addgene

A better way to share plasmids

CRISPR 101

A Desktop Resource

Created and Compiled by Addgene
January 2021 (3rd Edition)

www.addgene.org

Introduction to Addgene's CRISPR eBook

If you work in molecular biology chances are you've heard of CRISPR. This amazing technology, best known for its use in fast and easy genome editing, has accelerated research in many basic and applied disciplines.

As the tools developed around CRISPR have grown, so too have the resources available at Addgene thanks to the generous contributions of depositors working in the field. To better help you understand how to use these resources, we've been steadfast in our efforts to keep the CRISPR educational content on both our website and our blog up to date.

Now, to help you best utilize CRISPR and the plasmids kindly deposited by your colleagues we've compiled this comprehensive eBook. Whether you're looking to use CRISPR for the first time, need some advice on a particular CRISPR technique, or would simply like to learn more about how CRISPR came about, we hope that you'll find something interesting.

If you have any questions about the eBook, have suggestions for new content, or would just like to send us a note, please contact us at blog@addgene.org. Happy reading and good luck with your experiments!

- The Addgene Team

Join our community!

@addgene

blog.addgene.org

www.addgene.org

addgene
The nonprofit plasmid repository

Table Of Contents

addgene
The nonprofit plasmid repository

Table Of Contents

Table Of Contents

Chapter 1

Introduction to CRISPR

Components of CRISPR/Cas9

By Joel McDade | Feb 2, 2016, Updated Mar 26, 2020 by Jennifer Tsang

At their most basic level, CRISPR/Cas9 genome editing systems use a non-specific endonuclease (Cas9 or closely related Cpf1) to cut the genome and a small RNA (gRNA) to guide this nuclease to a user-defined cut site. After reading this article, we hope you will be caught up on much of the major CRISPR lingo and will be able to describe the functions of the various CRISPR/Cas9 components. Please note that while this article is intended to provide a general overview of CRISPR components, new Cas9 variants are being discovered all the time and the requirements of these different systems can vary (for example, xCas9 is a variant with increased PAM flexibiliy and eSpCas9/SpCas9-HF1 have increased targeting specificity).

Cas9 endonuclease functions

While native CRISPR/Cas systems have a variety of enzymes responsible for processing foreign DNA as well as the RNA guides required for endonuclease function, when used for genome engineering, the only CRISPR protein required is the Cas9 endonuclease or a variant thereof. This individual protein has all the components necessary to:

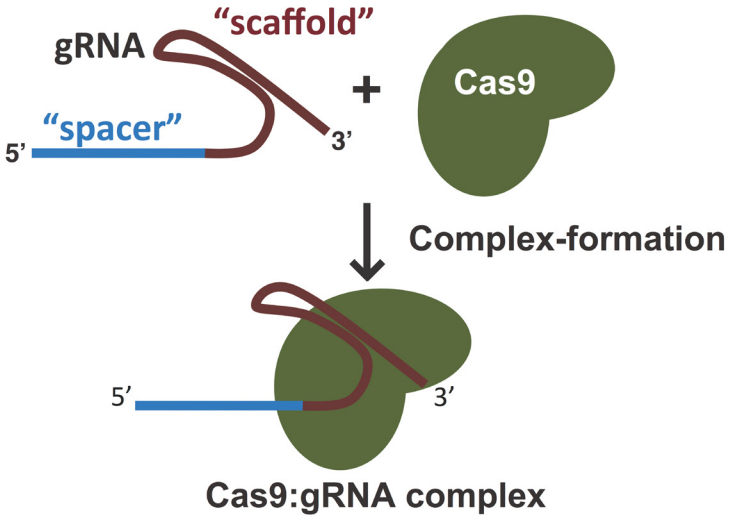


Figure 1: The gRNA, which contains a spacer and scaffold sequence binds to Cas9 to form the Cas9:gRNA complex.

1. Bind to a guide RNA
The guide RNA, which will be described in more detail below, enables Cas9 to cut a specific genomic locus of many possible loci. Without binding to the guide RNA, Cas9 cannot cut.

2. Bind to target DNA in the presence of a guide RNA provided that target is upstream (5') of a protospacer adjacent motif (PAM)
The gRNA and the Cas9 endonuclease come together to form the Cas9:gRNA complex. Cas9 endonuclease binding to the target genomic locus is mediated both by the target sequence contained within the guide RNA and a 3-base pair sequence known as the Protospacer Adjacent Motif or PAM. In order for dsDNA to be cut by Cas9, it must contain a PAM sequence immediately downstream (3') of the site targeted by the guide RNA. In the absence of either the guide RNA or a PAM sequence, Cas9 will neither bind nor cut the target. Cas9 homologs from different organisms or Cas9 mutants developed in a variety of labs (see table below) have different PAM requirements. These different PAM requirements allow researchers to target many different genomic loci.

Components of CRISPR/Cas9 (CONT'D)

Cas9 and its variants have two endonuclease domains: the n-terminal RuvC-like nuclease domain and the HNH-like nuclease domain near the center of the protein. Upon target binding, Cas9 undergoes a conformational change that positions the nuclease domains to cleave opposite strands of the target DNA. Thus, the end result of Cas9-mediated DNA damage is a DSB within the target DNA ~3-4 nucleotides upstream of the PAM sequence.

Table 1: Cas9 species/variants and PAM sequences

Species/Variant of Cas9	PAM Sequence
Streptococcus pyogenes (SP); Sp-Cas9	3' NGG
SpCas9 D1135E variant	3' NGG (reduced NAG binding)
SpCas9 VRER variant	3' NGCG
SpCas9 EQR variant	3' NGAG
SpCas9 VQR variant	3' NGAN or NGNG
xCas9	3' NG, GAA, or GAT
SpCas9-NG	3' NG
Staphylococcus aureus (SA); Sa-Cas9	3' NNGRRT or NNGRR(N)
Acidaminococcus sp. (AsCpf1) and *Lachnospiraceae bacterium* (Lb-Cpf1)	5' TTTV
AsCpf1 RR variant	5' TYCV
LbCpf1 RR variant	5' TYCV
AsCpf1 RVR variant	5' TATV
Campylobacter jejuni (CJ)	3' NNNNRYAC
Neisseria meningitidis (NM)	3' NNNNGATT
Streptococcus thermophilus (ST)	3' NNAGAAW
Treponema denticola (TD)	3' NAAAAC

Components of CRISPR/Cas9 (CONT'D)

The synthetic guide RNA or gRNA (sometimes sgRNA)

In the native Type II CRISPR/Cas system, Cas9 is guided to its target sites with the aid of two RNAs: the crRNA which defines the genomic target for Cas9, and the tracrRNA which acts as a scaffold linking the crRNA to Cas9 and facilitates processing of mature crRNAs from pre-crRNAs derived from CRISPR arrays. In most systems used for CRISPR-mediated genome editing, these two small RNAs have been condensed into one RNA sequence known as the guide RNA (gRNA) or single guide RNA (sgRNA). We'll refer to this RNA complex as the "gRNA". The gRNA contains both the 20 nucleotide target sequence to direct Cas9 to a specific genomic locus and the scaffolding sequence necessary for Cas9 binding. When using CRISPR/Cas9 for genome editing, researchers simply need to express a gRNA designed to direct Cas9 to their target sequence of choice (see tips for designing a gRNA) and their prefered Cas9 variant (with the appropriate PAM sequence) to modify the desired genomic locus.

Anti-CRISPR Proteins Switch Off CRISPR-Cas Systems

By Beth Kenkel | May 23, 2017, Updated Jul 23, 2020 by Jennifer Tsang

CRISPR-Cas technology is constantly evolving. Variants of Cas proteins can be used for genome editing, activating gene expression, repressing gene expression, and much more. But there's one thing that was missing: a way to shut off Cas's activity. The concern is that the longer Cas remains active in a cell, the greater chances there are for off-target edits to occur. Although methods to switch on Cas activity using light or drugs have been developed, the field lacked an "off-switch" for Cas proteins.

Discovery of anti-CRISPR proteins

That is, until the discovery of anti-CRISPR proteins in 2012 (Bondy-Denomy et al., 2012). Anti-CRISPR (Acr) proteins are phage-derived small protein inhibitors of CRISPR-Cas systems that help phages evade the CRISPR-Cas immune system of bacteria. These Acr proteins were originally found in type I CRISPR-Cas systems.

Four years later, the Sontheimer and Davidson labs discovered three anti-CRISPRs for a type II system in *Neisseria meningitidis* (Nme) (Pawluk et al., 2016). While *Streptococcus pyogenes* Cas9 (SpyCas9) is the most used and well studied CRISPR system, *N. meningitidis* can also be used for genome editing. In this study, the labs found that anti-CRISPR prevents Cas9 activity in HEK293 by reducing editing from ~30% in control samples to 0-10% in Acr samples. Acr proteins can also be used to prevent dCas9 tethering to DNA.

Since these early anti-CRISPR studies, scientists have discovered over 50 anti-CRISPR proteins that interact with CRISPR-Cas systems such as Cascade-Cas3, Cas9, Cas12, and Cas13 (Marino et al., 2020). Some Acr proteins are specific to a particular Cas protein while others can inhibit CRISPR enzymes from multiple species of bacteria.

A note on anti-CRISPR nomenclature

Acr family proteins are named for the type of CRISPR-Cas system they block and are numbered in order of discovery (Bondy-Denomy et al., 2018). For example, AcrIIA4 mentioned above indicates that it's the fourth Acr protein discovered that blocks type II-A CRISPR-Cas systems. A full list of Acr proteins organized by the type of CRISPR-Cas system they block can be found here.

How do anti-CRISPR proteins work?

Anti-CRISPRs proteins are highly diverse but most block CRISPR activity one of three ways:

Inhibit DNA binding: AcrIIA4 blocks Cas9's interaction with the PAM site (Dong et al., 2017), and AcrIIC3 causes Cas proteins to dimerize which blocks the PAM recognition site (Harrington et al., 2017).

