacer, Cas9 handle and *S. pyogenes* terminator (T Sp).

Figure 3a shows the genetic context of the 64r protospacer located in the *slr*0168 (SGL_RS12680) neutral site which expands from base position 2,300,515 to 2,302,446 in the genome of Synechocystis. Neutral sites do not contain critical genes, regulatory elements, or other essential genetic functions; hence, they are commonly targeted for genetic manipulation, including insertions without disrupting essential functions [41,43,55]. Gene *slr*0168 was selected as the target site for the incorporation of genes encoding heterologous proteins (cargo) because it has been identified as a highly effective neutral site in Synechocystis [43], and therefore incorporation of the cargo at this site can avoid any undesired phenotypical changes to the host. Within *slr*0168, the 64r protospacer is located from base position 2,301,781 to 2,301,799 in the reverse orientation on the bottom strand, and the GGG PAM is located downstream of the protospacer from base position 2,301,778 to 2,301,380 (Figure 3a). The 64r protospacer consists of a genomic sequence of 19 nucleotides with 1 non-genomic nucleotide (A) added due to promoter affinity preferences at the 5′ end of the protospacer sequence.

2.2. Design of Donor DNA Templates

Two donor DNA templates were designed in this study, sgRNA_GGR-dDNA (Figure 3b) and ShMKS-dDNA (Figure 3c). The first template encompassed the coding sequence corresponding to the enzyme digeranylgeranylglycerophospholipid reductase (GGR) from *Sulfolobus acidocaldarius* (1359 bp) whilst the other template encompassed the coding sequences corresponding to enzymes MKS1 and MKS2 belonging to the methyl ketone synthesis operon from *Solanum habrochaites* (798 bp and 627 bp, respectively). The gene sequences were codon optimized for *Synechocystis* using the IDT (Integrated DNA Technologies) codon optimization tool. The coding sequences in both templates were placed under the control of the anhydrotetracycline-inducible PL22 promoter, the B0034 RBS and the T7 terminator. PL22 was chosen as it has been reported to provide a suitable expression of CRISPR elements [56,57]. In both templates, the assembled genetic circuits were placed in between homology arms comprising approximately 400 bp upstream and downstream of the chosen protospacer sequence in *slr*0168 (Figure 3a, exact base positions on the genome and accession numbers are provided in Supplementary Tables S2 and S3). Care was taken to omit the PAM sequence from inclusion in the homology arms to prevent unintended and continuous cleavage from Cas9 at the target site.

Once the basic design of the donor DNA was complete, two divergent approaches were taken with respect to further design. The construct harbouring the *GGR* gene was chosen to generate a markerless knock-in mutant for the evaluation of Cas9-based genome editing in *Synechocystis*. Therefore, an antibiotic selection marker was not included in the template. The total size of the cargo excluding the homology arms was 1551 bp and was titled sgRNA_GGR-dDNA. The genetic circuit harbouring the coding sequence for the designed sgRNA under the control of the PL22 promoter and the *S. pyogenes* terminator (T Sp) was placed in the vicinity of the cargo. The assembly was placed in the pEERM4 plasmid backbone. Conversely, the construct harbouring the *ShMKS1* and *ShMKS2* genes was chosen to generate a knock-in mutant with chloramphenicol resistance as the selection marker. The assembly was placed in the pEERM3 plasmid backbone. The total size of the cargo excluding the homology arms was 2500 bp and was titled ShMKS-dDNA. The genetic circuit harbouring the coding sequence for the designed sgRNA under the control of the PL22 promoter in this template was inserted into the BssHII site of the backbone.

Synechocystis bearing the gene encoding the *S. pyogenes* Cas9 on plasmid pPMQK1 under the control of the PL22 promoter was used as the host organism. This strain was independently conjugated with either donor DNA template and subjected to a single round of induction with aTc.

