

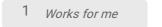


Aug 15, 2022

Student Guide

Brian Teague¹

¹University of Wisconsin - Stout



This document is published without a DOI.

Yeast ORFans CURE

Brian Teague University of Wisconsin - Stout

ABSTRACT

This is the student guide for the Yeast ORFans CURE. Associated protocols and an instructor guide are found in the **Yeast ORFans CURE** workspace.

Share

DOCUMENT CITATION

Brian Teague 2022. Student Guide. **protocols.io** https://protocols.io/view/student-guide-cew3tfgn

KEYWORDS

undergraduate, research, teaching, laboratory

LICENSE

This is an open access document distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

CREATED

Aug 05, 2022

LAST MODIFIED

Aug 15, 2022

DOCUMENT INTEGER ID

68283

ABSTRACT

This is the student guide for the Yeast ORFans CURE. Associated protocols and an instructor guide are found in the **Yeast ORFans CURE** workspace.



Introduction

The yeast *Saccharomyces cerevisiae* -- bakers' and brewers' yeast, the same organism that makes bread and beer -- has over 6,000 protein-coding genes. However, despite decades of study, over 600 of them still have no known function! From an evolutionary perspective, the fact that they've stuck around to make functional proteins suggest that they must have *some* function, confer some fitness advantage, but we don't know what that function is.

The goal of this research experience is to try to figure out what some of these genes do! The approach we'll take is called "reverse genetics." I like to explain it with a silly story, which follows:

"Imagine you are an alien who, after watching a bunch of Earth television from orbit, has a pretty good idea of how a car works. To learn more, you steal a car from a parking lot and are amazed to find that it is a hugely complicated machine with lots of smaller parts. How can you figure out what each part does? Well, one approach is to remove the part and see how that changes the car's function. For example, you could remove the wheels — and you'd discover that the wheels are pretty important to the function of the car! Or you could remove just the hubcaps — and you'd find that not much had changed.

What happens, though, if you remove the windshield wipers? If you were driving the car on a sunny day, it wouldn't have much impact. However, if you drove on a dark and stormy night, you would find that they were very important indeed -- but only under particular conditions."

That last point is particularly important. Most of the genes that are absolutely 100% required for yeast cells to grow have already been found. It is suspected that many of the others have *conditional* phenotypes – their importance only becomes apparent under particular growth conditions or (more likely) in response to particular stresses (just like the windshield wipers are only important for operating a car when it's raining.)

Looking at the big picture, then, we have three tasks.

- 1. Prepare the reagents for knocking out a gene: a plasmid (a short, circular piece of DNA) and a PCR amplicon (a linear piece of DNA).
- 2. Actually knock out the gene, and verify that it was knocked out successfully.
- 3. Perform an experiment (or perhaps many experiments!) to compare the knockout strain(s) to the wild-type strain, to see if there's an observable difference.

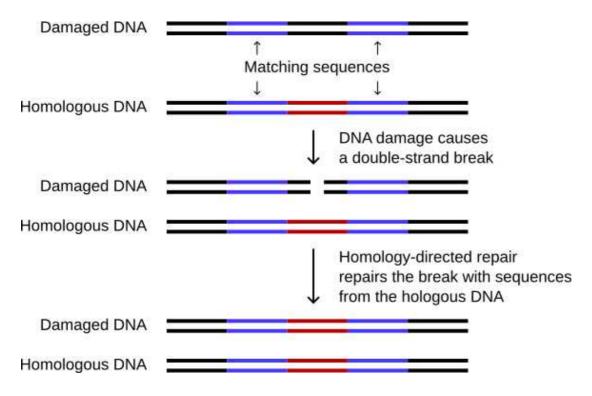
The rest of this guide focuses on tasks #1 and #2, only touching on #3 briefly. But before we can dive into the nitty-gritty of exactly how we'll knock out genes, we need to talk a little bit about some advanced molecular genetics — a topic called "homology-directed repair."

Homology-directed repair (HDR)

What does a (eukaryotic) cell do when its DNA is damaged? Often the cell just commits "cell suicide", a process called *apoptosis* or *programmed cell death*. (This is why, a few days after a bad sunburn, your skin starts to peel -- this is all of those cells identifying DNA damage and destroying themselves, rather than risk becoming cancerous.)



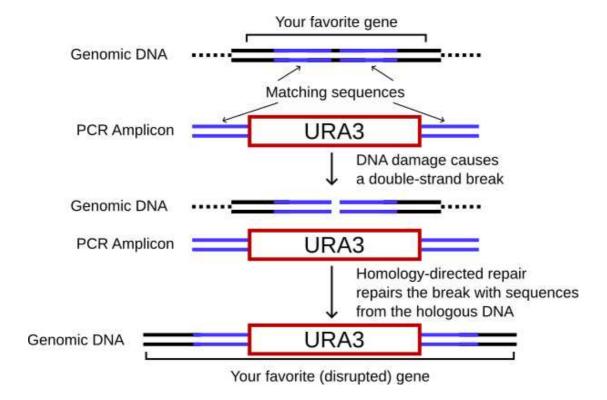
While apoptosis is common, it's not the only mechanism that cells use. Particularly if the DNA damage results in a complete break of both DNA strands, another option that the cell may use is *homology-directed repair (HDR)*. If there's a piece of DNA in the cell whose sequence is similar (or "homologous") to the damaged DNA, the the cell will use the homologous DNA to repair the damaged DNA molecule. The basics of the process are outlined in the figure below.



The key to homology-directed repair is the matching sequences on either side of the break -- they "tell" the cell which DNA to use to repair the damage. But this still leaves the question -- how could we use homology-directed repair to disable a gene?