Anti-CRISPR Proteins Switch Off CRISPR-Cas Systems (CONT'D)

Inhibit crRNA loading: AcrIIC2 interactions with Cas9 interferes and prevents the correct assembly of the crRNA-Cas complex (Zhu et al., 2019).

Inhibiting DNA cleavage: AcrIIC1 binds to the HNH endonuclease domain of Cas9 and prevents target DNA cleavage (Harrington et al., 2017).

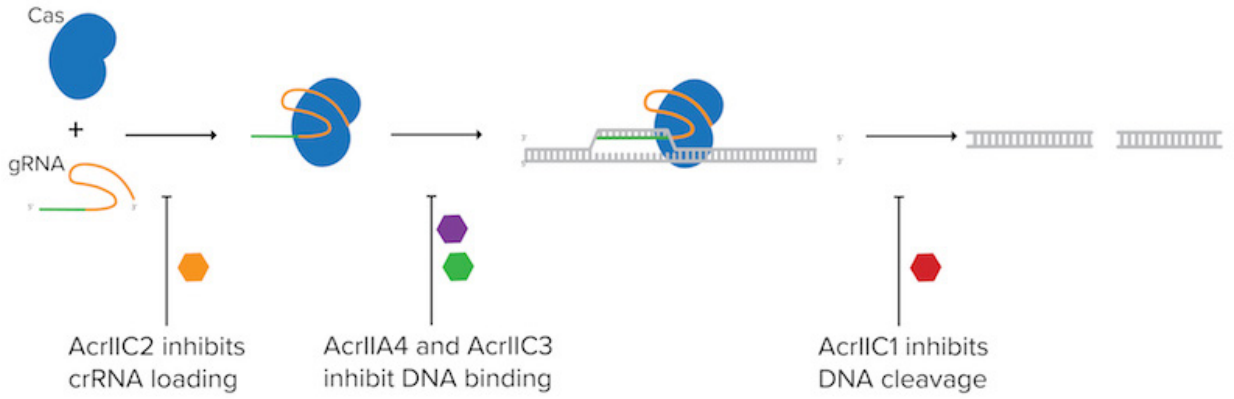


Figure 1: Anti-CRISPR proteins (Acr's) can block CRISPR activity in many ways such as inhibiting crRNA loading, inhibiting DNA binding, or inhibiting DNA cleavage.

Using anti-CRISPR proteins in your experiments

Reducing off-target effects

Prolonged Cas activity can increase the chance of off-target editing particularly after the majority of on-target editing has occurred. Anti-CRISPR proteins can be used to limit this off-target editing, but when is the best time to shut off Cas activity? Using inhibitor timing experiments, the Doudna lab found that at least 50% of on-target Cas9 edits happens within the first six hours. When added to human cells six hours after the introduction of Cas9 RNPs, AcrIIA4 reduced off-target editing while still allowing on-target editing (Shin et al., 2017).

Acrs could also be used to reduce off-target edits of base editors, which convert a nucleotide base pair to another base pair at a specific site.

Temporal, spatial, or conditional control of CRISPR activity

Regulating Acr proteins by inducible promoters, light, or small molecules could achieve rapid, dynamic, and spatial control of CRISPR-Cas activity. For example, fusing AcrA4II to the photo-sensitive LOV2

Anti-CRISPR Proteins Switch Off CRISPR-Cas Systems (CONT'D)

domain allows anti-CRISPR activity to be controlled by blue light (Bubeck et al., 2018).

Reduce cytotoxicity of CRISPR-Cas editing in tissues

The small size of Acr proteins (~50-200 amino acids) means they can be delivered *in vivo* using AAV vectors. Having a way to shut off CRISPR-Cas systems prevents the off-target effects and cytotoxicity that result from excessive nuclease activity in tissues (Li et al., 2018).

Selection marker while engineering viruses

Acr genes can be used as positive selectable markers in engineering viruses. For example, an Acr gene was used to replace genes in the difficult-to-engineer *Sulfolobus islandicus* rod-shaped virus 2 (Mayo-Muñoz et al., 2018). Only viral particles that have the gene deletion could replicate when challenged with the *S. islandicus* CRISPR-Cas system.

Phage therapy

Phage therapy uses phage to treat bacterial infections as an alternative to antibiotics. But because many bacterial pathogens possess CRISPR-Cas systems, phage therapy might not be as effective. This is where Acr proteins come in. Acr proteins are small enough that they could be engineered to "armor" therapeutic phages to protect them from pathogenic bacteria's CRISPR-Cas system.

Conclusion

CRISPR is a great way to make those genome edits. But, just as important is a way to flip the switch and turn off CRISPR activity. The ability to turn off CRISPR activity is an important step towards CRISPR-based therapies and for more precise editing in the lab.

References

Dong D, Guo M, Wang S, Zhu Y, Wang S, Xiong Z, Yang J, Xu Z, Huang Z (2017) Structural basis of CRISPR–SpyCas9 inhibition by an anti-CRISPR protein. Nature 546:436–439 . https://doi.org/10.1038/nature22377

Harrington LB, Doxzen KW, Ma E, Liu J-J, Knott GJ, Edraki A, Garcia B, Amrani N, Chen JS, Cofsky JC, Kranzusch PJ, Sontheimer EJ, Davidson AR, Maxwell KL, Doudna JA (2017) A Broad-Spectrum Inhibitor of CRISPR-Cas9. Cell 170:1224-1233.e15 . https://doi.org/10.1016/j.cell.2017.07.037

Anti-CRISPR Proteins Switch Off CRISPR-Cas Systems (CONT'D)

Li C, Psatha N, Gil S, Wang H, Papayannopoulou T, Lieber A (2018) HDAd5/35++ Adenovirus Vector Expressing Anti-CRISPR Peptides Decreases CRISPR/Cas9 Toxicity in Human Hematopoietic Stem Cells. Molecular Therapy - Methods & Clinical Development 9:390–401 . https://doi.org/10.1016/j.omtm.2018.04.008

Marino ND, Pinilla-Redondo R, Csörgő B, Bondy-Denomy J (2020) Anti-CRISPR protein applications: natural brakes for CRISPR-Cas technologies. Nat Methods 17:471–479 . https://doi.org/10.1038/s41592-020-0771-6

Mayo-Muñoz D, He F, Jørgensen J, Madsen P, Bhoobalan-Chitty Y, Peng X (2018) Anti-CRISPR-Based and CRISPR-Based Genome Editing of *Sulfolobus islandicus* Rod-Shaped Virus 2. Viruses 10:695 . https://doi.org/10.3390/v10120695

Pawluk A, Amrani N, Zhang Y, Garcia B, Hidalgo-Reyes Y, Lee J, Edraki A, Shah M, Sontheimer EJ, Maxwell KL, Davidson AR (2016) Naturally Occurring Off-Switches for CRISPR-Cas9. Cell 167:1829-1838.e9 . https://doi.org/10.1016/j.cell.2016.11.017

Shin J, Jiang F, Liu J-J, Bray NL, Rauch BJ, Baik SH, Nogales E, Bondy-Denomy J, Corn JE, Doudna JA (2017) Disabling Cas9 by an anti-CRISPR DNA mimic. Sci Adv 3:e1701620 . https://doi.org/10.1126/sciadv.1701620

CRISPR Glossary

Base editor
Fusion of a Cas protein to a deaminase that enables direct base conversion in RNA or DNA without a DNA double-strand break

Cas
CRISPR Associated Protein, includes nucleases like Cas9 and Cas12a (also known as Cpf1)

CRISPR
Clustered Regularly Interspaced Short Palindromic Repeat, a bacterial genomic region used in pathogen defense

CRISPRa
CRISPR Activation; using a dCas9 or dCas9-activator with a gRNA to increase transcription of a target gene

CRISPRi
CRISPR Interference; using a dCas9 or dCas9-repressor with a gRNA to repress/decrease transcription of a target gene

Cut
A DNA double-strand break; the wild type function of most Cas nucleases

dCas9
Nuclease dead Cas9, an enzymatically inactive form of Cas9; can bind, but cannot cleave DNA

DSB
Double-Strand Break, a break in both strands of DNA

Dual nickase/Double nick
A method to decrease off-target effects by using a single Cas9 nickase and 2 different gRNAs, which bind in close proximity on opposite strands of the DNA, to create a DSB

enChIP
Engineered DNA-binding molecule-mediated ChIP, using a tagged-dCas9+gRNA to purify specific genomic regions to identify molecules associated with the genomic regions

CRISPR Glossary (CONT'D)

Genetic modification or manipulation
Any genetic perturbation, including genetic knockout, gene activation, or gene repression

gRNA
Guide RNA, a synthetic fusion of the endogenous bacterial crRNA and tracrRNA that provides both targeting specificity and scaffolding/binding ability for Cas9 nuclease. This synthetic fusion does not exist in nature and is also commonly referred to as an sgRNA.

gRNA scaffold sequence
The sequence within the gRNA that is responsible for Cas9 binding, it does not include the 20 bp spacer/targeting sequence that is used to guide Cas9 to target DNA

gRNA targeting sequence
The 20 nucleotides that precede the PAM sequence in the genomic DNA. This sequence is cloned into a gRNA expression plasmid but does NOT include the PAM sequence or the gRNA scaffold sequence

HDR
Homology Directed Repair, a DNA repair mechanism that uses a template to repair DNA nicks or DSBs

Indel
Insertion/deletion, a type of mutation that can result in the disruption of a gene by shifting the ORF and/or creating premature stop codons

NHEJ
Non-Homologous End Joining; a DNA repair mechanism that often introduces indels

Nick
A break in only one strand of dsDNA

Nickase
Cas9 with one of the two nuclease domains inactivated. This enzyme is capable of cleaving only one strand of target dsDNA.

