CRISPR-Cas9 System Allows Precise Nucleotide Changes in Yeast to Modify Genomes

Emily Wurst

Molecular Biology and Evolution, Volume BIMM 101, Issue 1, March 2022
Published: 11 March 2022

Abstract

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) endonuclease in bacteria and archaea are a natural defense mechanism against viral infection. This research paper uses Type II CRISPR-Cas9 system and lab-designed gRNA and HDR template to mutate *Saccharomyces cerevisiae's* ADE2 gene, causing LoF and accumulation of red pigment in mutated yeast cells. NHEJ produces low transformation efficiency (31.82 colonies/ug of plasmid(c/ugP)) compared to the positive control (146.35c/ug) and red phenotype (21.3% of all transformed colonies) for *Saccharomyces cerevisiae.* However, we show that when CRISPR-Cas9 is guided by HDR through the design of and HDR template to cause point mutations, the transformation rate of *Saccharomyces cerevisiae* doubles and produces colonies with a red phenotype of nearly 97%. Through the use of Cas9 + gRNA + HDR, we can drastically improve the transformation efficiency and mutation rate of CRISPR mutagenesis.

Keywords: CRISPR, Cas9, HDR, NHEJ, genome editing, molecular biology
Issue: Original Research

Introduction

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) are repetitive DNA sequences observed in bacteria. Between the CRISPR sequences, bacteria genomes contain “spacer” DNA that exactly matches viral sequences (Jackson, et al 2011). Upon viral infection, bacteria transcribe these CRISPR loci (CRISPR plus “spacer” DNA) to small RNAs (crRNAs/gRNAs) (Terns, et al 2011). crRNAs guide CRISPR-associated (Cas) endonucleases to the invading target genome effectively recognizing then destroying the invading DNA through double-strand breaks (DSBs) and subsequent mutation through error-prone non-homologous end joining (NHEJ) repair (Hsu, et al 2014).

Researchers realized that this adaptive bacterial immune system could be exploited to create mutagenic DNA through the manipulation of gRNA and the use of Cas-9. By engineering synthetic gRNA independently from