2.3. Mutant Segregation Analysis

Induction of the CRISPR/Cas9 system was carried out in both strains with an aTc concentration of 200 ng/mL and white light for a first attempt and subsequently modified to 400 ng/mL of aTc and red light for a second attempt. The segregation state of the mutants was evaluated by PCR screening after a single round of selection/induction. Ten mutants for each construct were evaluated. For the first attempt, neither mutant showed full mutant segregation. However, full mutant segregation was achieved on the second attempt (Figure 4) in 7 out of 10 screened mutants after a single round of induction for *Synechocystis*::Cas9_sgRNA_GGR-dDNA, showing an editing efficiency of 70%, and 8 out of 10 screened mutants for *Synechocystis*::Cas9_sgRNA_ShMKS-dDNA, showing an editing efficiency of 80%.
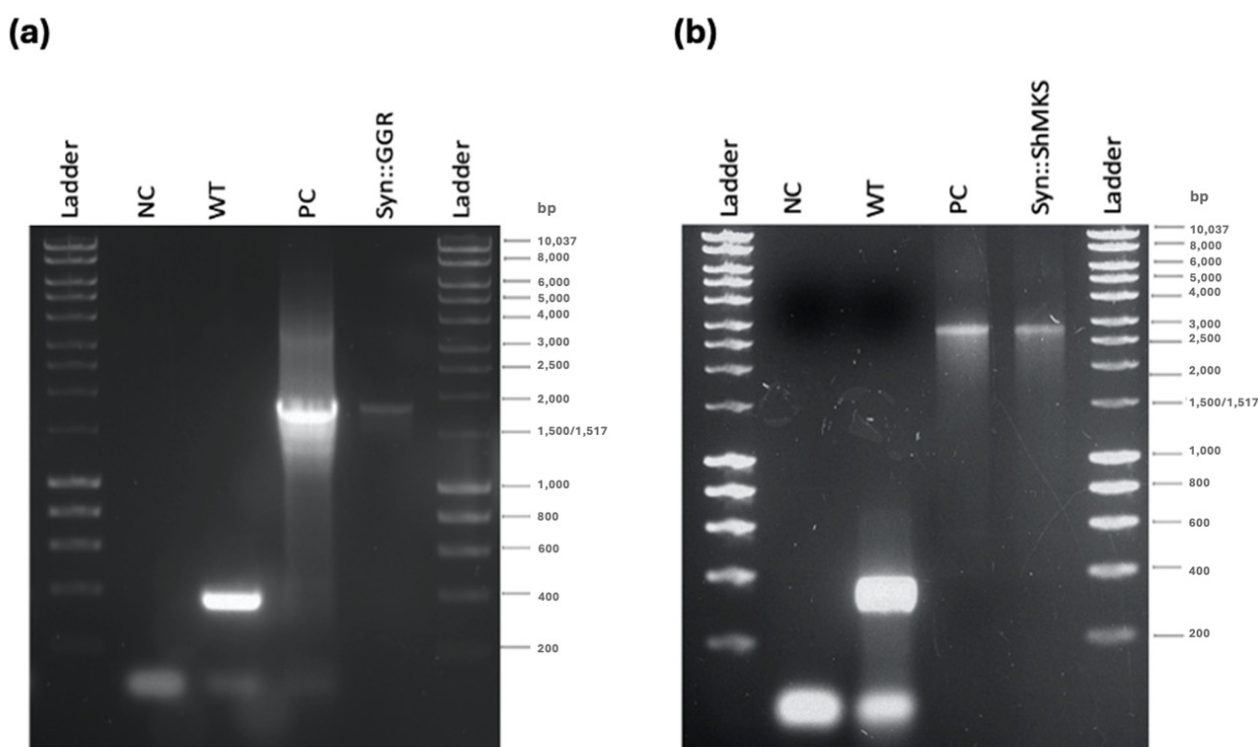

Figure 4. (**a**) Mutant segregation analysis of *Synechocystis*::Cas9_sgRNA_GGR-dDNA and (**b**) *Synechocystis*::Cas9_sgRNA_ShMKS-dDNA. Showing a negative control (NC); *Synechocystis* WT displaying a band corresponding to the homology regions present in the GGR-dDNA, and ShMKS-dDNA taken from the target site in the genome of *Synechocystis* WT; a positive control (PC), displaying a amplicon of the desired size; (**a**) *Synechocystis*::Cas9_sgRNA_GGR-dDNA for and (**b**) *Synechocystis*::Cas9_sgRNA_shMKS-dDNA, both showing a band of the desired size that coincides with the desired band on their respective positive control as well as the lack of additional bands for both.

The representative mutant segregation analysis in Figure 4 shows a band of approximately 400 base pairs for the *Synechocystis* wild-type strain (WT). This band corresponds to the homology regions (H1 and H2) flanking GGR-dDNA as well as ShMKS-dDNA and taken from the target neutral site *slr*0168 in the genome of *Synechocystis* WT. Therefore, this band is expected to appear in the wild-type genome. Moreover, the amplicon from the pEERM4 vector bearing GGR-dDNA shows a band of approximately 1800 bp, corresponding to the sequence of GGR-dDNA, being amplified from the homology regions. Similarly, the amplicon from the pEERM3 vector bearing the ShMKS-dDNA shows a band of approximately 2800 bp, corresponding to the sequence of ShMKS-dDNA, amplified from the homology regions. Both bands are expected not only since they are positive

controls (PCs), representing the desired bands aimed to be present in the analysis of both mutants, respectively.