CRISPR Glossary (CONT'D)

Off-target effects or off-target activity
Cas9 cleavage at undesired locations due to gRNA targeting sequence with sufficient homology to recruit Cas9 to unintended genomic locations

On-target activity
Cas9 cleavage at a desired location specified by a gRNA target sequence

ORF
Open Reading Frame; the translated codons that make up a gene

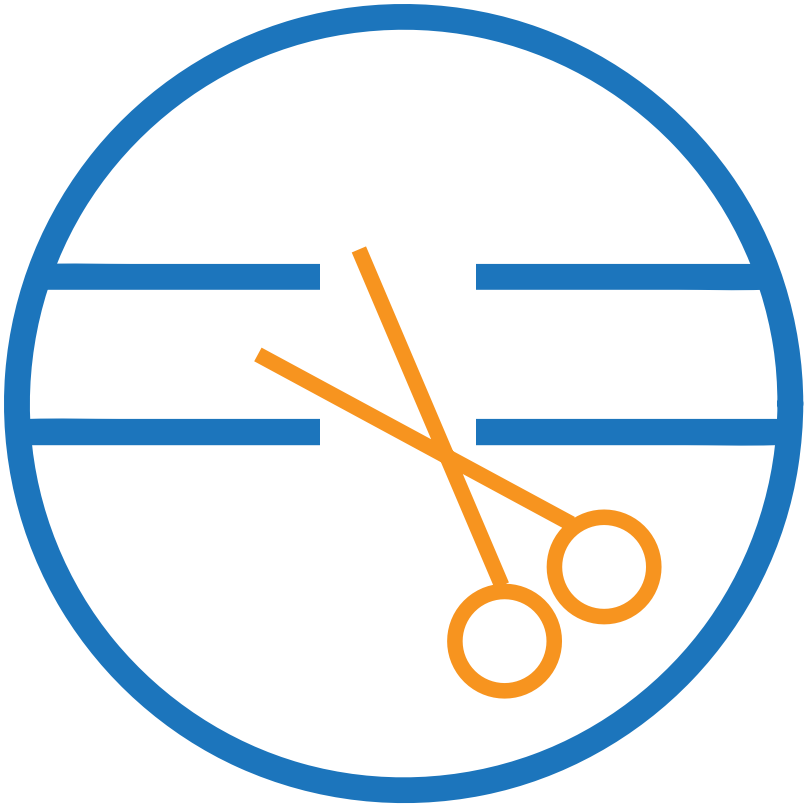
PAM
Protospacer Adjacent Motif; sequence adjacent to the target sequence that is necessary for Cas enzymes to bind target DNA

PCR
Polymerase Chain Reaction; used to amplify a specific sequence of DNA

Target locus
Genomic target of the gRNA. The sequence includes the unique ~20 bp target specified by the gRNA plus the genomic PAM sequence

Chapter 2

Cas Functions: Generating Knockout



Homology Directed Repair

By Chari Cortez | Mar 12, 2015, Updated Jul 27, 2020 by Christina Mork

DNA lesions are sites of structural or base-pairing damage of DNA. Perhaps the most harmful type of lesion results from breakage of both DNA strands – a double-strand break (DSB) – as repair of DSBs is paramount for genome stability. DSBs can be caused by intracellular factors such as nucleases and reactive oxygen species, or external forces such as ionizing radiation and ultraviolet light; however, these types of breaks occur randomly and unpredictably. To provide some control over the location of the DNA break, scientists have engineered plasmid-based systems that can target and cut DNA at specified sites. Regardless of what causes the DSB, the repair mechanisms function in the same way.

In this article, we will describe the general mechanism of homology directed repair with a focus on repairing breaks engineered in the lab for genome modification purposes.

How does homology directed repair fix double-strand breaks?

Genome stability requires the correct and efficient repair of DSBs. In eukaryotic cells, mechanistic repair of DSBs occurs primarily by two pathways: Non-Homologous End-Joining (NHEJ) and Homology Directed Repair (HDR). NHEJ is the canonical homology-independent pathway as it involves the alignment of only one to a few complementary bases for the re-ligation of two ends, whereas HDR uses longer stretches of sequence homology to repair DNA lesions.

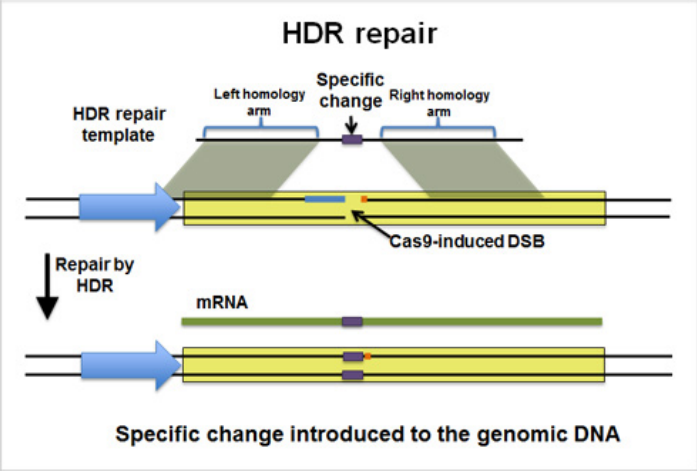


Figure 1: The HDR repair template contains a specific change in the middle of sequences homologous to the genomic DNA. After a Cas9-induced DSB, the break is repaired by HDR which introduces the specific change into the genomic DNA.

This article focuses on HDR, which is the more accurate mechanism for DSB repair due to the requirement of higher sequence homology between the damaged and intact donor strands of DNA. The process is error-free if the DNA template used for repair is identical to the original DNA sequence at the DSB, or it can introduce specific mutations into the damaged DNA.

There are four different HDR pathways to repair DSBs. Here are three central steps of the HDR pathways:

- The 5'-ended DNA strand is resected at the break to create a 3' overhang. This will serve as both a substrate for proteins required for strand invasion and a primer for DNA repair synthesis.
- The invasive strand can then displace one strand of the homologous DNA duplex and pair with the

Homology Directed Repair (CONT'D)

other; this results in the formation of the hybrid DNA referred to as the displacement loop (D loop). This is the defining point of HDR.

- The recombination intermediates can then be resolved to complete the DNA repair process.

Homology directed repair pathways

HDR can occur either non-conservatively or conservatively. The non-conservative method is composed of the single-strand annealing (SSA) pathway and is more error prone. The conservative methods, characterized by the accurate repair of the DSB by means of a homologous donor (e.g., sister chromatid, plasmid, etc), are composed of three pathways: the classical double-strand break repair (DSBR), synthesis-dependent strand-annealing (SDSA), and break-induced repair (BIR).

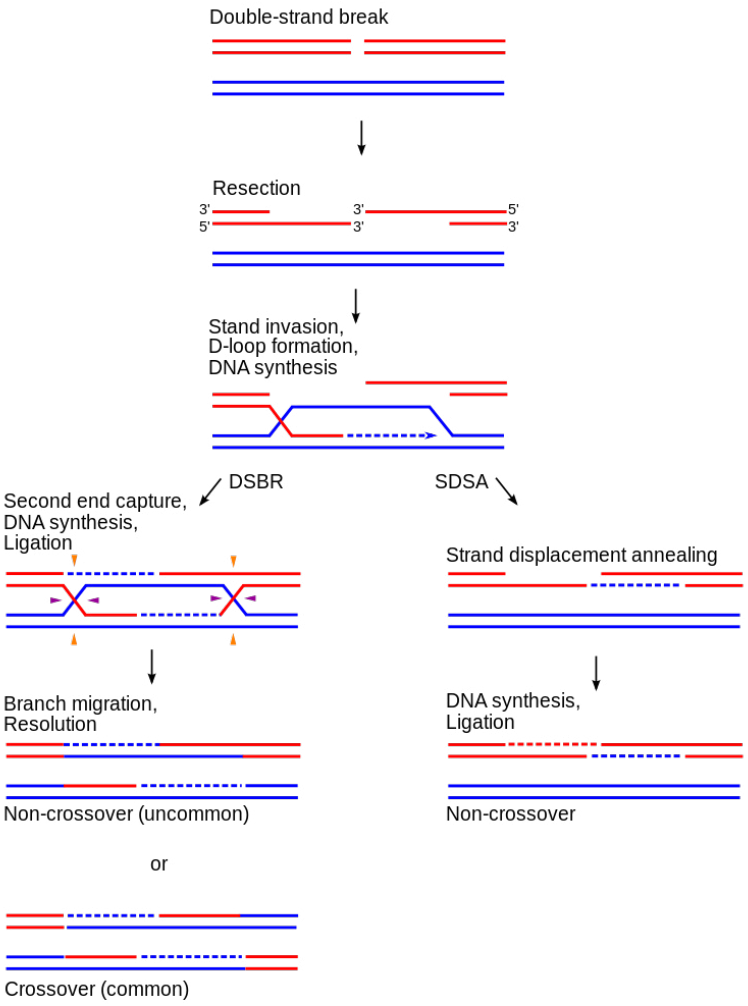


Figure 2: Illustration of double strand break repair by DSBR and SSDA. Image from Emw2012.

Classical double-strand break repair (DSBR)

In the classical DSBR pathway, the 3' ends invade an intact homologous template to serve as a primer for DNA repair synthesis, ultimately leading to the formation of double Holliday junctions (dHJs). dHJs are four-stranded branched structures that form when elongation of the invasive strand "captures" and synthesizes DNA from the second DSB end. The individual HJs are resolved via cleavage in one of two ways. Looking at the left branch of the figure, each junction resolution could happen on the crossing strand (horizontally at the purple arrows) or on the non-crossing strand (vertically at the orange arrows). If resolved dissimilarly (e.g. one junction is resolved on the crossing strand and the other on the non-crossing strand), a crossover event will occur; however, if both HJs are resolved in the same manner, this results in a non-crossover event. DSBR is semi-conservative, as crossover events are most common. This animation nicely illustrates the DSBR pathway.

Homology Directed Repair (CONT'D)

Synthesis-dependent strand-annealing (SDSA) pathway

As illustrated on the right branch of the figure above, SDSA is conservative, and results exclusively in noncrossover events. This means all newly synthesized sequences are present on the same molecule. Unlike DSBR, following strand invasion and D loop formation in SDSA, the newly synthesized portion of the invasive strand is displaced from the template and returned to the processed end of the non-invading strand at the other DSB end. The 3' end of the non-invasive strand is elongated and ligated to fill the gap, thus completing SDSA.