Fig. 1

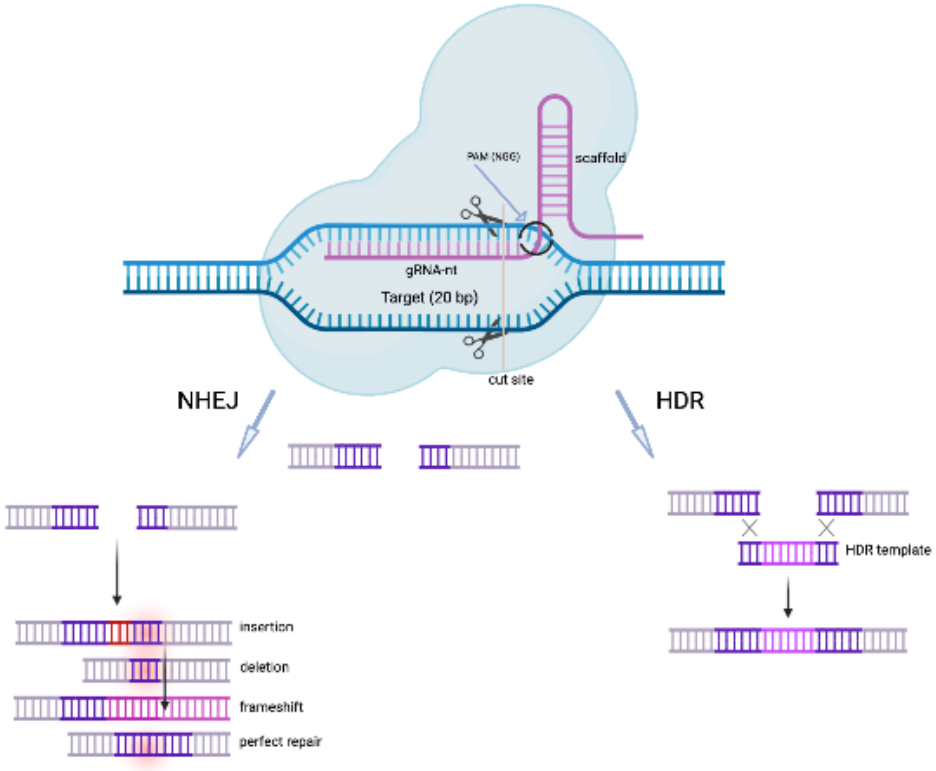


Figure 1 illustrates the Cas9 + gRNA system and two repair methods (NHEJ and HDR) both of which are explored in this paper. gRNA is an RNA copy of the target sequence (20bp). Attached to the gRNA is a scaffolding sequence. The gRNA leads Cas9 to the DNA complimentary to the target sequence, and cuts 3 bps 5' to the PAM sequence. After this double straded cut by Cas9, the cell will go through NHEJ where joining will either result in insertion, deletion, frameshift, or perfect repair. If a homologous template is provided, the DNA will repair itself with the template resulting in HDR.

Cas-9, researchers are able to improve CRISPR's specificity and control for off-target (Allen, et al 2021). However, manipulating gRNA alone does not ensure precise mutations (Figure 1)–researchers must control for repair as well. Because natural non-homologous end-joining is imprecise, we introduced homology-directed repair (HDR). HDR is a system or repair that uses two DNA fragments 25-50bps long (HDR template) on either side of the gRNA sequence to direct repair (Figure 1), resulting in higher mutation accuracy (Wang, et al). Instead of possibly having deletion, insertion, or frame shift, researchers can introduce specific point mutations outside the 20bp gRNA and even insert engineered sequences to produce mutations (Sander, et al).

For our experiment, we modified the pML104 plasmid containing DNA for Cas9 and a unique gRNA to mutate and cause LoF of the ADE2 gene in Saccharomyces cerevisiae (yeast, strain BY4743). We designed our gRNA to cause a nonsense mutation introducing a premature stop codon effectively stopping gene translation of ADE2, and causing a red phenotype in transformed yeast cells.

Additionally, we compared the results of NHEJ and HDR. We designed an HDR template that introduces five mutations (of six bps total), four of which would cause an early stop codon ending mutation. In addition, our last mutation was designed to mutate the PAM sequence to stop further Cas9 interference and ensure. Through PCR, we amplified out HDR template and ligated it to our plasmid, transforming our yeast. To ensure that our CRISPR-Cas9 system effectively produced the desired mutations, we sequenced each phenotype for each transformation.

Through our experiment, we have shown that the CRISPS+Cas9 complex in addition to HDR is an effective and precise gene mutation method for creating loss of function (LoF) in the ADE2 gene of Saccharomyces cerevisiae

Results

CRISPR-Cas9 NHEJ Mutagenesis

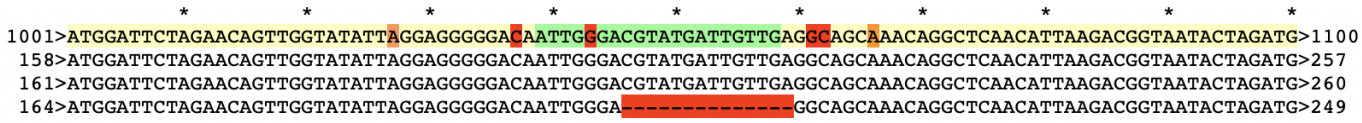
The ADE2 gene was targeted with gRNA. To understand the CRISPR-Cas9 mutation success with a cell's natural non-homologous end-joining, no HDR sequence was added to aid direct repair. Thus, gRNA targeted G→T turning glycine into a terminator 50 bps from the start codon. With NHEJ, three scenarios are possible: perfect repair, insertion of small insertions, or small deletions (Figure 1). The bottom sequence of Figure 2A shows a small deletion which caused this colony's ADE2 gene to lose function. If NHEJ were to cause deletion or insertion that terminates or frameshifts away from the ORF of the ADE2 gene, the mutation would be successful and red colonies would be observed. But, NHEJ is also likely to correctly repair (Figure 2A) the sequence resulting in a failed mutation and white colonies. Additionally, because the NHEJ repair is error-prone, it is very likely to see a lower transformation rate due to cell death.

Out results found that while NHEJ was able to result in some mutations and produce a red phenotype, it was much more likely to correctly repair the broken DNA and result in white colonies. Our transformation efficiency is 31.82 avg c/ug P(regardless of phenotype) (Figure 3A). Compared to the positive control, pML104 + gRNA was only able to transform about 21.7%. This is relatively low.