Furthermore, the mutant segregation analysis for *Synechocystis*::Cas9_sgRNA_GGR-dDNA (Figure 4a) shows a band of the desired size which coincides with the desired band on the positive control. This suggests that the GGR-dDNA sequence is present in the genome of such mutants, and hence that knock-in of GGR-dDNA via CRISPR/Cas9 is effective. Furthermore, the desired band is present as a single band, highlighting the absence of additional bands. This indicates that only complete sequences of the GGR-dDNA are found in this mutant population and that the insertion of the GGR-dDNA into all chromosomal copies of all the individuals of the *Synechocystis*::Cas9_sgRNA_GGR-dDNA population is successful. It must be highlighted that this screening was made after a single round of induction. According to these results, the designed sgRNA is effective in the knock-in of a single gene up to 1551 bp. Placing the sgRNA in close vicinity to the cargo does not affect the incorporation of the cargo at *slr*0168. Furthermore, mutant segregation does not seem to be affected, although the template DNA did not bear a selection marker.

Comparably, the mutant segregation analysis for *Synechocystis*::Cas9_sgRNA_ShMKS-dDNA (Figure 4b) also shows that the expected band of the desired size is present for the CRISPR/Cas9 mutant, which coincides with the desired band on the positive control. Furthermore, the desired band is present as a single band in the lane and there are no other additional bands, similar to the GGR mutant. The results indicate the insertion of ShMKS-dDNA into all chromosomal copies of *Synechocystis* was successful. Thus, only complete sequences of the ShMKS-dDNA are present in all chromosomal copies within each cell and that all of these mutants in the culture bear the insert in their genome. It must be highlighted that this screening was made after a single round of antibiotic induction and that the ShMKS mutant bears a dual operon construct and a resistance marker gene, considerably larger than the GGR mutant.

According to these results, the size and complexity of the dDNA, up to 2500 bp, appears to have no effect on the ability of CRISPR/Cas9 to favoring mutant segregation after a single round of selection/induction. Furthermore, mutant segregation does not seem to be affected when sgRNA and dDNA are constructed as distinct inserts in the same vector for inserts bearing a selection marker.

3. Discussion

Genetic engineering aims to express desired heterologous traits in a specific chassis to achieve industrial-scale production of a target product. For this purpose, full segregation of the introduced mutation is essential. This ensures that the new traits are stable and homogeneously present across all chromosomal copies in every individual, preventing mutation dilution, loss of mutants, and reduced or null product yields [1,3]. Achieving complete mutant segregation is particularly challenging in polyploid genomes, such as that of *Synechocystis* [3,58,59]. However, CRISPR/Cas9 offers a promising solution, having been demonstrated as highly effective and reliable [27,30].

In this study, an sgRNA was designed for the successful knock-in of heterologous genes up to 2500 bp at the neutral *slr*0168 site in the highly polyploid *Synechocystis* genome. Complete mutant segregation was achieved in a single round of selection/induction using CRISPR/Cas9. Protospacer 64r was selected for its minimal mismatches and off-target effects, facilitating precise insertions. The proper secondary structure formation of the Cas9 handle was also optimized, ensuring effective sgRNA-Cas9 complex formation for intracellular double-strand break induction. A critical factor in achieving high efficiency was the use of red light, rather than white light, when utilizing the aTc-inducible PL22 promoter. Initial attempts under white light failed to achieve full mutant segregation, likely

due to the sensitivity of aTc to white light [58] and the weaker expression of PL22 under such conditions [60]. In contrast, red light significantly enhanced the expression of the CRISPR/Cas9 components driven by PL22. Furthermore, increasing the concentration of aTc from 200 ng/mL to 400 ng/mL improved mutant segregation efficiency, and this concentration is recommended for optimal performance in *Synechocystis*.

Both knock-in experiments demonstrated complete mutant segregation after a single round of selection/induction, regardless of the size or complexity of the donor DNA (dDNA). Constructs up to 2500 bp, including operons with two genes and a resistance marker, were successfully managed by the CRISPR/Cas9 system. Notably, the proximity of CRISPR/Cas9 elements to the cargo on the same plasmid had no effect on segregation efficiency. Additionally, sgRNA was equally effective for markerless mutants and those bearing selection markers.