Knocking out genes with HDR

In the following, I'll use "YFG" as short for "your favorite gene" -- ie, the gene you're trying to disrupt. Now imagine we took a big piece of completely unrelated DNA and inserted it right in the middle of YFG. As you might imagine, doing so has a pretty good chance of disrupting YFG. And that's exactly what we'll do -- using HDR and a piece of DNA we'll make using a method called the polymerase chain reaction, or PCR. There will be more about PCR below, but for the moment, assume we have made such a piece of DNA -- and on either end, there are homologous sequences that match sequences in YFG! The following figure illustrates what happens when we put that piece of linear DNA into a yeast cell:



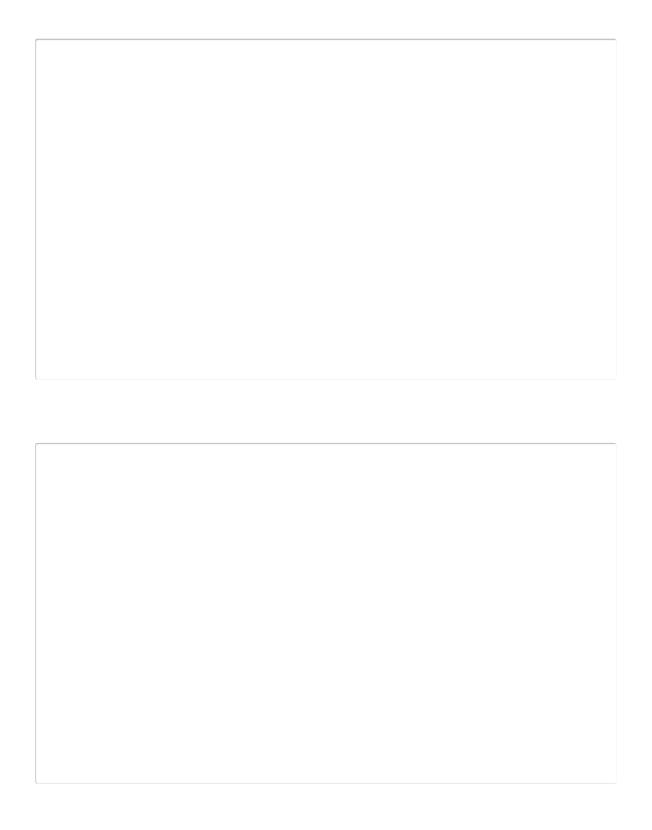
Now how this figure is similar to the last one -- because we used PCR to make a linear piece of DNA with sequences homologous to the region of the double-stranded break, the cell repairs the break with our PCR amplicon. (The piece it's repaired with is named URA3 -- it will be explained below.)

However, this process still assumes that a double-stranded break will happen -- is that a safe assumption? In general, the answer is *yes* -- there is enough random DNA damage that happens to yeast that scientists have been using this method to knock out yeast genes for decades. However, doing so is highly inefficient -- we could make it more efficient by *creating* a double-stranded break for the cell to repair (with our URA3 "patch"). Is there a way that we can point at a spot in the genome and say "I want a double-stranded break right there, please?"

Yes, there is. It's called CRISPR/Cas9.

Stimulating HDR with Cas9

You may have heard of CRISPR or Cas9 -- they are names for a technology that scientists use to edit DNA in living cells. The following two Youtube videos give a good overview of what CRISPR/Cas9 is and how it works:



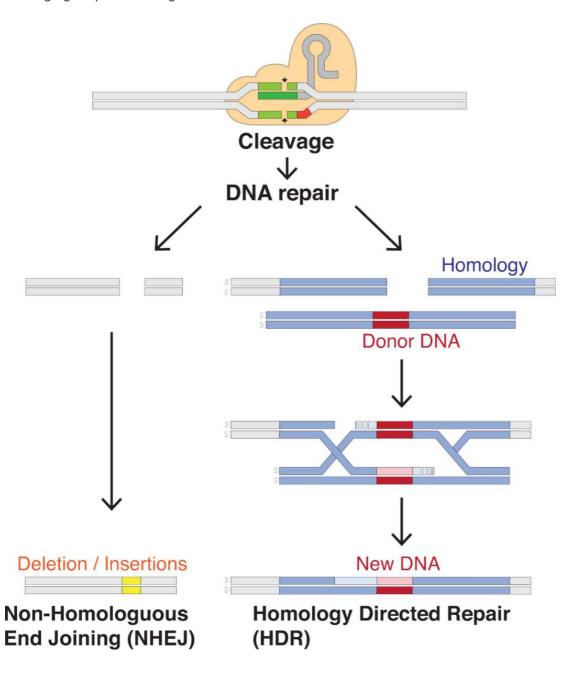
There are three key takeaways from those Youtube clips:

- 1. Cas9 is an enzyme that cuts DNA.
- 2. Cas9's guide RNA (gRNA) tells is what sequence of DNA to cut.
- 3. This makes Cas9 programmable -- change the gRNA and you change where Cas9 cuts.

Hopefully you're beginning to see how we could use Cas9 in our scheme to disable a gene in S.



cerevisiae. Using a PCR amplicon to knock out a gene depends on a double-stranded break to initiate homology-directed repair. And Cas9 lets us point at a spot in the genome and say "I would like a double stranded break right here please." If we (a) create DNA damage in a gene, and (b) give the cell a piece of DNA to repair that damage with, the cell will insert our "patch" and permanently disable the gene. The following figure puts it all together:



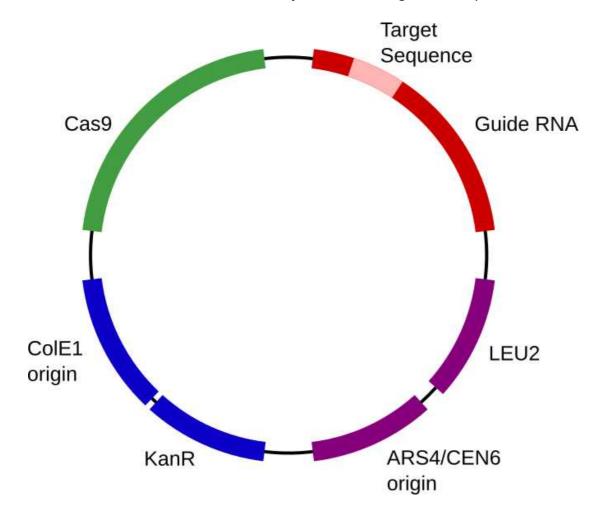
Mariuswalter, CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0, via Wikimedia Commons