Fig. 2

2.A

2.B



2.C

Figure 2A shows ADE2 gene on the top strand, gRNA + HDR on the middle strand, and gRNA no HDR on the bottom strand. As shown, gRNA alone implements one mutation while gRNA + HDR implements five mutations (over six nucleotides). Figure 2B shows the original ADE2 gene (unmutated) and three genes sequences after NHEJ. The two middle sequences are from a white colony and the sequence reflects perfect repair. The bottom sequence came from a red colony and the deletion (shown in red on the alignment) shows where the mutation causing the red phenotype happened.

The red phenotype—indicating successful mutation of ADE2—was observed 21.3% of all pML104 + HDR colonies and white was observed for the remaining 78.7% of colonies (Figure 3B). This low transformation efficiency and low red phenotype percentage indicates that NHEJ is an inefficient repair system. Ultimately, to use CRISPR with an amount of accuracy for mutation and high transformation efficiency, other repair methods should be explored.

Mutations were confirmed with sanger sequencing (Figure 2B).

CRISPR-Cas9 HDR Mutagenesis

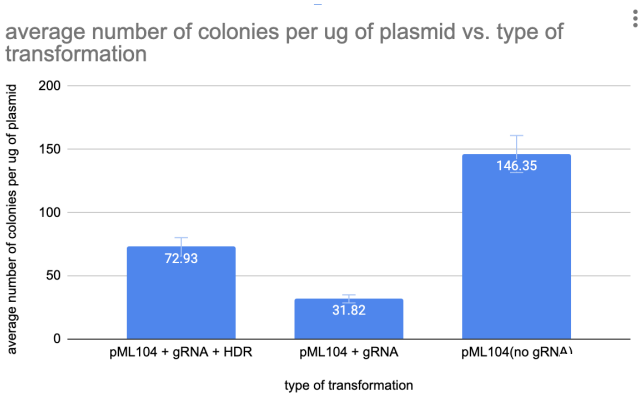
The ADE2 gene was targeted with gRNA + HDR template. The gRNA plus HDR template would not only target G→T, terminating ADE2 50bp with gRNA, but would also introduce four other mutations (three single nucleotides and one double nucleotide), three of which are terminator codons, creating a higher probability for successful LoF. A successful aspect of the gRNA + HDR design, in addition to its multiple stop codon, is the mutation/stop codon located 60 bps away from the beginning of the start codon at the original site for the PAM sequence (Figure 2C). Because gRNA + HDR causes multiple mutation and mutates the PAM sequence (preventing recutting and remutating by Cas9), it is more successful a higher percentage of LoF.

Class results confirm this data. The transformation efficiency was 72.93 c/ug P, double the 31.82 c/ng P observed in pML104 + gRNA (Figure 3B). Compared to the positive control, pML104 + gRNA + HDR has much lower efficiency (49.8% compared to pML104(no gRNA) (Figure 3A).

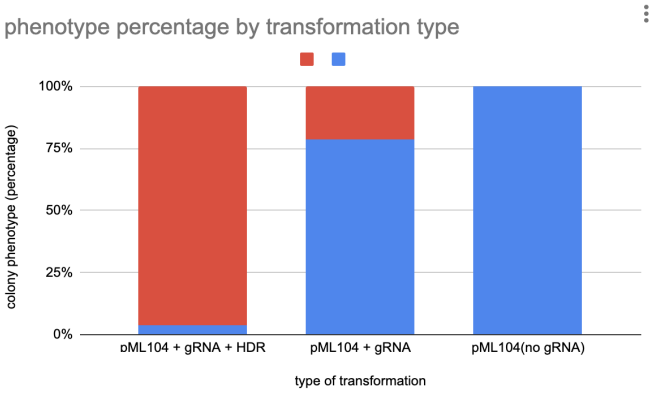
Class results also indicate that pML104 + gRNA + HDR has an extremely high red phenotype as 96.4% of all colonies from pML104 + gRNA + HDR were observed to be red (Figure 3B). This indicates that with HDR, the CRISPR+Cas9 system can correctly mutate nearly all yeast colonies transformed.