Break-induced repair (BIR) pathway

BIR is not as well characterized as either DSBR or SDSA, but one central feature of this pathway is the presence of only one invasive end at a DSB that can be used for repair. This single invasive strand invades a homologous sequence and initiates both leading and lagging strand synthesis, which results in the formation of one HJ. This HJ is resolved by cleavage of the crossed strand. While this pathway may not be immediately applicable in DSB-induced gene targeting or relevant to plasmid-based genome engineering, it may have biological importance for the repair of chromosome ends that have no second end that would enable DSBR or SDSA.

Homology directed repair and genome engineering

Plasmid-based methods that induce DSBs have employed homologous recombination for genome engineering. Zinc finger nucleases (ZFNs), TAL effector nucleases (TALENs), and CRISPR can all direct a nuclease to cause a specific DSB. The ease of creating guides, the speed of the system, and its targeting versatility have not just reinvigorated genome engineering, but have really revolutionized the field.

General considerations for designing a repair template to create mutations

HDR templates used to create specific mutations or insert new elements into a gene require a certain amount of homology surrounding the target sequence that will be modified. Scientists have been most successful using homology arms that start at the CRISPR-induced DSB. In general, the insertion sites of the modification should be very close to the DSB, ideally less than 10 bp away if possible. Check out our article on Improving HDR efficiency for more details on how "Cut-to-Mutation distance" affects editing efficiency.

One important point to note is that the CRISPR enzymes may continue to cleave DNA once a DSB is introduced and repaired. As long as the gRNA target site/PAM site remain intact, the Cas9

Homology Directed Repair (CONT'D)

endonuclease will keep cutting and repairing the DNA. This repeated editing is problematic if you are trying to introduce a very specific mutation or sequence. To get around this, you should design the HDR template in such a way that will ultimately block further Cas9 targeting after the initial DSB is repaired. Two ways to block further editing are mutating the PAM or the gRNA seed sequence.

What makes the best repair template: plasmid DNA or single-stranded donor oligonucleotide (ssODN)?

When designing your repair template, the size of your intended edit is a big factor. ssDNA templates (referred to as ssODNs) are commonly used for smaller modifications. Small insertions/edits may require as little as 30-50 bases for each homology arm, but keep in mind these numbers may vary based on your locus of interest and experimental system. 50-80 base homology arms are commonly used. Asymmetric homology arms (36 bases distal to the PAM and 91 bases proximal to the PAM) supported HDR efficiencies up to 60% (Corn et al., 2016).

Due to the difficulties associated with creating ssODNs longer than 200 bases, researchers have used dsDNA plasmid templates for larger insertions such as fluorescent proteins or selection cassettes. These templates should have homology arms of at least 800 bp. HDR efficiency with plasmid templates is generally low; to increase the frequency of edits, researchers have designed self-cleaving plasmids that contain gRNA target sites flanking the template. When the Cas enzyme and the appropriate gRNA(s) are present, the template is liberated from the vector. To avoid plasmid cloning, scientists also use PCR-generated long dsDNA templates, but these templates may be toxic to cells, thus lowering editing efficiency. Another disadvantage of dsDNA templates is their ability to bluntly integrate into the genome, duplicating the homology arm sequences.

Easi-CRISPR is a technique to allow researchers making large mutations to take advantage of the benefits of ssODNs (Quadros et al., 2017). To create ssODNs longer than 200 bases, you *in vitro* transcribe RNA encoding your repair template, then use reverse transcriptase to create the complementary ssDNA. Easi-CRISPR works well in mouse knock-in models, increasing editing efficiency from 1-10% with dsDNA to 25-50% with ssODNs. Although HDR efficiency varies across loci and experimental systems, ssODN templates generally provide the highest frequency of HDR edits.

Promoting HDR over NHEJ

To promote HDR over the more common NHEJ repair mechanism, strategies such as inhibiting NHEJ and activating HDR have been used. Rees et al. and Nambiar et al. have used HDR factors such as Rad variants like Rad18 (Nambiar et al., 2019) and hRad51 (Rees et al., 2019) to activate HDR. Inhibitors such as CYREN can be used to inhibit NHEJ (Arnoult et al., 2017), while a Cas9 fusion to dominant negative 53BP1 both inhibits NHEJ and activates HDR (Jayavaradhan et al., 2019).

3 Tips to Improve HDR Efficiency for CRISPR Editing in Human Cells

By Dominik Paquet and Dylan Kwart | Sept 5, 2017

The CRISPR/Cas9 system is a versatile tool for precise gene editing in many organisms and model systems. We have used CRISPR/Cas9 extensively to make sequence-specific changes in human induced pluripotent stem cells (iPSCs). The CRISPR/Cas9 complex is very efficient at introducing double stranded breaks (DSBs) into genomic DNA in many cell types and often results in biallelic modifications. Most commonly, DSBs are repaired by the nonhomologous end-joining (NHEJ) pathway, leading to non-specific nucleotide insertions, deletions or other mutations, referred to as 'indels'. While this is convenient for generating gene knockouts, NHEJ repair does not allow introduction of specific sequence changes.

To generate a specific sequence change the cell must undergo homology-directed repair (HDR), a distinct cellular DNA repair pathway. Accomplishing this usually involves simultaneously introducing a homologous DNA repair template, such as a single-stranded oligodeoxynucleotide (ssODN), which contains the intended sequence changes to be incorporated into the edited genome. In mammalian cells such as stem cells, HDR is relatively rare and DSBs are predominantly repaired by NHEJ. While other groups have focused on strategies to improve overall HDR rates, we have shown recently that desired genome-editing events by HDR can be generated more efficiently by (1) improving editing accuracy by preventing re-editing and (2) optimizing "cut-to-mutation distance". Furthermore (3), we have developed a framework that, by combining these two strategies allows you to specifically incorporate homo- or heterozygous mutations.

1. Increasing HDR accuracy with CRISPR/Cas-blocking mutations

After genome editing with CRISPR/Cas9 in human cells was described in 2013 (Cong et al., 2013, Mai et al., 2013) we were very excited to set up a CRISPR-based editing platform in our lab. Our goal was to use the system to introduce disease-causing mutations in iPSCs (or correct them back to wild type), differentiate them into somatic cells affected by the disease, and study what's wrong with the cells. However, our excitement turned into frustration when we sequenced our cells and found how inefficiently our edits were being introduced. Although we were able to introduce the desired mutations in genes associated with dementia, such as APP and PSEN1, in most cases the intended HDR events were corrupted by unwanted indels on the same allele, probably due to concomitant NHEJ repair. At the PSEN1 locus, over 90% of the HDR edited alleles had unwanted insertions/deletions, therefore making those alleles useless for our studies. If one now also takes into account that HDR is rare, usually 2-5% of the edited cells in our hands, these combined pitfalls made it virtually impossible to find the 0.2-0.5% correctly edited alleles, when screening reasonable amounts of single cell clones. Furthermore, CRISPR/Cas9 editing is often bi-allelic, meaning that even if one allele is accurately edited, the other likely contains unwanted indels.

Following this observation, we tested a strategy that was discussed in the field at the time but had never been systematically analyzed: introducing silent CRISPR/Cas9 blocking mutations in either the

3 Tips to Improve HDR Efficiency for CRISPR Editing in Human Cells (CONT'D)

PAM (changing one of the G's in SpCas9's PAM to an A, T or C – but avoiding NAG) or guide RNA target ("seed") sequence. The idea was that these "blocking mutations" mutations prevent CRISPR from re-cutting the target sequence once the desired edit has been introduced.

We found that blocking mutations increased the accuracy of editing by up to 10-fold per allele. This is a 100-fold increase in editing accuracy if both alleles in a single cell are edited (Paquet, Kwart, et al., 2016). For those of you involved in the tedious business of iPSC clone picking, a 100-fold decrease is the difference between picking 100 versus 10,000 clones! We would therefore strongly encourage any researcher to incorporate CRISPR/Cas-blocking mutations at both targeted alleles when gene editing.