By the way, why do we need a "patch" DNA -- the red and blue strand labeled "Donor DNA" in the above picture -- anyway? As the left side shows, just cutting the DNA will usually cause damage. There are two reasons. First, simply cutting the DNA in a gene is not always enough to damage the gene. The cell tries to repair the gene -- and sometimes it succeeds! If not, the damage is usually

slight -- a few extra or deleted base pairs -- which sometimes isn't enough to disable the gene. And second, we need a way to tell which cells have the disabled gene and which don't. The processes we use to get DNA into cells are extremely inefficient, so we need to be able to *select* for cells that have been modified. That's what the URA3 in the "patch" is for -- more details below.

Getting Cas9 and a guide RNA into yeast cells - with plasmids!

Here's another important point that you might have missed in the videos about Cas9. Yeast don't make the Cas9 protein. (The one that we're using comes from a bacterium called Streptococcus pyogenes.) To get Cas9 and a guide RNA to target your favorite gene, we need to give the yeast cells instructions for making them. We encode these instructions on a plasmid—a small circular piece of DNA that we can build in a test tube, then insert into the yeast cells. A diagram of the plasmid is below.



Each colored box is a functional sequence — and there are a lot of them! Ignore the blue and purple ones for a moment, and note that the green sequence has instructions for the cell to make the Cas9 protein, and the red sequence is instructions for the cell to make the guide RNA. In lighter red, I've noted the target sequence. Once you've made this plasmid, you'll put it into the yeast cell, and the yeast cell will make the Cas9 protein and a guide RNA that targets the Cas9 to your favorite gene.

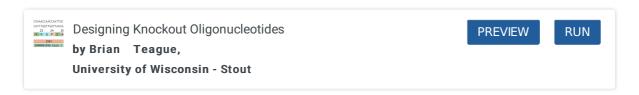
Remember, though, that the target sequence is different for every gene that we want to target. So if

each group is targeting a different gene, then each group will need to make a different plasmid. Fortunately, only the target sequence needs to change -- it's just 20 basepairs (20 DNA "letters"). Which begs the question -- how are we actually going to modify those bases?

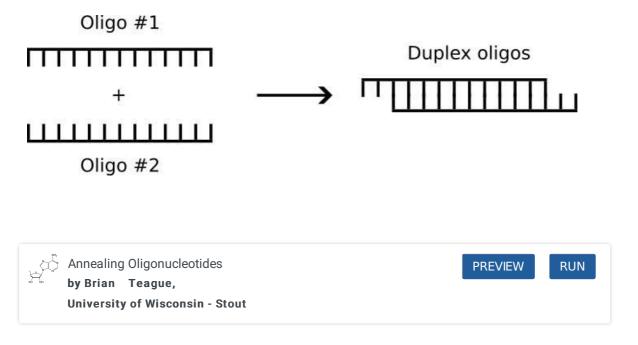
Customizing the plasmid: annealing and ligation

Fortunately, this plasmid has been designed to be easily customizable. Doing so will require three steps.

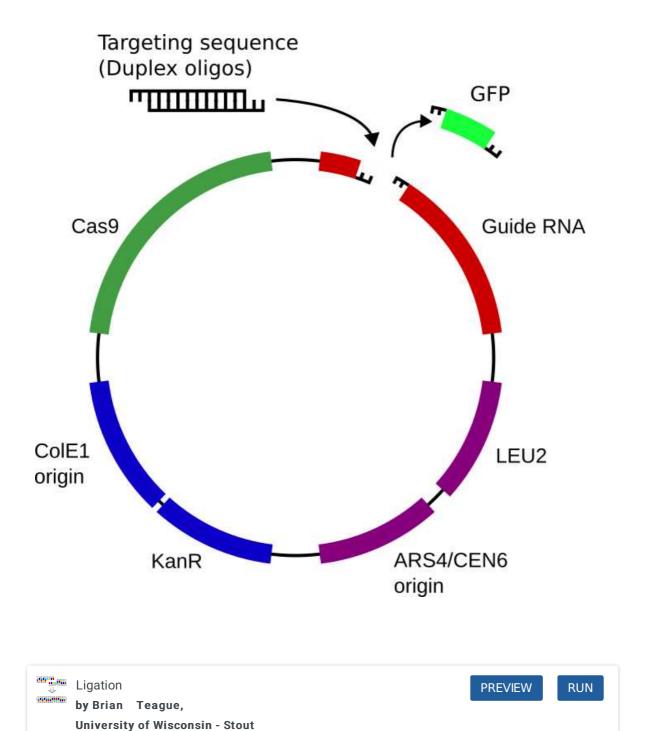
First, you'll need to design two short pieces DNA, which your instructor will order. The protocol linked below walks you through how to do so. The pieces will have the 20-base targeting sequences, with a few bases on either side that will help it "stick" to the plasmid backbone.



Second, we'll need to *anneal* the oligonucleotides together. Most DNA that is made by cells is double-stranded -- but the DNA molecules that are chemically synthesized are single-stranded. You designed both strands, so now we need to "stick" them together into a short double-stranded piece of DNA we'll call a "duplex."