Previous studies have reported reduced editing efficiencies for larger constructs. For example, Cengic et al. (2022) achieved only 8.3% efficiency for a 2240 bp knock-in with 350 bp homology arms, yielding just one fully segregated colony out of 12 screened [61]. To address this and to conserve the cost of DNA synthesis of the inserts and reserve the maximum nucleotide positions for the cargo, we extended the homology regions by 50 bp beyond those used by Cengic et al., conserving DNA synthesis costs while enhancing recombination efficiency. This approach resulted in an 80% editing efficiency for constructs up to 2500 bp, supporting the hypothesis that longer homology arms improve editing outcomes. Additionally, tri-parental mating [61] facilitated the delivery of genetic elements for all CRISPR/Cas9 transformations in *Synechocystis*.

In conclusion, this study developed an effective sgRNA and optimized methodology for CRISPR/Cas9-mediated knock-ins of heterologous genes up to 2500 bp in the neutral *slr*0168 site of *Synechocystis*. Complete mutant segregation was achieved after a single selection/induction step. This approach represents a valuable addition to the synthetic biology toolkit for genome editing in *Synechocystis*, a highly promising yet challenging polyploid host organism.

4. Materials and Methods

4.1. Strains and Media

All *E. coli* strains were cultivated at 37 °C in sterile lysogeny broth (LB) medium with an orbital agitation of 150 rpm for aeration. *E. coli* 10-beta was used for all the molecular cloning experiments, plasmid propagation and maintenance. *E. coli* CopyCutter EPI400 cells, bearing the pPMQK1 vector, were used for the propagation of the plasmid bearing the *Cas9* gene. *E. coli* MC1061 cells were used as a helper strain in all tri-parental mating events. The medium was supplemented with kanamycin, at a concentration of 50 µg/mL, chloramphenicol at a concentration of 25 µg/mL and ampicillin at a concentration of 100 µg/mL where appropriate.

The *Synechocystis* wild-type (WT) strain (*Synechocystis* sp. PCC 6803) used for the CRISPR mutagenesis was purchased from the Pasteur Culture Collection of Cyanobacteria (PCC). All *Synechocystis* cells were grown photoautotrophically in sterile Blue-Green liquid medium (BG-11) at a temperature of 25 °C, under continuous orbital agitation conditions of 100 rpm for aeration. Initially, the mutants were placed under white light, with an intensity of 125 µmol/m^2/s; subsequently, this was changed to red light, with an intensity of 20 µmol/m^2/s. The media for *Synechocystis* mutants bearing chloramphenicol resistance was supplemented with chloramphenicol at a concentration of 12 µg/mL.

4.2. Insert and Vector Construction

The pPMQK1 vector (Supplementary Figure S1), which was a kind gift from Dr. Paul Hudson and his group (Karolinska Institute, Solna, Sweden), carried the *Streptococcus pyogenes* Cas9 nuclease expression cassette and was conjugated into *Synechocystis* for its expression. The pEERM4 Cm (Addgene plasmid # 64026) (Supplementary Figure S2) and the pEERM3 Km vectors (Addgene plasmid # 64025) (Supplementary Figure S3) were kind gifts from Pia Lindberg [60]. The QIAGEN Miniprep Spin Kit was used following manufacturer's instructions for plasmid extraction from *E. coli* strains. For plasmid manipulation, appropriate enzymes were sourced from New England Biolabs (Oxfordshire, UK), and manufacturer's instructions were followed. Chemically competent *E. coli* cells were purchased from New England Biolabs (Oxfordshire, UK) and the heat shock method was performed following manufacturer's instructions for plasmid verification and maintenance. Once designed, the sgRNA and donor DNA sequences were synthesized as custom genes by Eurofins Genomics.

4.3. sgRNA Selection

All available protospacer candidates in *slr*0168 were identified using CasOT web-based off-target prediction tool (single version) [53]. To enable the identification of suitable protospacer sequences, the sequence of *slr*0168 was submitted to CasOT as a FASTA file and the following parameters were set as the selection criteria. The -NGG Protospacer Adjacent Motif (PAM) only option (option "A-NGG only") was selected to generate only the candidates to match this PAM. The "target and off-target" mode was selected to allow the generation of all the possible protospacer candidates followed by off-target analysis. A maximum of 2 mismatches were allowed in the seed region. The "non-5'G" requirement option was selected as the T7 promoter was not used in the sgRNA. Finally, a "length" of 19 to 22 bp for the base-pairing region of the candidates was allowed. For the off-target analysis, each protospacer candidate was evaluated for the number of possible mismatches and off-targets they presented.