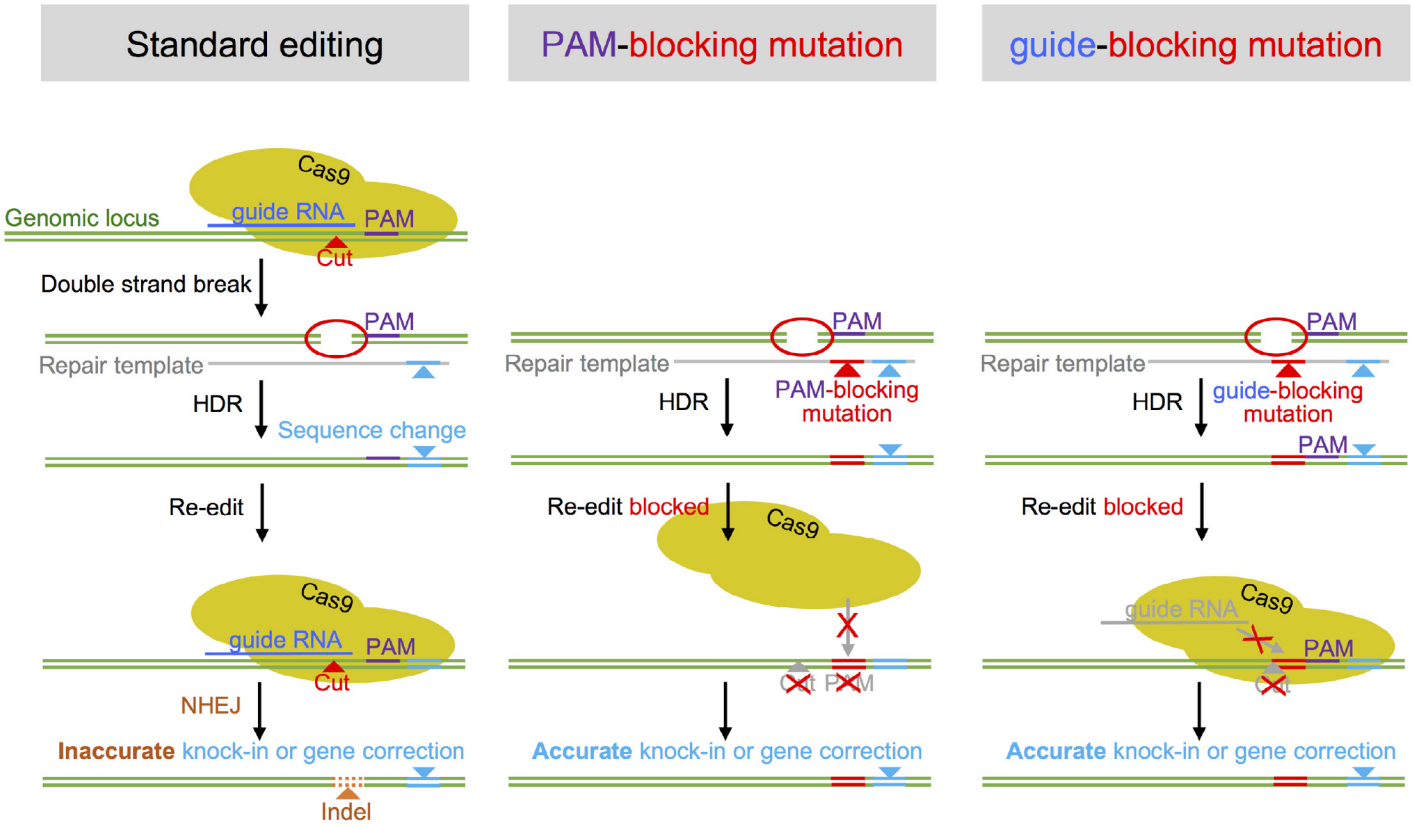


Figure 1: Techniques to improve editing efficiency.

But what's better: PAM-blocking or guide RNA blocking mutations? This depends entirely on the locus you are editing, and if the blocking mutation needs to be silent. In our view, it's always best to change the PAM site, however we have also demonstrated the efficiency of guide RNA blocking mutations. So if the locus does not allow blocking the PAM sequence (e.g. if it would lead to a missense or nonsense mutation), blocking the guide RNA binding site works well. However, since some guide RNAs may

3 Tips to Improve HDR Efficiency for CRISPR Editing in Human Cells (CONT'D)

tolerate a mismatch in the binding sequence, we recommend changing several bases as close to the PAM as possible (see below to find out why).

What if your locus does not allow you to use permanent blocking mutations at all, such as when you are editing a non-coding DNA region? To this end we developed a technology called CORRECT, which, by undergoing two rounds of gene-editing enables scarless introduction of an intended mutation while taking advantage of the improved efficiency provided by blocking mutations. Essentially, you:

Step 1 - Introduce the blocking mutation together with the intended sequence change and screen a few hundred clones to find the right one.

Step 2 – Edit the cells a second time using a repair template that corrects the blocking mutation and again identify the right clone from a few hundred.

From these two steps, you'll be screening a total of a few hundred clones - Still better than 10,000, right? (Kwart, Paquet, et al 2017).

2. Optimizing "cut-to-mutation distance"

Even though we had controlled over-active Cas9 and repeated our experiments with blocking mutations, we still weren't entirely happy with our editing outcomes and in many cases still didn't get what we wanted. The indels were gone from most HDR reads that we studied by deep sequencing, but interestingly (and to our frustration) in many cases we only saw incorporation of the blocking mutation. Despite the fact that both the blocking mutation and our intended mutation were on the same oligo, the intended mutation was rarely incorporated. It seemed that the cell was only using a part of the repair template and mutations further away from the cut site were not efficiently incorporated during HDR.

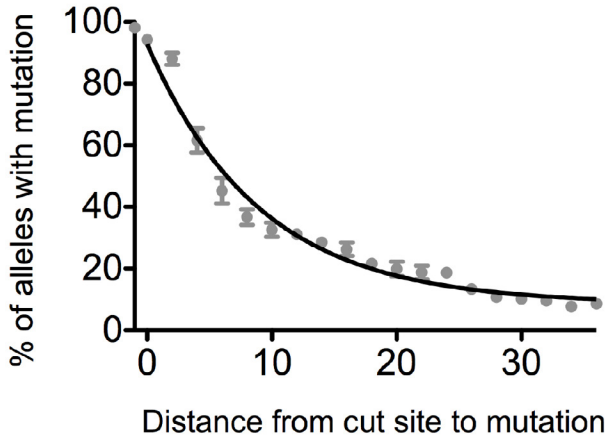


Figure 2: Relationship between cut site to mutation distance and editing efficiency.

We decided to systematically characterize this phenomenon and found a general relationship between the distance of a mutation from the CRISPR/Cas9 cut site and how frequently it is incorporated. We were surprised to see how rapidly the efficiency of mutation incorporation falls with increasing distance from the cut site. If the distance is only 10 bp, the efficiency already drops by half and after only about 30 bp, it is no longer feasible to incorporate mutations without screening thousands of clones to find

3 Tips to Improve HDR Efficiency for CRISPR Editing in Human Cells (CONT'D)

a positive one. So if you have better things to do than clone picking, choose your guide RNAs wisely (and let them cut close).

3. Optimizing "cut-to-mutation distance" for homo or heterozygous mutations

When thinking further about the distance relationship described above, we realized that, in addition to increasing mutation incorporation efficiency, we might exploit this relationship to direct the zygosity of the cells. This was important for us because the Alzheimer mutations we study are heterozygous in patients. As the probability of mutation incorporation drops with increasing cut-to-mutation distance, the mutation may only be inserted at one of the alleles, but not the other – the efficiency of editing is just too low for one to expect two successful mutations in a single cell.

We calculated the likelihood of obtaining homo- and heterozygous mutations by multiplying the single allelic mutation incorporation probabilities our data revealed, and found that for generating a homozygous HDR event with optimal efficiency, a guide RNA targeting a cut <10 bp from the desired mutation should be used. For a heterozygous event, the cut should ideally be around 5 to 20 bp away from the cut site. Lastly, if a heterozygous event is desired but only guide RNAs that target very close to the intended mutation site are available, a repair template containing blocking and intended mutations can be equally mixed with another ssODN template that contains the same CRISPR/Cas-blocking mutation but not the intended mutation. Here, the blocking-only templates will compete with the blocking/intended mutation templates resulting in cells that use one or the other on each target site.

Taken together, for efficient and specific introduction of a desired sequence change by HDR, the optimal guide RNAs should not only have sufficient on-target and very low off-target activity, but should also mediate a DSB at optimal distance from the intended mutation site, depending on the desired zygosity. Furthermore, re-editing activity of the guide RNA should be minimized by incorporating CRISPR/Cas-blocking mutations into the HDR repair template.

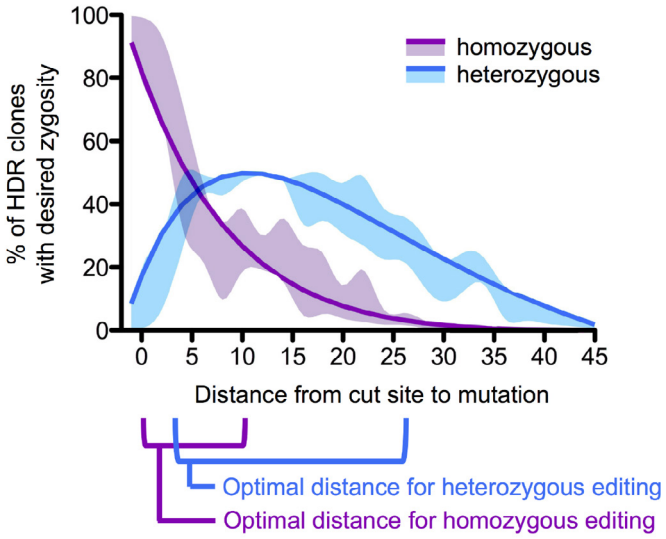


Figure 3: Optimal cut site to mutation distances for heterozygous and homozygous mutations.

3 Tips to Improve HDR Efficiency for CRISPR Editing in Human Cells (CONT'D)

References

1. Cong, Le, et al. "Multiplex genome engineering using CRISPR/Cas systems." Science 339.6121 (2013): 819-823. PubMed PMID: 23287718. PubMed Central PMCID: PMC3795411.

2. Mali, Prashant, et al. "RNA-guided human genome engineering via Cas9." Science 339.6121 (2013): 823-826. PubMed PMID: 23287722. PubMed Central PMCID: PMC3712628.

3. Paquet, Dominik, et al. "Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9." Nature 533.7601 (2016): 125-125. PubMed PMID: 27120160.

4. Kwart, Dylan, et al. "Precise and efficient scarless genome editing in stem cells using CORRECT." Nature protocols 12.2 (2017): 329-354. PubMed PMID: 28102837.

Non-homologous End Joining

By David Wyatt and Dale Ramsden | Apr 16, 2015, Updated Sep 1, 2020 by Christina Mork

One advantage to using the CRISPR/Cas system for genome engineering is the fact that Cas9 can be easily programmed to make a DNA double strand break (DSB) in the genome wherever the user chooses. After the initial cut, the next steps in the process involve repairing chromosomal DSBs. It is important to know that cells possess two major repair pathways – Non-Homologous End Joining (NHEJ) and Homology Directed Repair (HDR) – and how these pathways work, as this could be relevant when planning your experiment. Here, we'll discuss NHEJ, and how it impacts what happens to Cas9-mediated DSBs in the genome.

Non-homologous end joining steps

Unlike HDR, NHEJ is active throughout the cell cycle and has a higher capacity for repair, as there is no requirement for a repair template (sister chromatid or homologue) or extensive DNA synthesis. NHEJ also finishes repair of most types of breaks in tens of minutes – an order of magnitude faster than HDR. NHEJ is consequently the principle means by which CRISPR/Cas9-introduced breaks are repaired.