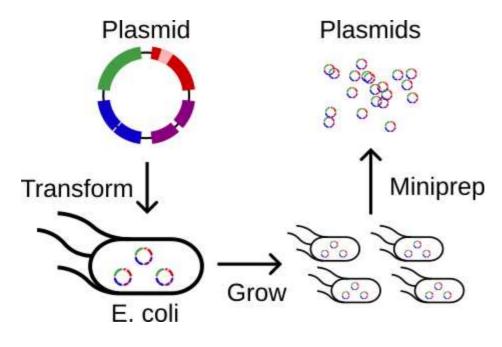
Finally, we'll *ligate* the duplex into the plasmid. To do so, the circular plasmid is first "cut" open with a (different) DNA-cutting enzyme that doesn't make a "clean" cut. Instead, it leaves a few bases of single-stranded DNA on each side of the cut -- bases that exactly match the "sticky" single-stranded ends on the duplex. You'll mix the cut-open plasmid backbone, your annealed oligos, and a *ligase* enzyme that will "glue" them together -- and an hour later, you'll have a complete, circular plasmid with your target sequence in the right place in the guide RNA.

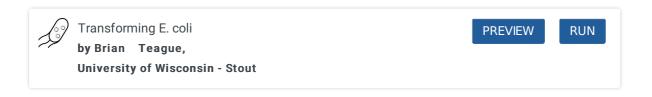


Amplifying the plasmid

There are two problems with the plan above to "customize" your plasmid. The first is that it doesn't make very much DNA -- you're only starting with about $\Box 50$ ng (nanograms) of DNA, while we need at least $\Box 1$ μg for future steps. And second -- the ligation doesn't always work! Sometimes the sticky ends of the plasmid just stick back to eachother. Also, the place where the plasmid was cut open has a piece of DNA in it called "GFP" that gets cut out -- sometimes the GFP goes back in instead of your duplex.

We'll solve the first problem -- "not enough DNA" -- by *cloning* your plasmid. The key idea here is that we can use a common laboratory bacterium called *E. coli* as a kind of "DNA copier." To make more plasmid DNA, we'll put the plasmid in the bacteria E. coli (called *transformation*), grow a bunch of them, and then purify the plasmid DNA back out (a protocol called a *miniprep*.)



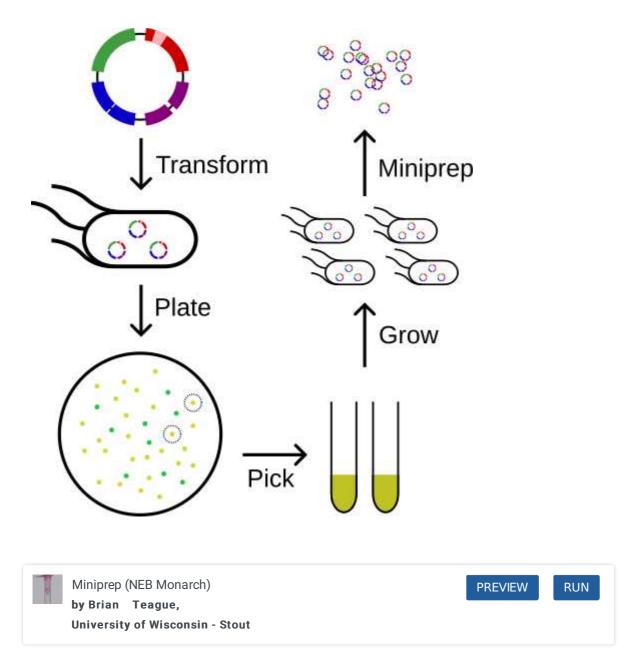


Unfortunately, transformation is *extremely* inefficient. If you try to transform a billion E. coli, you might only get 100 that have your plasmid in them — and those are the only ones we want to grow and miniprep. We need a way to *select* for just the successfully transformed E. coli, and we do by making the transformed E. coli antibiotic resistant.

Look back at the plasmid map, above, and pay attention to the blue boxes this time. The "ColE1 origin" is a sequence that tells E. coli, "copy this plasmid when you're copying your own genome." Thus, as the culture of E. coli grows, the number of plasmids increase too. The second box, "KanR", is a gene that makes an E. coli with the plasmid resistant to the antibiotic *kanamycin*. Now, imagine that after we transform our billion E. coli, we put them on a petri dish with growth media that contains kanamycin. The antibiotic kills the E. coli without the plasmid, but the small number that did pick it up survive and begin to multiply. After a night in the incubator, each individual cell has turned into a colony of cells on the petri dish. If your transformation works, it will look something like the image below:



Each "spot" on that petri dish is a colony of E. coli that started out as a single cell. You or your instructor will "pick" two of these colonies and grow them up overnight in test tubes containing some liquid broth (and some kanamycin). Under these growth conditions, the cells will divide every 30 minutes -- and the next day, you'll have *billions* E. coli, each of which contains 25-30 copies of your plasmid. Then, you can purify the DNA back out.



There's one additional benefit here. Recall that each colony started from *one single E. coli cell that picked up one single plasmid*. That cell doubled, and its children double, and doubled and doubled and doubled, until there was a colony that you could see with your naked eye -- but they all started from *one* cell. That means that, as long as the DNA copying process was very accurate (and it is), that all of E. coli from a single colony are *clones* -- they are genetically identical. It also means that all of the molecules of plasmid DNA that you get out of the miniprep are identical clones as well. This is why the process is called *cloning*.

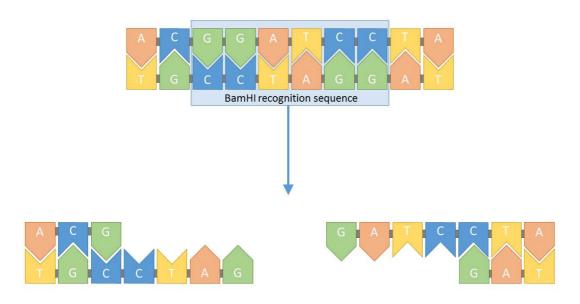
Why is this useful? Remember that we have also have a second issue -- that the process of ligation doesn't always work. Look carefully at the diagram above and note that some of the colonies on the petri dish are green. Remember that the duplex with our targeting sequence is replacing a GFP gene. "GFP" stands for "green fluorescent protein", which means that those are colonies where the GFP gene ligated back into the plasmid backbone -- or it wasn't properly cut out in the first place -- glow bright green. Those plasmids don't have the right instructions to make the yeast cell express a guide RNA, so we don't want them! Instead, we want "white" colonies, the ones that aren't glowing green.