Mutations were confirmed with sanger sequencing (Figure 2C).

Fig. 3

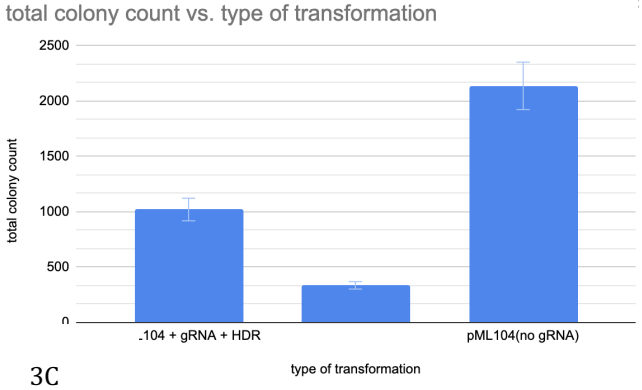


3A



3B

Figure 3 shows colony averages per transformation, phenotype percentages per transformation, and total colony count per transformation (note: negative control of water was omitted as it produced 0 yeast colonies as expected. Figure 3A shows that colony transformation was much lower for both pML104 + gRNA + HDR and pML104 + gRNA compared to pML104(no gRNA). Figure 3B representation that HDR has the highest mutation proportion compared to NHEJ and the control. Figure 3C gives colony counts for reader to compare to averages, giving perspective to 3A.



3C

Discussion

This paper explores the transformation efficiency and mutation rate for the CRISPR+Cas9 system, comparing NHEJ and HDR. Our paper found that while NHEJ can accurately cause a LoF for the ADE2 gene, it can only do so for around 20% of colonies. Additionally, NHEJ provides a low transformation, nearly ⅕ the efficiency of the control. On the other hand, HDR gives researchers much higher mutation rates. We have shown that an effectively designed gRNA + HDR template can turn near 97% of all transformed colonies red. This indicates that when a gene is targeted with Cas9 + gRNA + HDR, researchers can cause LoF in nearly all colonies transformed. However, despite the high LoF rate, HDR still shows a much lower transformation when compared to positive control (nearly ½ as many colonies transformed). This could be for a few reasons.

CRISPR-Cas9 could have targeted a mutation outside of the ADE2 gene causing cell death. In this experiment, we did not perform control transformation to check for off-target CRISPR-Cas9 activity. Although our gRNA was designed to target a unique sequence in the yeast genome, it is possible that there were genes unaccounted for in our design. Since we did not run an experiment, we cannot check. By mutating genes outside the target gene, the CRISPR-Cas9 system could have caused cell death. Additionally, if our gRNA + HDR complexes were being used to target off-target genes, even if these mutations did not cause cell death and resulted in an otherwise functional yeast, it could have resulted in a lower percentage of red colonies as not enough gRNA + HDR were added (Cho, et al). Additionally, the use of Sanger sequencing as opposed to genome-wide sequencing before transformation and after-transformation limits are the ability to deny the possibility of off-target mutation. If we were to perform CRISPR-Cas9 mutations on higher-level organisms, it would be worth the resources to genome-wide sequence the parent generations and prodigy to check for off-target mutations.

CRISPR-Cas9 is known to fail instances of nuclease-induced toxicity. This can occur in both on-target and off-target genes, and researchers posit this toxicity is due to DNA damage (Morgens, et al). Toxicity can result in cell death and lower transformation rates. It is very likely that the lower HDR transformation rate was due to CRISPR/Cas9 toxicity. This indicates that more studies should be run on DNA damage both during the Cas9 cutting and DNA Damage during HDR repair.

Another possible reason for low transformation efficiency/white colonies in the HDR group is failed mutations of the PAM sequence. If the PAM sequence fails to mutate, CRISPR-Cas9 can recut and yeast could repair itself, or the recut/repair could cause further toxicity and a lower transformation rate.