The following factors are required for NHEJ repair regardless of end structure, and dictate the major events of the pathway:

1. Broken ends are recognized by loading of the Ku70/Ku80 heterodimer.
2. Ku then acts as a scaffold for recruitment of a kinase (DNA-PKcs) and a two subunit DNA ligase (XRCC4-ligase IV); together with some accessory factors (PAXX, XLF), this complex holds a pair of DNA ends together, forming a paired end complex.
3. The paired end complex then ligates compatible DNA ends together, thus repairing the break.

This is a simplified, streamlined version of this pathway and does not consider the missing or damaged nucleotides that are common to biological sources of DSBs, and which need to be processed. Processing occurs prior to ligation as incompatible DNA ends interfere with that step. Accordingly, NHEJ has a vast toolbox of processing factors, including polymerases (Pol μ and Pol λ), nucleases (Artemis), and structure-specific end cleaning enzymes (Aprataxin, Tdp2) that function to make ends better substrates for ligation. Although we do not describe these steps here, the processing of DNA ends tends to be the point where mutations are introduced.

Repair of Cas9-induced breaks by NHEJ

As illustrated on the next page, NHEJ-mediated repair of Cas9-generated breaks is useful if the intent is to make a knockout in your gene of interest, as it is prone to generating indel errors. Indel errors generated in the course of repair by NHEJ are typically small (1-10 bp) but extremely heterogeneous. There is consequently about a two-thirds chance of causing a frameshift mutation. Of some importance, the deletion can be less heterogeneous when constrained by sequence identities in

Non-homologous End Joining (CONT'D)

flanking sequence ("microhomologies").

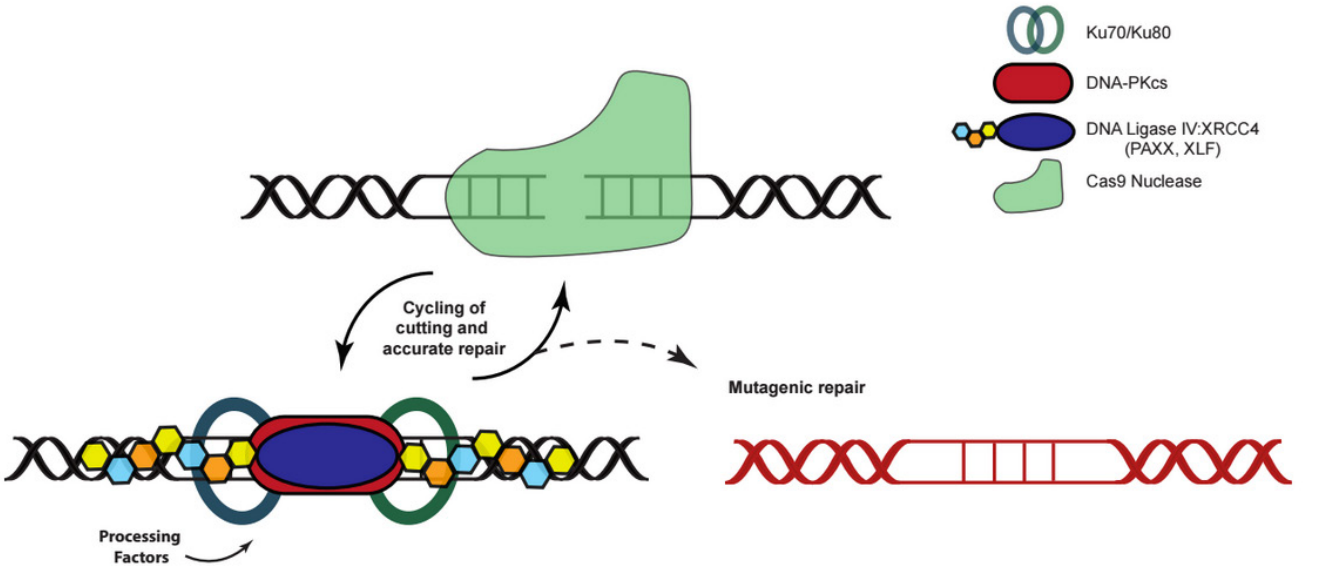


Image courtesy of D. Wyatt and D. Ramsden, UNC at Chapel Hill

Figure 1: The components involved in repairing a DSB using NHEJ.

NHEJ doesn't obligatorily introduce indels. Given the end structure of the Cas9 DSB (blunt or near-blunt ends without nucleotide damage) such products are rare, probably accounting for less than 5% of repair events. However, the products of accurate repair are easily re-cleaved while indel products aren't, so repeated cycles will favor accumulation of the latter products. As noted above, a single cycle of cleavage and accurate repair should take less than an hour, thus a population of cells constitutively expressing a targeted Cas9 should possess indels in the majority of their chromosomes within a day. Another factor expected to impact repair is that the Cas9 protein doesn't immediately release from the broken end after cleavage, which may interfere with loading of Ku and normal NHEJ activity. Other factors can also influence NHEJ activity. For example, repressing RecA and overexpressing the NHEJ machinery improved NHEJ accuracy in *M. smegmatis* (Yan et al., 2020).

NHEJ can also be engaged by variants of the canonical Cas9 approach. A pair of CRISPR guides that flank regions of hundreds or more DNA base pairs can simultaneously introduce a pair of chromosome breaks, and could result in deletion of the intervening DNA ("pop-out" deletions) if NHEJ joins the distal ends together. Similarly, it may be possible to directly insert an exogenous DNA fragment at a Cas9 targeted break (or pair of breaks) by NHEJ-dependent repair ("pop-in" insertion) provided a template containing compatible overhangs is available. Cas9 can also be altered to generate a targeted single strand break when two such breaks are introduced near each other, in opposite strands. This "double nickase" strategy vastly reduces breaks and mutations at off-target sites.

PITChing MMEJ as an Alternative Route for Gene Editing

By Mary Gearing | Feb 23, 2016

If you follow CRISPR research, you know all about using non-homologous end-joining (NHEJ) to make deletions or homology-directed repair (HDR) to create precise genome edits. But have you heard of another double-stranded break repair mechanism: MMEJ (microhomology-mediated end-joining)? MMEJ, a form of alternative end-joining, requires only very small homology regions (5-25 bp) for repair, making it easier to construct targeting vectors. Addgene depositor Takashi Yamamoto's lab has harnessed MMEJ to create a new method for CRISPR gene knock-in, termed PITCh (Precise Integration into Target Chromosomes). Using their PITCh plasmids, GFP knock-in cell lines can be created in about a month and a half, without the need for complicated cloning of homology arms.

MMEJ: An Introduction

There are three primary methods for repairing DNA after a double-stranded break. HDR copies the sequence from a repair template with flanking sequence homology for error-free DSB repair. NHEJ joins the ends of a DSB in an error-prone fashion, with insertions and deletions common. In contrast, MMEJ uses regions with 5-25 bp of microhomology flanking a DSB to repair DNA. The DNA ends are chewed back to reveal homology, allowing the strands to anneal. DNA synthesis then fills in the gaps. The end result is a deletion of the region between the microhomology and the retention of a single microhomology sequence.

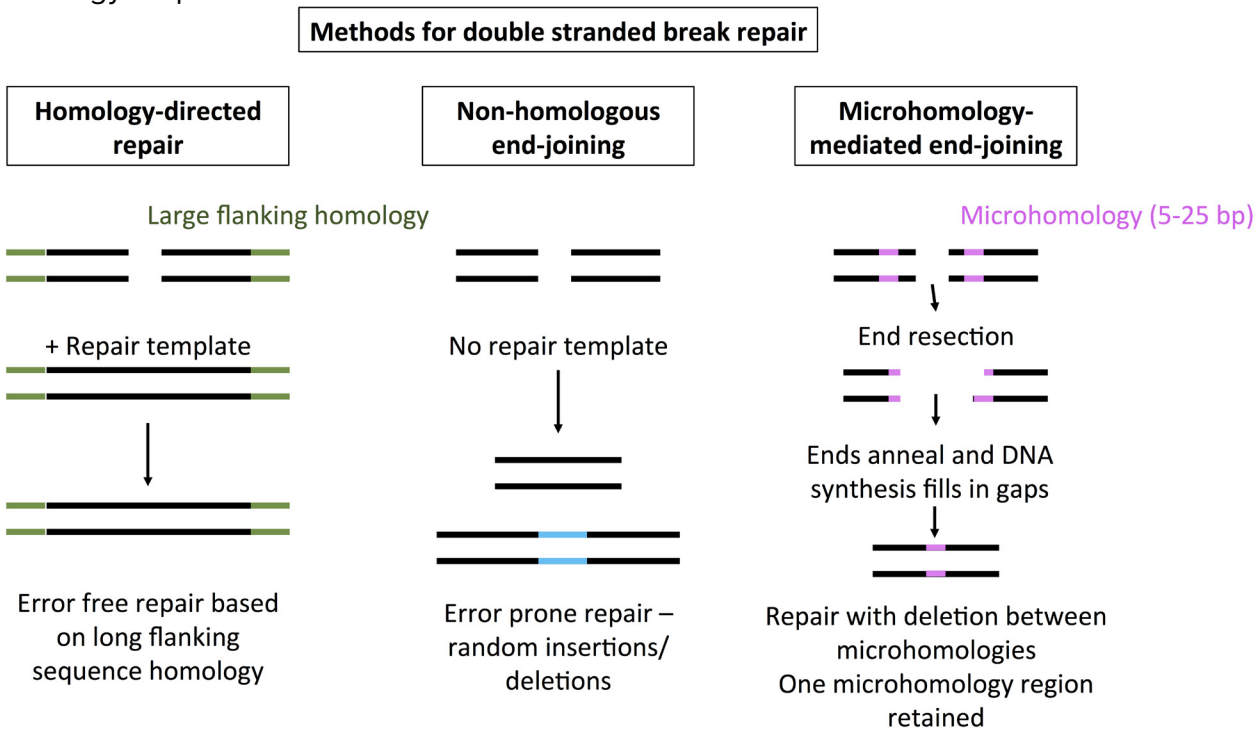


Figure 1: Three methods for double stranded break repair.