Still, whether the colony glows green or not is not a whole lot of evidence that the plasmid we purified with the miniprep is the plasmid we wanted. What additional evidence can we get about whether our DNA manipulations worked?

Verifying the plasmid with a restriction digest

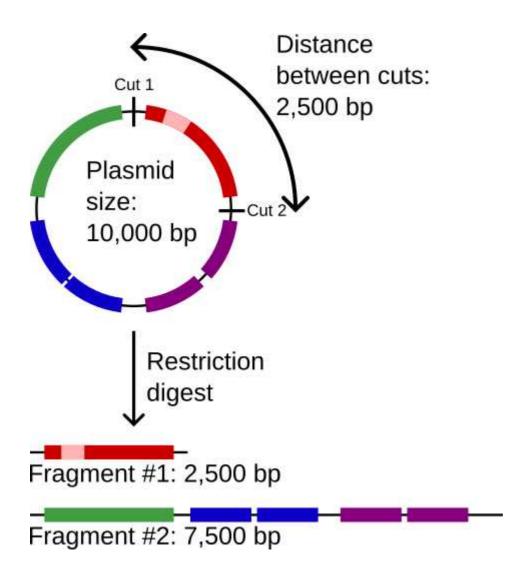
The best answer to the question "does this plasmid have the DNA sequence I think it does?" would be to sequence it. And while the cost of doing so has come down, it's still prohibitively expensive for routine DNA manipulations, particularly in a teaching laboratory.

So instead, we'll do a *diagnostic restriction digest*. Cas9 is not the only enzyme that cuts DNA -- in fact, hundreds of them are known and sold as purified protein by several biotechnology companies. Cas9 is special, though, because it's *programmable*. Most DNA-cutting enzymes (called "restriction endonucleases") cut a specific, non-changeable sequence, anywhere from 4 to 8 bases long. For example, the restriction endonuclease BamHI recognizes the sequence GGATCC and cuts it as depicted below:

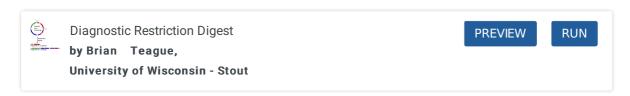


Simon Caulton, CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0, via Wikimedia Commons

How can we use a restriction enzyme like BamHI to give us some more evidence that the plasmid we have is the plasmid we want? The answer hinges on an important observation - we know the DNA sequence of the plasmid that we're trying to construct, and we know the DNA sequence that these restriction enzymes cut. So, we can use this information to predict where the enzyme will cut on the plasmid, and thus we can predict the lengths of the DNA fragments that will result if the enzyme cuts there. For example, if our plasmid is 10,000 bp long, and our enzyme has two cut sites 2,500 bp apart from eachother, we can predict the lengths of the resulting fragments:



Once we predict the fragment lengths (based on the known sequence of the plasmid and the predicted places our restriction enzyme will cut), then all we have to do is actually perform the digest — and see if the result matches our prediction.



Analyzing the restriction digest with agarose gel electrophoresis

The result of a restriction digest is a mixture of DNA fragments. How can we "see" those fragments to determine their size? By using *agarose gel electrophoresis*. Agarose gel works kind of like a "molecular sieve", separating pieces of DNA by size. We'll pipette the DNA mixture into a well in the gel, then apply an electric field. Because DNA is negatively charged, it moves through the gel towards the positive electrode — but not fragments migrate at the same rate. Instead, the bigger pieces get more "tangled up" in the gel than smaller pieces, and migrate more slowly. The result looks something like this:

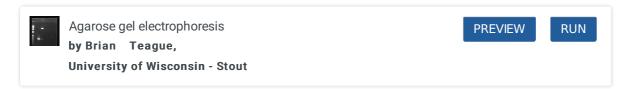


3 different restriction enzyme digests of plasmid DNA

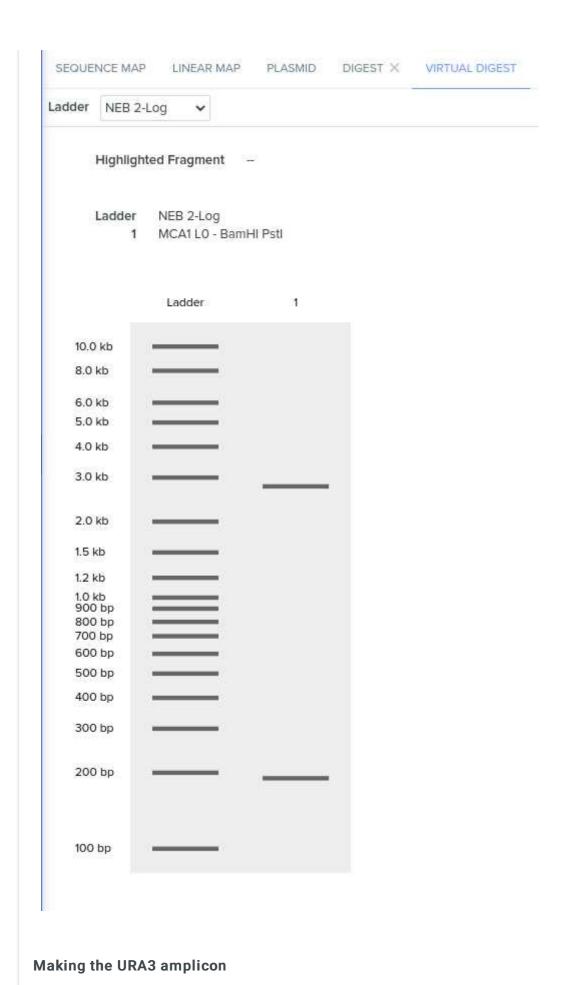
wells

DNA migration

Each lane of the gel is a separate sample -- the plasmid DNA in the first three lanes and a molecular weight standard in the fourth lane. The molecular weight standard is a commercial product, and importantly, we know the size (in bp) of each fragment in the standard. Thus, we can estimate the size of the unknown fragments in the digest that we're analyzing, and see if they match the ones that we predict that we'll see. If the bands on the gel are about the right size, that is good evidence that the plasmid DNA in our sample has the sequence that we expect it to have.

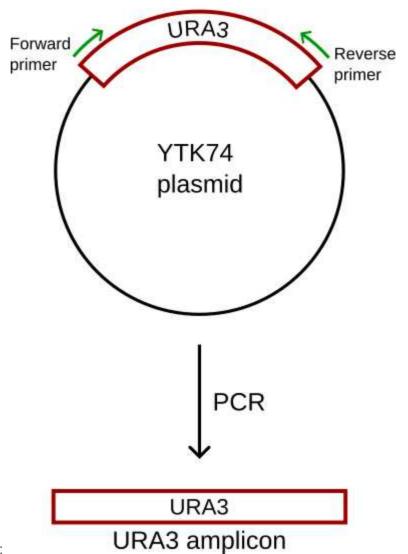


Modern molecular biology software such as Benchling will make this even easier — if you give it the sequence of the plasmid and the enzyme you're using, it will simulate the restriction digest *and the gel,* giving you an image such as the one below:



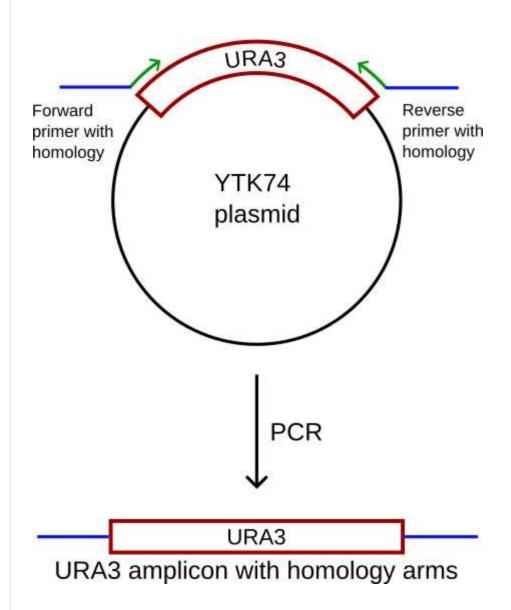
So now we have a plasmid to put in our yeast cells, one that instructs the cells to make the Cas9 protein and a guide RNA that targets it to our gene. The Cas9 protein will make a double-stranded break in the cell's DNA, but we still need the URA3 "patch" for the cells to repair the double-stranded break with. We'll make it with a method called the <i>polymerase chain reaction</i> , or PCR. Here's a quick video overview:
As the video shows, we need four things for a PCR: The template - the DNA that we want to copy Primers to specify what sequence of DNA to copy A polymerase enzyme, to actually do the copying A bunch of nucleotides to make the new copies of DNA.
The polymerase enzyme and the nucleotides are all in a commercial "master mix", so we'll take them as a given. What template are we using, and what primers? Well, as we saw above, we want our amplicon to contain the URA3 gene (whose purpose will become clear below), which means we need a template with the URA3 gene. And in fact, we have a plasmid that carries the URA3 gene.

need a template with the URA3 gene. And in fact, we have a plasmid that carries the URA3 gene, called YTK74. We'll use that, and a set of primers that bind to the plasmid before and after the

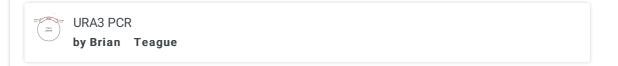


URA3 gene, to amplify it with PCR:

However, PCR isn't only good for *amplifying* (copying) DNA sequences – it can also *modify* the DNA sequence as it's being amplified. For example, if you add nucleotides to the ends of the primers, those nucleotides will be added on to the ends of the amplicon, as so:



And *this* is how we can "customize" our URA3 patch to disable your favorite gene. The oligo design protocol walks you through designing these PCR primers so that your URA3 amplicon will end up with ends ("homology arms") that are the same as the sequences that flank the cut that Cas9 will make. Then, when Cas9 makes a double-stranded break there, the cell will repair the break with the URA3 "patch", (hopefully) completely disabling the gene.



Verifying the PCR with agarose gel electrophoresis

Unfortunately, nothing in molecular biology works 100% of the time -- so we need to make sure that

our PCR worked as expected. Here, we have two questions to ask: (1) did we get any DNA at all out of the PCR? And (2) is it the DNA we expect?

We could use a diagnostic restriction digest, like we did with our plasmid, to answer (2). However, as the video above points out, PCR is very specific, and we know the DNA sequence that we were trying to amplify. So, we'll just ask "is the DNA the *size* we expect?"

And of course, we can answer both of those questions using agarose gel electrophoresis. The theory and the protocol are identical to the one above.