4.4. Cas9 Handle Secondary Structure Analysis

The RNAfold online tool from the ViennaRNA Package was used to evaluate whether the Cas9 handle formed a proper secondary structure after the addition of protospacer 64r [54]. Protospacer 64r 20 nucleotides were appended to the 42 nucleotide sequences of the Cas9 handle, for a total of 62 nucleotide sequences, and subsequently input into the algorithm to evaluate its secondary structure formation. The minimum free energy (MFE) and partition function, avoidance of isolated base pairs, absence of dangling energies, and the Turner model (1999) options for RNA parameters and the default criteria for the output options were selected.

4.5. Strain Construction and CRISPR/Cas9 Induction

The pPMQK1 vector was conjugated into *Synechocystis* to generate *Synechocystis*::Cas9 through standard protocol for tri-parental mating in cyanobacteria [61]. Subsequently, the pEERM4::sgRNA_GGR-dDNA and the pEERM3::ShMKS-dDNA vectors were conjugated independently into *Synechocystis*::Cas9 following a standard protocol for tri-parental mating in cyanobacteria to generate *Synechocystis*::Cas9_sgRNA_GGR-dDNA and *Synechocystis*::Cas9_sgRNA_ShMKS-dDNA, respectively. The transformants were transferred to non-selective BG-11 plates for 24 h for recovery and then subjected to both selection and induction. Only ShMKS mutants were subjected to both selection and induction while GGR mutants were only subjected to induction, effectuated on BG-11 plates supplemented with anhydrotetracycline (aTc) at a concentration of 200 ng/mL while using white light

for the first attempt and a subsequent one with an aTc concentration of 400 ng/mL under red light.

4.6. PCR Analysis of Mutant Segregation

Segregation analysis was performed on genomic DNA extracts of *Synechocystis* and *Synechocystis* mutants obtained using the Gen-Elute Bacterial Genomic DNA Kit (Sigma-Aldrich, Burlington, MA, USA) by effectuating a screen for GGR and ShMKS incorporation into *slr*0168 using PCR. PCR analysis was performed using the GoTaq Green Master Mix (Promega). Segregation primers (Supplementary Table S4) amplifying GGR-dDNA and ShMKS-dDNA from the homology regions were designed and used for analysis. The mutants were screened after a single round of selection/induction to determine their state of mutant segregation.

References

1. Caicedo-Burbano, P.; Smit, T.; Herna, H.P.; Du, W.; Branco dos Santos, F. Construction of fully segregated genomic libraries in polyploid organisms such as *Synechocystis* sp. PCC 6803. *ACS Synth. Biol.* **2020**, *9*, 2632–2638. [CrossRef] [PubMed]
2. Hitchcock, A.; Hunter, C.N.; Canniffe, D.P. Progress and challenges in engineering cyanobacteria as chassis for light-driven biotechnology. *Microb. Biotechnol.* **2020**, *13*, 363–367. [CrossRef] [PubMed]
3. Pope, M.A. Improving *Synechocystis* sp. PCC 6803 as a Model Organism. Ph.D. Thesis, Imperial College London, London, UK, 2020.
4. Taton, A.; Ma, A.T.; Ota, M.; Golden, S.S.; Golden, J.W. NOT Gate Genetic Circuits to Control Gene Expression in Cyanobacteria. *ACS Synth. Biol.* **2017**, *6*, 2175–2182. [CrossRef] [PubMed]
5. Gao, X.; Sun, T.; Pei, G.; Chen, L.; Zhang, W. Cyanobacterial chassis engineering for enhancing production of biofuels and chemicals. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 3401–3413. [CrossRef]
6. Harrison, K.W.; Harvey, B.G. High cetane renewable diesel fuels prepared from bio-based methyl ketones and diols. *Sustain. Energy Fuels* **2018**, *2*, 367–371. [CrossRef]

7. Heidorn, T.; Camsund, D.; Huang, H.-H.; Lindberg, P.; Oliveira, P.; Stensjö, K.; Lindblad, P. Synthetic Biology in Cyanobacteria: Engineering and Analyzing Novel Functions. In *Methods in Enzymology*; Academic Press: Cambridge, MA, USA, 2011; Volume 497, pp. 539–579.