A large limitation of this study is the number of scientists running separate experiments (in-fashion with classroom laboratory science) and data reporting. Although instructions were the same for all students, it is ultimately individual lab pairs calculating amounts and running experiments independently. This most likely has caused issues in the data since there was not much control beyond initial instruction to normalize experimental design. Also, with more students running the same experiment once, rather than a small team running the experiment multiple times, errors are more likely to occur. Thus, when cleaning the data of the experiments, many outliers were excluded. When a colony count was 0 for white colonies and red colonies in both experimental groups and zero in the positive control and experimental groups, this data was removed to account for experimental error. Additionally, plates that displayed colonies outside the two yeast phenotypes (white or red) were scrubbed from the data due to contamination. Lastly, outliers for colony counts in the thousands were excluded from the data. This is because the authors of this particular report were unable to obtain proof of any plate with colonies of that count.

Another limitation was colony reporting. Colonies count/type were reported on a large spreadsheet accessed by over 70 students. In the spreadsheet, there were obviously incorrect data. For example, students may have marked putting 250ng of HDR in a sample they labeled to be without HDR. The data was scrubbed of these values with obvious mistakes. Despite these mistakes needing correction, there is no doubt that the errors (caught and uncaught) and cleaning of this data interferes with the accuracy of these results and presents a big limitation of this study.

If this study were to be replicated I would reccomen more control for experimental execution and checks on data reporting to give more accurate results.

Materials and Methods

pML104 plasmid design

The pML104 plasmid was sourced from *addgene.org*, pML104 has 11,247 bps. Plasmid pML104 has is grown in *E. coli*, ligated with our gRNA, than used as a vector to "infect" our yeast cells with to target and mutate the ADE2 gene. A few important design features of this plasmid are: promoters and origins of replication for both

prokaryotes (E. Coli) and eukaryotes (yeast), selectable genetic markers for both E. coli and yeast, BClI and SwaI restriction enzymes to clone gRNA, DNA encoding for Cas9, and DNA encoding for engineered gRNA.

Cultures in Lab/Plasmid Purification and Growth

Dam-/Dcm- E. coli (C2925) WT is not antibiotic resistant. The *E. coli* is used to grow pML104 plasmids. This strain is used because the *Dam-/Dcm-* genotype has no longer active methylases that would methylate the *BclI* site. This ensures that enzymes will be able to digest the *BclI* site to the gRNA sequence (and HDR) in the stuffer region of the plasmid (Marinus, et al).

Saccharomyces cerevisiae BY4743 is used since is auxotrophic and can be selected against using G418 (geneticin). Meaning we can control growth by plating on a uracil poor plate which we do after transformation.

Ecoli lysis and Plasmid Extraction

To extract plasmid DNA from *E. coli* we used the "QIAprep Miniprep" which lyses and denatures DNA at the same time. By doing this, then neutralizing the solution in acid and high salt buffers, the DNA is renatured. Once centrifuging the solution, the plasmid becomes the supernant and can be used for further experimentation. This leaves plasmids in either the supercoiled or nicked form, both forming distinct lines in a gel. The supernant is run via gel electrophoresis and two bands (one for supercoiled and one for nicked) is observed to confirm proper plasmid

gRNA and HDR Design

gRNA is complimentary RNA to a target sequence, thus when before designing gRNA it is necessary to locate the target sequence. The ADE2 gene start codon (ATG) is located 190 base pairs after ADE2's promoter region (TGACTC). Fifty-five base pairs after the start codon is the PAM sequence (NGG), in this case, arginine (AGG). Cas9 recognizes the PAM sequence triggering the double-helix to unwind for some of the DNA, making the DNA accessible to the gRNA. Thus, we designed the gRNA to target this PAM sequence fifty-five base pairs after the start codon to trigger a nonsense mutation that would enter a premature stop codon. A gRNA targeting a PAM sequence early after the start codon ensures that the gene will lose function before it is able to transcribe too many amino acids, thus making this an effective target sequence for LoF of ADE2.

Specifically, our gRNA would turn a G→T turning glycine into a terminator. Additionally we would introduce two other mutations using HDR. Two base pairs before the gRNA we turned C→T making glutamine a terminator (TAA) and two base pairs after the gRNA turning GC→TA turning alanine to a terminator (TAA) and mutating the PAM sequence. If either the mutation in the gRNA or one of the mutations introduced by HDR fail, the other two mutations would ensure the LoF for ADE2. However, most important is the mutation GC→TA because this mutation is a terminator while mutating the PAM, ensuring Cas9 will no longer cut that DNA.