Agarose gel electrophoresis

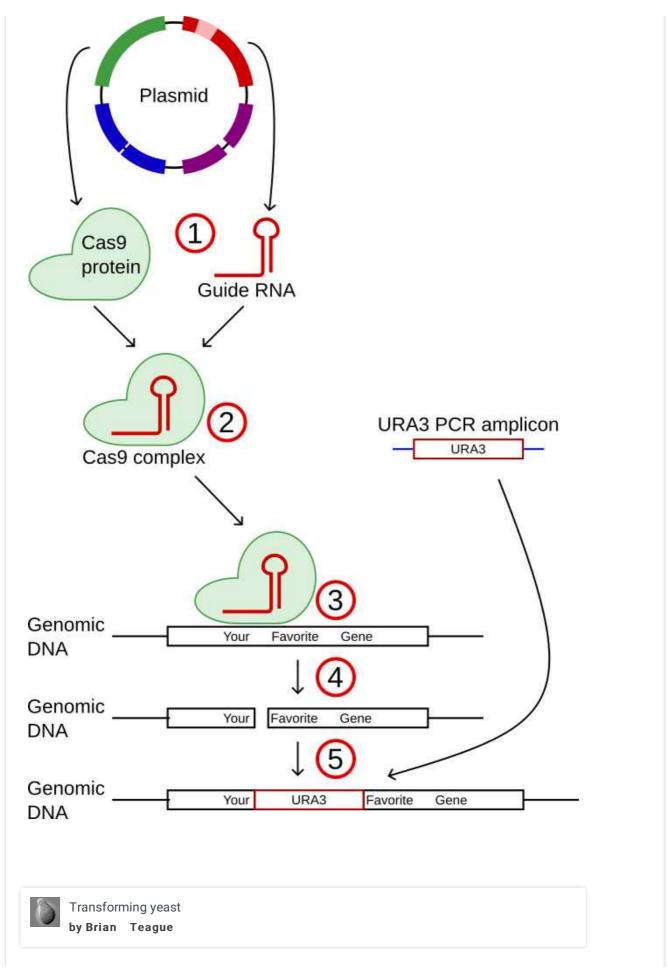
by Brian Teague

Transforming yeast to KNOCK OUT YOUR GENE!

It's finally time to knock out your gene! Using a process called "transformation", you'll put your plasmid and your URA3 PCR amplicon into yeast cells. If everything goes as its supposed to, the following will happen:

- 1. The plasmid will instruct the yeast cell to make the Cas9 protein and the guide RNA
- 2. The guide RNA will complex with the Cas9 protein, targeting it to your gene
- 3. The Cas9 protein will bind to the cell's DNA at your gene
- 4. ...and create a double-stranded break to stimulate homology-directed repair (HDR)
- 5. The cell will repair the break using the URA3 amplicon as "patch". This will completely disable the gene.

The process is depicted below:



Selecting for transformants

Remember how, when we transformed E. coli, I said that the process was extremely inefficient? The same is true here -- despite starting with many, many yeast cells, only a few will be transformed. How will we find them?

Recall also that the answer for E. coli was "kill the cells that didn't get transformed." Specifically, the media we grew our E. coli on had an antibiotic that was deadly for E. coli. Cells that took up our plasmid, however, became resistant to the antibiotic and could survive.

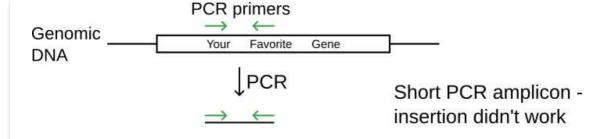
We have the same problem with our yeast transformation, but we're going to use a slightly different approach to solve it. Instead of spreading our yeast cells on media that contains an antibiotic, we're going to use media that *lacks a nutrient the cells need to live*. We'll use two different kinds of media, one that lacks *leucine* (an amino acid needed to make proteins) and another that lacks *uracil* (a base needed to make RNA). Cells that survive on these media have been successfully transformed — the LEU2 gene on the plasmid allows cells to make their own leucine, and the URA3 gene on the PCR amplicon allows cells to make their own uracil.

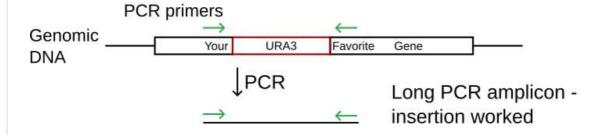
One other note – yeast grows much more slowly than E. coli. After two days you may see colonies, but it may take as long as three days. If there's nothing on the plate the morning after you transform, don't panic! After several days in the 30° incubator, the plates should like something like this:

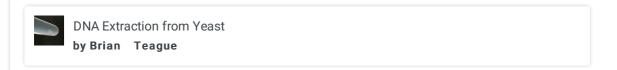


Verifying the knockout

As with every other manipulation, we need to verify that our knockout worked. (The fact that cells can make their own uracil only tells us that the URA3 gene went successfully into the genome -- it doesn't tell us where!) We'll do this by extracting the genomic DNA from your yeast cells and using PCR to amplify the region of the genome that URA3 *should* have gone into. Because we know the sequence of the genomic DNA there, we can predict how big the amplicon we should see if the knockout worked. The figure below shows how:

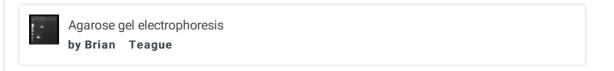






Knockout PCR
by Brian Teague

Once we've done the genomic DNA extraction and the PCR, we can analyze the PCR using agarose gel electrophoresis. Again, the two questions we want to answer are (1) did the PCR work? and (2) how long (approximately) is the PCR product? Because we know how long the URA3 amplicon is, and how far apart the PCR primers are in the unmodified genome, we can predict the length of the amplicon we'd expect to see if the knockout worked.



Testing for phenotypes

Good experimental design is beyond the scope of this document (: but I will drop a few important

ideas here:

Don't reinvent the wheel. There has been LOTS of work on yeast in the last 50 years -- leverage it! Here's a good, if somewhat dated, survey of yeast phenotypes:

Hampsey M (1997). A review of phenotypes in Saccharomyces cerevisiae.. Yeast (Chichester, England).

Run a pilot experiment or two. If you are testing a stressor like, say, salt -- do you know how much salt to use? You'll probably need to design a separate experiment to figure this out.