8. Oliver, N.J.; Rabinovitch-Deere, C.A.; Carroll, A.L.; Nozzi, N.E.; Case, A.E.; Atsumi, S. Cyanobacterial metabolic engineering for biofuel and chemical production. *Curr. Opin. Chem. Biol.* **2016**, *35*, 43–50. [CrossRef] [PubMed]

9. Pade, N.; Erdmann, S.; Enke, H.; Dethloff, F.; Dühring, U.; Georg, J.; Wambutt, J.; Kopka, J.; Hess, W.R.; Zimmermann, R.; et al. Insights into isoprene production using the cyanobacterium *Synechocystis* sp. PCC 6803. *Biotechnol. Biofuels* **2016**, *9*, 89. [CrossRef]

10. Shabestary, K.; Hudson, E.P. Computational metabolic engineering strategies for growth-coupled biofuel production by *Synechocystis*. *Metab. Eng. Commun.* **2016**, *3*, 216–226. [CrossRef]

11. Deshpande, A.; Vue, J.; Morgan, J. Combining random mutagenesis and metabolic engineering for enhanced tryptophan production in *Synechocystis* sp. strain PCC 6803. *Appl. Environ. Microbiol.* **2020**, *86*, e02816–e02819. [CrossRef] [PubMed]

12. Pope, M.A.; Hodge, J.A.; Nixon, P.J. An improved natural transformation protocol for the cyanobacterium *Synechocystis* sp. PCC 6803. *Front. Plant Sci.* **2020**, *11*, 372. [CrossRef] [PubMed]

13. Ranade, S.; He, Q. Escherichia coli AraJ boosts utilization of arabinose in metabolically engineered cyanobacterium *Synechocystis* sp. PCC 6803. *AMB Express* **2021**, *11*, 115. [PubMed]

14. Gao, E.-B.; Kyere-Yeboah, K.; Wu, J.; Qiu, H. Photoautotrophic production of p-Coumaric acid using genetically engineered *Synechocystis* sp. Pasteur Culture Collection 6803. *Algal Res.* **2021**, *54*, 102180. [CrossRef]

15. Reinsvold, R.E.; Jinkerson, R.E.; Radakovits, R.; Posewitz, M.C.; Basu, C. The production of the sesquiterpene β-caryophyllene in a transgenic strain of the cyanobacterium *Synechocystis*. *J. Plant Physiol.* **2011**, *168*, 848–852. [CrossRef]

16. Begemann, M.B.; Zess, E.K.; Walters, E.M.; Schmitt, E.F.; Markley, A.L.; Pfleger, B.F. An Organic Acid Based Counter Selection System for Cyanobacteria. *PLoS ONE* **2013**, *8*, e76594. [CrossRef] [PubMed]

17. Berla, B.M.; Saha, R.; Immethun, C.M.; Maranas, C.D.; Moon, T.S.; Pakrasi, H.B. Synthetic biology of cyanobacteria: Unique challenges and opportunities. *Front. Microbiol.* **2013**, *4*, 246. [CrossRef] [PubMed]

18. Cheah, Y.E.; Albers, S.C.; Peebles, C.A.M. A novel counter-selection method for markerless genetic modification in *Synechocystis* sp. PCC 6803. *Biotechnol. Prog.* **2013**, *29*, 23–30. [CrossRef] [PubMed]

19. Doudna, J.A.; Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **2014**, *346*, 6213. [CrossRef]

20. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821. [CrossRef]

21. Miura, H.; Quadros, R.M.; Gurumurthy, C.B.; Ohtsuka, M. Easi-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors. *Nat. Protoc.* **2018**, *13*, 195–215. [CrossRef]

22. Ran, F.A.; Patrick, D.H.; Jason, W.; Vineeta, A.; David, A.S.; Feng, Z. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **2013**, *8*, 2281. [CrossRef]

23. Sternberg, S.H.; Redding, S.; Jinek, M.; Greene, E.C.; Doudna, J.A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **2014**, *507*, 62–67. [CrossRef] [PubMed]

24. Davis, L.; Maizels, N. Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E924–E932. [CrossRef] [PubMed]

25. Deltcheva, E.; Chylinski, K.; Sharma, C.M.; Gonzales, K.; Chao, Y.; Pirzada, Z.A.; Eckert, M.R.; Vogel, J.; Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **2011**, *471*, 602–607. [CrossRef]

26. Overballe-Petersen, S.; Harms, K.; Orlando, L.A.A.; Mayar, J.V.M.; Rasmussen, S.; Dahl, T.W.; Rosing, M.T.; Poole, A.M.; Sicheritz-Ponten, T.; Brunak, S.; et al. Bacterial natural transformation by highly fragmented and damaged DNA. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19860–19865. [CrossRef] [PubMed]

27. Anzalone, A.V.; Koblan, L.W.; Liu, D.R. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **2020**, *38*, 824–844. [CrossRef] [PubMed]