PITChing MMEJ as an Alternative Route for Gene Editing (CONT'D)

Compared to its counterparts NHEJ and HDR, MMEJ doesn't get a lot of press. However, this process accounts for a percentage of the mutations seen with TALENs and CRISPR. MMEJ is active during M and early S phases, when HDR is inactive, and the balance of NHEJ, HDR, and MMEJ repair varies from organism to organism. MMEJ doesn't yield a "perfect repair" like HDR, but it's much more predictable than NHEJ. As seen in Figure 1, short (5-25 bp) regions of homology flanking a double-stranded break yield precise deletion of the sequence between the microhomologies.

MMEJ has certain advantages over HDR. Some species don't have good HDR systems, and NHEJ will be favored even if a repair template is present. HDR also presents a cloning dilemma - the longer the homology, the more efficient the recombination, but with longer homology arms comes more time spent cloning. In contrast, the short homology required by MMEJ can easily be added via PCR amplification. Given the inefficiency of HDR for knock-ins, some labs have used NHEJ for whole plasmid integration; however, since NHEJ is error-prone, such a system is likely to introduce additional nucleotides flanking the insertion. If the DNA ends anneal incorrectly, MMEJ may also introduce, substitute, or delete nucleotides in addition to the expected deletion, but the frequency should be lower than that observed with NHEJ.

PITCh: Using MMEJ for gene knock-in

Building on the lab's previous work, Sakuma et al. describe a detailed protocol for MMEJ-mediated knock-in of a GFP-Puro cassette into a given locus, just upstream of a stop codon. Briefly, the PITCh vector should be constructed with 5' and 3' microhomology to the insertion locus flanking the GFP-Puro cassette. Three double stranded breaks are necessary for knock-in: one on either side of the GFP-Puro cassette and one in between the 5' and 3' microhomologies in the genomic locus. The first two breaks are induced via a generic PITCh-gRNA; the third break by an insertion locus-specific gRNA. These double stranded breaks allow for two sets of microhomologies (5' and 3') to anneal, knocking the GFP-Puro cassette into the locus (see Figure 2). The double-MMEJ strategy looks very similar to HDR, but it is accomplished using much smaller regions of homology, which facilitates easier cloning.

Abbreviated protocol for PITCh

Step 1: Generate microhomologies in the PITCh vector
~20 bp 5' and 3' microhomologies are added to the GFP-Puro cassette via PCR, and this construct is inserted into the PITCh vector via In-Fusion or SLIC cloning.

Step 2: Design an insertion locus-specific gRNA
The gRNA should target near the last coding exon of your gene of interest. For ideal use, this gRNA should be cloned into a vector containing Cas9 and the PITCh-gRNA.

PITChing MMEJ as an Alternative Route for Gene Editing (CONT'D)

Step 3: Contransfect the PITCh vector with the vector carrying Cas9 and the PITCh- and locus-specific gRNAs

Step 4: Select for puromycin resistant cells

Step 5: PCR amplify and sequence the locus to verify correct GFP-puro insertion
To lower the risk of off-target effects, Sakuma et al. optimized the PITCh-gRNA for minimal off-target binding in mammalian genomes as assessed by CRISPR design tools at crispr.mit.edu. They tested their system in HEK293 cells, integrating the GFP-Puro cassette into the FBL locus. Upon sequencing of puromycin resistant clones, they found that 80% and 50% of clones displayed proper insertion at the 5' and 3' junctions, respectively. These results indicate that PITCh is a robust method for genomic insertion. PITCh can also be adapted for whole-plasmid integration if you'd like to integrate a larger amount of material into the genome.

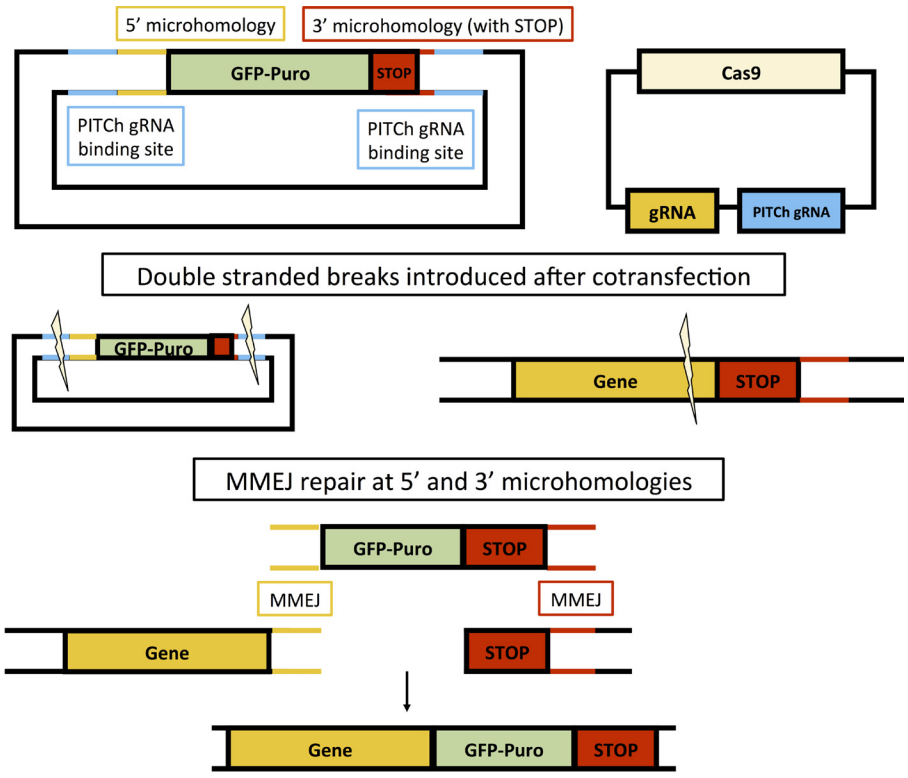


Figure 2: An overview of PITCh.

PITChing MMEJ as an Alternative Route for Gene Editing (CONT'D)

Open questions and alternative systems

The ready-made PITCh plasmids available from Addgene are perfect for expressing GFP from a given promoter, and the technique can be adapted to other transgenes. It's important to note that the fluorescence level observed will be dependent on both the endogenous promoter and the 3' UTR of the locus of interest, since the GFP-Puro will be inserted just upstream of a stop codon. One potential concern is if the GFP-Puro will alter expression of the gene it follows.

For increased versatility, it would be advantageous to adapt PITCh to insert genes into AAVS1, the "safe harbor locus" of the human genome, as shown by Dalvai et al., who used HDR to insert cDNA constructs into this locus. One important question to ask is how the efficiency of PITCh-based genomic insertion would compare to CRISPR sticky-end insertion using the nuclease Cpf1. Since Cpf1 cuts in a staggered pattern, it is thought to be ideal for HDR-independent knock-ins, but this possibility is still being explored.

More broadly, their use of MMEJ represents another strategy researchers can exploit for CRISPR gene editing. In organisms where HDR is downregulated, MMEJ represents another method for making targeted modifications. A recent publication by Zhang et al. shows just that - using MMEJ to insert FLAG tags into the genome of the pathogenic fungus *Aspergillus fumigatus*, which has been difficult to modify due to NHEJ's dominance over HDR in this species. As CRISPR technology continues to develop, it's become clear that the power of this editing platform lies in the diversity of nucleases and their applications. It will be interesting to see what new editing possibilities MMEJ can enable.

References

1. Sakuma, Tatsushi, et al. (2016). "MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems." Nat Protoc. 11(1): 118-33. PMID: 26678082.

2. Dalvai, Mathieu, et al. (2015). "A Scalable Genome-Editing Based Approach for Mapping Multiprotein Complexes in Human Cells." Cell Rep. 13: 621-633. PMID: 26456817.

3. Zhang, Chi, Xiuhua Meng, Xiaolei Wei, and Ling Lu. (2016). "Highly efficient CRISPR mutagenesis by microhomology-mediated end joining in *Aspergillus fumigatus*." Fungal Genet Biol. 86: 47-57. PMID: 26701308.

Validating Your Genome Edit

By Melina Fan | Jul 30, 2015 , Updated Oct 9, 2020 by Michael Lemieux

You've created your gRNA expression construct and used Cas9 to introduce it into your target cells. Hooray! You're ready to begin reading out data, right? Almost. In this article, we'll explain how to verify that your cells were appropriately edited. We'll also cover the basic techniques for detecting insertion, deletion, and mutation events.

Process overview

The method for validating your genome edit will vary by species and the type of edit. In this article, we will focus on diploid mammalian cells, but many of the principles will hold across different model organisms.