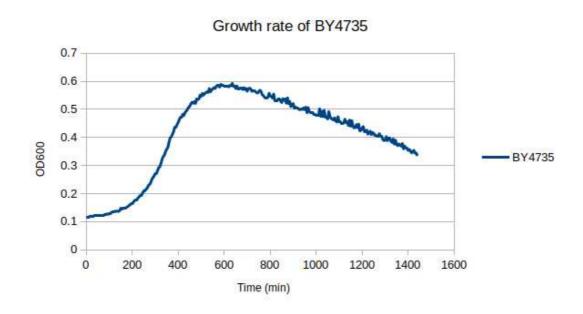
Only vary one thing at a time. Your final variable is "genotype" -- ie, whether a gene is knocked out (and if so, which gene.)

Be quantitative. Often, a phenotype experiment involves analyzing a growth curve. Don't just eyeball it -- derive some quantitative results! More information about analyzing a growth-rate phenotype below.

Analyzing a growth rate phenotype

"Is this knock-out strain more or less sensitive to a particular stress?"

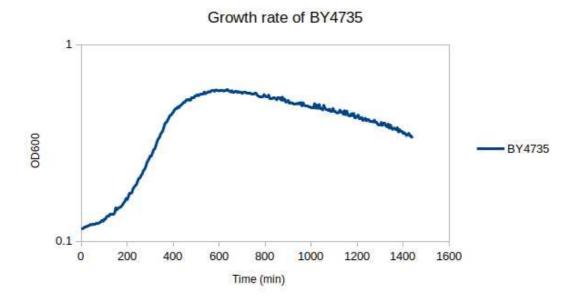
The most common way to ask "how happy is this yeast?" is to ask "how fast is it growing?" And so your experiment may generate growth-curve data such as the plot below (for the wild-type strain, growing in YPD media):





We can even answer the question "how fast is this yeast strain growing?" with a single number: the doubling-time when the yeast is growing exponentially. To determine the doubling time, there are four steps.

1. Plot the growth curve with the Y axis scale set to "logarithmic"

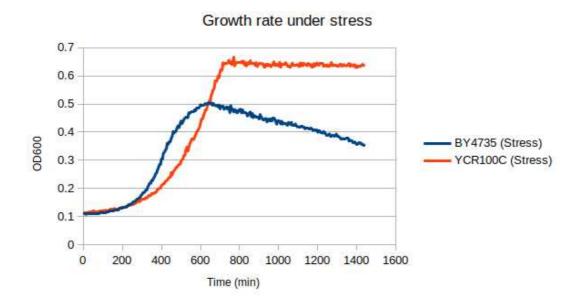


- 1. Find the part of the curve that is linear. In this case, I might choose from 250 minutes to 350 minutes.
- 2. Find the OD600 reading at these two times. (Remember, OD600 is the "optical density" of the yeast how much yeast is in the well.) At 250 minutes, my OD600 is 0.2086; at 350 minutes, my OD600 is 0.3583.
- 3. Compute the doubling time using a bit of math:

$$T_d=(t_2-t_1)\cdotrac{\ln(2)}{\ln(rac{q_2}{q_1})}.$$

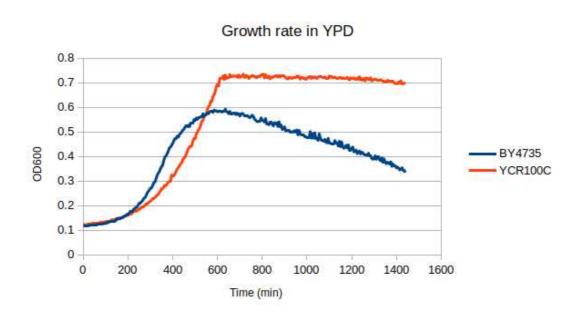
Where t_1 and t_2 are the start and end times you found in step 2, and q_1 and q_2 are the OD600 at those times. Thus, for my BY4735 strain growing in YPD, the doubling time is 128 minutes.

So let's compare a knock-out strain YCR100 to the wild-type BY4735 strain under some stress condition. Here's the pair of growth curves:



The BY4735's doubling time under these conditions is 129 minutes (not much different than in YPD), while the YCR100C's doubling time is 183 minutes. So there's clearly a difference. However, while it's possible that the stress is slowing down the growth of the YCR100C knock-out strain, there's another possible explanation: *it could just be that the YCR100C strain grows more slowly under all conditions*.

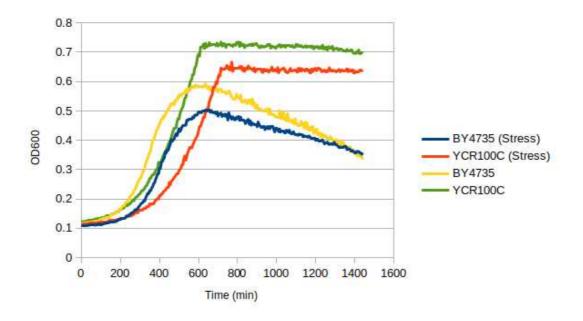
To see whether that's the case, we need to compare BY4735 and the YCR100C growing in YPD media without any stress. Here's that data:





Well, that plot looks very similar! The doubling time of BY4725 (without stress) is 128 minutes, while the doubling time of YCR100C (without stress) is 185 minutes. Thus, the YCR100C knock-out strain doesn't seem to be any more or less sensitive to this stress -- it just grows more slowly in general. (Which is a phenotype in and of itself.)

Here's all four curves plotted together:



One more thing. The doubling time doesn't have to be the only conclusion you draw from these plots – feel free to note other changes. For example, I might say that the YCR100C knock-out strain grows more slowly, but reaches a higher OD600 and doesn't decline nearly as much after it saturates. Those dynamics don't appear to change when I stress them using this particular stress.