28. Ran, F.A.; Hsu, P.D.; Lin, C.Y.; Gootenberg, J.S.; Konermann, S.; Trevino, A.E.; Scott, D.A.; Inoue, A.; Matoba, S.; Zhang, Y.; et al. Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. *Cell* **2013**, *154*, 1380–1389. [CrossRef]

29. Barrangou, R.; Marraffini, L.A. CRISPR-cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell* **2014**, *54*, 234–244. [CrossRef] [PubMed]

30. Jinek, M.; East, A.; Cheng, A.; Lin, S.; Ma, E.; Doudna, J. RNA-programmed genome editing in human cells. *eLife* **2013**, *2013*, e00471. [CrossRef] [PubMed]

31. Ramey, C.J.; Barón-Sola, Á.; Aucoin, H.R.; Boyle, N.R. Genome Engineering in Cyanobacteria: Where We Are and Where We Need to Go. *ACS Synth. Biol.* **2015**, *4*, 1186–1196. [CrossRef] [PubMed]

32. Hsu, P.D.; Lander, E.S.; Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **2014**, *157*, 1262–1278. [CrossRef] [PubMed]

33.   Labuhn, M.; Adams, F.F.; Ng, M.; Knoess, S.; Schambach, A.; Charpentier, E.M.; Schwarzer, A.; Mateo, J.L.; Klusmann, J.-H.; Heckl, D. Refined sgRNA efficacy prediction improves large- and small-scale CRISPR-Cas9 applications. *Nucleic Acids Res.* **2018**, *46*, 1375–1385. [CrossRef] [PubMed]

34.   Wright, A.V.; Nuñez, J.K.; Doudna, J.A. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell* **2016**, *164*, 29–44. [CrossRef]

35.   Biot-Pelletier, D.; Martin, V.J.J. Seamless site-directed mutagenesis of the Saccharomyces cerevisiae genome using CRISPR-Cas9. *J. Biol. Eng.* **2016**, *10*, 6. [CrossRef]

36.   Choi, K.R.; Lee, S.Y. CRISPR technologies for bacterial systems: Current achievements and future directions. *Biotechnol. Adv.* **2016**, *34*, 1180–1209. [CrossRef] [PubMed]

37.   Heidi, L. CRISPR babies. *Nature* **2019**, *570*, 293–296. [CrossRef]

38.   Li, H.; Shen, C.R.; Huang, C.H.; Sung, L.Y.; Wu, M.Y.; Hu, Y.C. CRISPR-Cas9 for the genome engineering of cyanobacteria and succinate production. *Metab. Eng.* **2016**, *38*, 293–302. [CrossRef] [PubMed]

39.   Wendt, K.E.; Ungerer, J.; Cobb, R.E.; Zhao, H.; Pakrasi, H.B. CRISPR/Cas9 mediated targeted mutagenesis of the fast growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Microb. Cell Factories* **2016**, *15*, 115. [CrossRef] [PubMed]

40.   Ungerer, J.; Pakrasi, H.B. Cpf1 Is A Versatile Tool for CRISPR Genome Editing Across Diverse Species of Cyanobacteria. *Sci. Rep.* **2016**, *6*, 39681. [CrossRef] [PubMed]

41.   Xiao, Y.; Wang, S.; Rommelfanger, S.; Balassy, A.; Barba-Ostria, C.; Gu, P.; Galazka, J.M.; Zhang, F. Developing a Cas9-based tool to engineer native plasmids in *Synechocystis* sp. PCC 6803. *Biotechnol. Bioeng.* **2018**, *115*, 2305–2314. [CrossRef]

42.   Cengic, I.; Cañadas, I.C.; Minton, N.P.; Hudson, E.P. Inducible CRISPR/Cas9 allows for multiplexed and rapidly segregated single-target genome editing in *Synechocystis* sp. PCC 6803. *ACS Synth. Biol.* **2022**, *11*, 3100–3113. [CrossRef] [PubMed]

43.   Pinto, F.; Pacheco, C.C.; Oliveira, P.; Montagud, A.; Landels, A.; Couto, N.; Wright, P.C.; Urchueguía, J.F.; Tamagnini, P. Improving a *Synechocystis*-based photoautotrophic chassis through systematic genome mapping and validation of neutral sites. *DNA Res.* **2015**, *22*, 425–437. [CrossRef]

44.   Alonso-Lerma, B.; Jabalera, Y.; Samperio, S.; Morin, M.; Fernandez, A.; Hille, L.T.; Silverstein, R.A.; Quesada-Ganuza, A.; Reifs, A.; Fernández-Peñalver, S. Evolution of CRISPR-associated endonucleases as inferred from resurrected proteins. *Nat. Microbiol.* **2023**, *8*, 77–90. [CrossRef] [PubMed]