Restriction enzyme to digest plasmids

Our designed gRNAs were added to the pML104 plasmid through restriction enzyme digestion at the SwaI and BclI cut sites, plasmid clean-up of double digested enzymes to remove "stuffer", and ligation of plasmid of gRNA oligos and double digested pML104 plasmid.

To add our designed gRNAs to the pML104 plasmid, four digestions were set up and executed: Swa1 only, *Bcl*I only, double digestion (Swa1 and *Bcl*I), and no digestion. Each digestion set up had 1ug of plasmid DNA, 10x NEBuffer 3.1 (5ul 1x), and up to 50 ul of nuclease-free water added to a tube along with 1ul (10 units) of the respective restriction enzyme (2ul for double digest and 0ul for the no digestion sample). (Note that plasmid DNA is given in grams, not concentration. This is weight, not concentration, was given so volume was incalculable since this experiment lab completed online). The tube was spun briefly then flicked to mix reagents to ensure that all reagents were present throughout the mixture and not stuck to the side of the tubes. The tube was incubated at 25C for 30-45 minutes. Then tubes were incubated at 50C for 15 mins.

5ul of each digest was added to a gel and checked for success. Although only double digestion plasmid is further ligated, the other digestions are performed and run on a gel to check for individual successful RE digestion.

Column clean-up

Prior to ligation, column clean-up was ran for digested plasmids. After RE digest, plasmids contain a "stuffer" region between the restriction enzyme sites that is removed during double digestion to add gRNA. Column cleanup to remove "stuffer" fragment and improve ligation results.

Although there were four digestions (Swa1 only, *Bcl*I only, double digestion (Swa1 and *Bcl*I), and no digestion) ran, we only will use double digestion for further experimentation and thus will only ligate double digested plasmid.

For column clean-up, firstly, elution buffer is warming at 37C to increase clean up product. Results are nanodropped to determine sample concentration.

**This portion of the lab was done online so nanodrop results are not given.

Ligation

After the plasmid is cleaned, the plasmid is then ligated with gRNA (and HDR template) using T4 DNA ligase which causes *in vivo* repair. Moles ratio of gRNA to plasmid are considered, and the amount of gRNA is increased (2:1).

After ligation, the results are ran on a gel to confirm that plasmids are in their linear form (indicating successful ligation and promising results that gRNA was added).

E. Coli transformation

The Dam-/Dcm- were previously treated by their curator (ZippyTM cells) to take up foreign DNA. Ligation product is added to the cells and they are incubated for 5 to 10 mins on ice then spread on LB (lysogeny broth) plate that is treated with ampicillin. Since the plasmid contains ampicillin resistance gene, only *E. coli* with a plasmid will take up gene.

After allowing for colony growth, a few colonies will go through PCR to ensure that colonies took up plasmid. First, primers are designed: the forward primer will anneal to the designed gRNA sequence and the reverse primer will anneal to the pML104 plasmid. The high temperature of the PCR will expose the *E.coli* DNA to the PCR, and the primers will simplify our sequence. Only colonies with pML104 and gRNA will anneal forward and backward. After, we ran the PCR product on a gel to ensure our plasmids with gRNA are present in E.coli.

After, we will lyse the remaining E. coli using "QIAprep Miniprep", and the supernant again becomes the plasmid to be used in yeast transformation.

In this step, PCR was also used to amplify HDR template which was ligated to plasmid and rePCRed to amplify gRNA + HDR.

Yeast Transformation

First, yeast is streaked on a YPD plate to grow colonies sufficient for transformation. After yeast has grown, transformations are set up. For this experiment, four transformations were used. Sample 1 was positive control: pML104(no gRNA) 1ug). The second transformation: pML104 + gRNA (ug), and was used to examine NHEJ efficiency and mutation rate. The third transformation: pML104 + gRNA (1ug) + HDR (250 ng), and was used to examine HDR efficiency and mutation rate. Lastly, forth transformation: no pML104 and instead sterile water, used to measure contamination.

After transformation mixtures were created and incubated, the mixtures were plated on a YC-ura plate, taking advantage of the plasmid's URA3 marker. After incubation, only yeast with the pML104 plasmid were observed.

Lastly, DNA was extracted and PCRed to magnify the ADE2 genome using primers. After, the PCR results were sent to be Sanger sequenced to show mutations and confirm phenotype to genotype.