45.   Takasugi, P.R.; Wang, S.; Truong, K.T.; Drage, E.P.; Kanishka, S.N.; Higbee, M.A.; Bamidele, N.; Ojelabi, O.; Sontheimer, E.J.; Gagnon, J.A. Orthogonal CRISPR-Cas tools for genome editing, inhibition, and CRISPR recording in zebrafish embryos. *Genetics* **2022**, *220*, iyab196. [CrossRef]

46.   Kenney, C.T.; Marraffini, L.A. Rarely acquired type II-A CRISPR-Cas spacers mediate anti-viral immunity through the targeting of a non-canonical PAM sequence. *Nucleic Acids Res.* **2023**, *51*, 7438–7450. [CrossRef]

47.   Wang, J.; Teng, Y.; Gong, X.; Zhang, J.; Wu, Y.; Lou, L.; Li, M.; Xie, Z.-R.; Yan, Y. Exploring and engineering PAM-diverse *Streptococci* Cas9 for PAM-directed bifunctional and titratable gene control in bacteria. *Metab. Eng.* **2023**, *75*, 68–77. [CrossRef] [PubMed]

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

49.   Jiang, W.; Bikard, D.; Cox, D.; Zhang, F.; Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **2013**, *31*, 233–239. [CrossRef]

50.   Liu, Z.; Shan, H.; Chen, S.; Chen, M.; Song, Y.; Lai, L.; Li, Z. Efficient base editing with expanded targeting scope using an engineered Spy-mac Cas9 variant. *Cell Discov.* **2019**, *5*, 58. [CrossRef]

51.   Guo, C.; Ma, X.; Gao, F.; Guo, Y. Off-target effects in CRISPR/Cas9 gene editing. *Front. Bioeng. Biotechnol.* **2023**, *11*, 1143157. [CrossRef] [PubMed]

52.   Rao, X.; Zhao, H.; Shao, C.; Yi, C. Characterizing off-target effects of genome editors. *Curr. Opin. Biomed. Eng.* **2023**, *28*, 100480. [CrossRef]

53.   Xiao, A.; Cheng, Z.; Kong, L.; Zhu, Z.; Lin, S.; Gao, G.; Zhang, B. CasOT: A genome-wide Cas9/gRNA off-target searching tool. *Bioinformatics* **2014**, *30*, 1180–1182. [CrossRef]

54.   Larson, M.H.; Gilbert, L.A.; Wang, X.; Lim, W.A.; Weissman, J.S.; Qi, L.S. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* **2013**, *8*, 2180–2196. [CrossRef]

55.   Albers, S.C.; Peebles, C.A. Evaluating light-induced promoters for the control of heterologous gene expression in *Synechocystis* sp. PCC 6803. *Biotechnol. Prog.* **2017**, *33*, 45–53. [CrossRef] [PubMed]

56.   Kojima, S.; Okumura, Y. Outer membrane-deprived cyanobacteria liberate periplasmic and thylakoid luminal components that support the growth of heterotrophs. *BioRxiv* **2020**, preprint. [CrossRef]

57.   Yao, L.; Cengic, I.; Anfelt, J.; Hudson, E.P. Multiple gene repression in cyanobacteria using CRISPRi. *ACS Synth. Biol.* **2016**, *5*, 207–212. [CrossRef]

58. Pecoraro, V.; Zerulla, K.; Lange, C.; Soppa, J. Quantification of ploidy in proteobacteria revealed the existence of monoploid, (mero-)oligoploid and polyploid species. *PLoS ONE* **2011**, *6*, e16392. [CrossRef] [PubMed]

59. Zerulla, K.; Ludt, K.; Soppa, J. The ploidy level of *Synechocystis* sp. PCC 6803 is highly variable and is influenced by growth phase and by chemical and physical external parameters. *Microbiology* **2016**, *162*, 730–739. [CrossRef] [PubMed]

60. Englund, E.; Andersen-Ranberg, J.; Miao, R.; Hamberger, B.; Lindberg, P. Metabolic Engineering of *Synechocystis* sp. PCC 6803 for Production of the Plant Diterpenoid Manoyl Oxide. *ACS Synth. Biol.* **2015**, *4*, 1270–1278. [CrossRef] [PubMed]

61. Gale, G.A.; Osorio, A.A.S.; Puzorjov, A.; Wang, B.; McCormick, A.J. Genetic modification of cyanobacteria by conjugation using the cyanogate modular cloning toolkit. *J. Vis. Exp.* **2019**, *152*, e60451. [CrossRef]