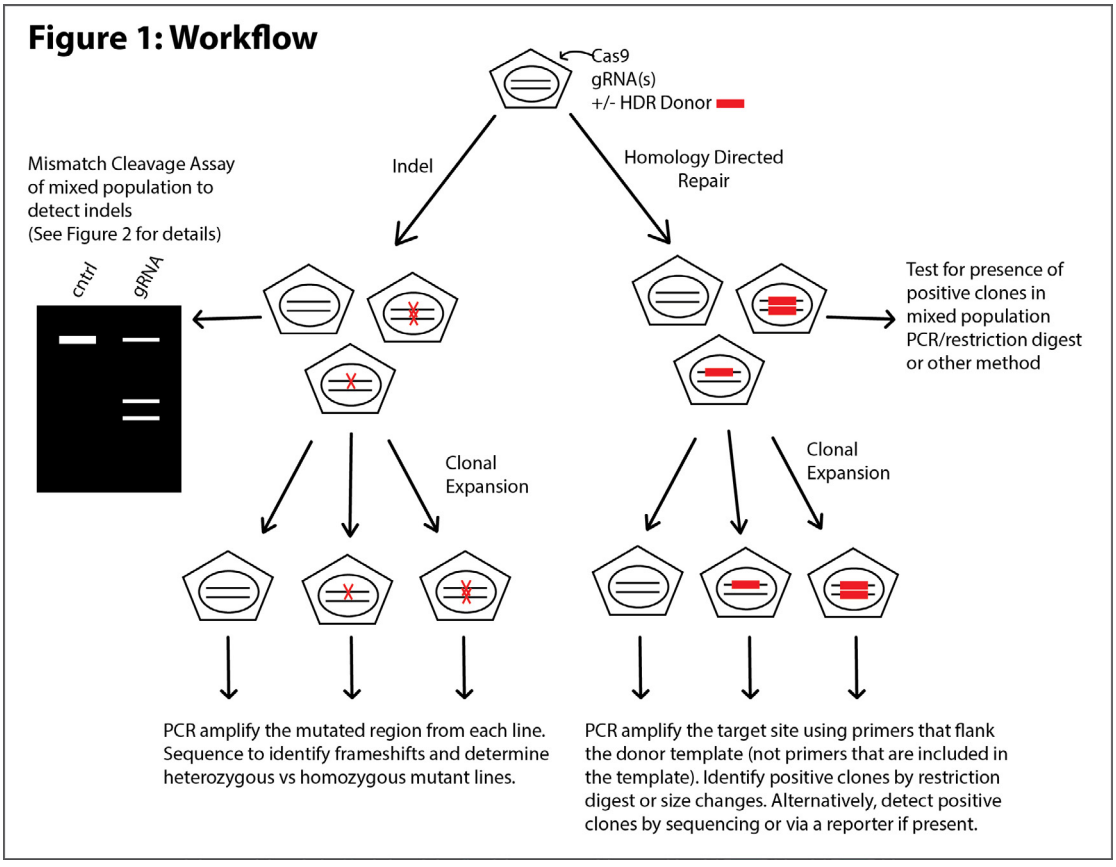


Figure 1: The workflow for verifying genome edits vary depending on type of edit (indel vs. HDR).

Introducing Cas9 and a gRNA into your cells (possibly along with a donor template) will result in a mixed population of cells. Following the introduction of a Cas9-mediated double strand break (DSB)

Validating Your Genome Edit (CONT'D)

in mammalian cells, cellular machinery repair the DSB by non-homologous end joining (NHEJ) or homology directed repair (HDR). Repair via the NHEJ pathway predominates in mammalian cells resulting in the creation of indel errors, short heterogeneous insertions and deletions of nucleic acid sequences, at the site of the DSB. In addition to the heterogeneity of indels introduced at Cas9-induced DSBs, allelic editing frequencies will vary as well. The HDR pathway requires the presence of a repair template, which is used to fix the DSB in a more specific manner. HDR faithfully copies the sequence of the repair template to the cut target sequence. Some cells will not be edited, some will have one allele edited, and some will have both alleles edited.

The first step in the validation process is to quickly assess whether a significant number of the cells have been edited (see Figure 1). For indels, this is visualized using a mismatch cleavage assay (see Figure 2). For HDR, this is often visualized by a change in the restriction pattern at the site of interest or via a reporter readout. For deletions (see Figure 3), this is visualized by a decrease in size of a PCR product produced by primers flanking the region to be deleted.

Once you know that a portion of your cells have been edited, you can go on to create clonal cell lines. Serial dilutions can be used to isolate single cells followed by an expansion period to generate these lines. If you have a fluorescent protein marker on your plasmid, you can use FACS to enrich the cells that received Cas9 and your gRNA. After expansion, assay each cell line and sequence the region of interest in order to validate the genome edit as described in Figure 1. But don't stop there, because you can potentially validate your edit at four different levels (DNA, RNA, protein, and phenotype). Depending on the nature of your edit, something like qRT-PCR could be used to demonstrate a difference in mRNA transcript levels. A western blot can assess protein expression, and phenotypic analyses will be crucial to demonstrate the functional consequences of your edit. We'll focus mostly on validation at the DNA level, so read on to learn more!

Mismatch cleavage assay to detect indels

A mismatch cleavage assay is a quick and easy way to detect indels. Surveyor™ nuclease is commonly used for this purpose, as it cleaves both DNA strands 3' to any mismatches. It can detect indels of up to 12 nucleotides and is sensitive to mutations present at frequencies as low as 1 in 32 copies.

Mismatch cleavage assays typically consist of four steps: 1) PCR amplify the region of interest, 2) denature the strands and rehybridize to allow for the mutant and wild-type strands to anneal, 3) treat annealed DNA with Surveyor™ nuclease to cleave heteroduplexes, and 4) analyze DNA on an agarose gel or other instrument that separates DNA based on size. Figure 2 illustrates how the assay works. In this example, both +gRNA lanes contain cleaved fragments of the expected sizes, indicating that the gRNAs successfully produced indels in the target region. This assay is often used semi-quantitatively, and in this case, gRNA1 appears to be more efficient at producing indels than gRNA2. This is because

Validating Your Genome Edit (CONT'D)

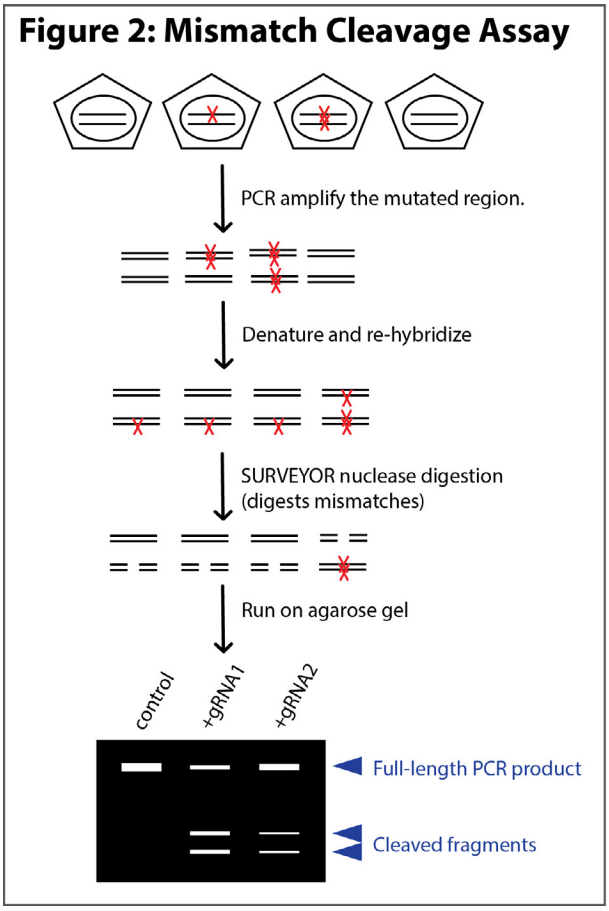


Figure 2: The mismatch cleavage assay quickly detects indels using a nuclease that cleaves mismatched strands after PCR.

of the greater intensity of the lower molecular weight bands on the gel for gRNA1, which suggests that it generated more indels that could be recognized and cleaved by Surveyor™ nuclease.

If you'd prefer an alternative to the mismatch cleavage assay as a means of semi-quantitatively detecting indels, you could also try RFLP analysis. RFLP stands for Restriction Fragment Length Polymorphism, and can be used when indels either create or abolish restriction endonuclease recognition sites. Following PCR amplification of the target region and subsequent digestion with appropriate enzymes, the DNA can be run on an agarose gel or other instrument that separates DNA based on size, and the restriction pattern analyzed to determine the degree of indel formation in your polyclonal cell population. RFLP analysis has an advantage over the mismatch cleavage assay because it can detect homozygous mutations. The major disadvantage is that you must have an altered restriction recognition site for this approach to work.

Detect homology directed repair

If you want to mutate your region of interest using HDR, it is advisable to first determine whether your gRNA efficiently cutts your target sequence by creating indels and conducting a mismatch cleavage assay. Once you've selected your optimal gRNA, introduce it along with Cas9 and your repair template to drive HDR.

When designing your HDR donor template, plan ahead for detection of integration events. For instance, you could purposefully introduce or remove restriction sites which would alter the digestion pattern of PCR products. Alternatively, you could include a reporter element for detection of HDR at the DNA, RNA, or protein level. A large insertion or deletion after integration could also be detected by a size change in the PCR product. Single nucleotide changes can be quickly assayed using restriction digests if the polymorphism creates/removes a restriction digest site. Otherwise, single nucleotide changes can be detected by TA cloning and Sanger sequencing, next generation sequencing, or droplet digital PCR genotyping.

Validating Your Genome Edit (CONT'D)

Remember that performing serial dilutions and generating clonal cell lines will help with validating your genomic edit. This is because the CRISPR-Cas9 system makes different edits across different cells in the population, and this can convolute the interpretation of your results. For example, if you sequence from a mixed population, you'll see many overlapping peaks in a Sanger sequencing chromatogram. Sanger amplicon sequencing can give you a general idea that some editing occurred, but you won't know which cells were successfully modified.

HDR events are generally less frequent than indels, so you will likely need to screen a larger number of colonies to create a clonal line. The number of clones that need to be screened will depend on both your transfection/transduction efficiency and HDR frequency. For example, if you have 40% transfection efficiency and 5% HDR efficiency, approximately 0.4x0.05=0.02 or 2% of your cells will have the recombined region. Thus, you should plan to screen at least 50 colonies. If you are able to select cells that have been successfully transfected/transduced using a marker, then you may be able to test fewer colonies.

PCR to detect deletions

Most deletions are created by using two gRNAs that direct Cas9 to cleave out the intervening region of DNA. The deletion can thus be detected by conducting a PCR using primers flanking the deleted region. The workflow is similar to that described in Figure 1. Figure 3 provides an example of PCR results obtained by screening a panel of clonal lines for deletions. In this example, clones 1, 5, and 7 are heterozygous for the deletion and clone 4 is homozygous for the deletion.

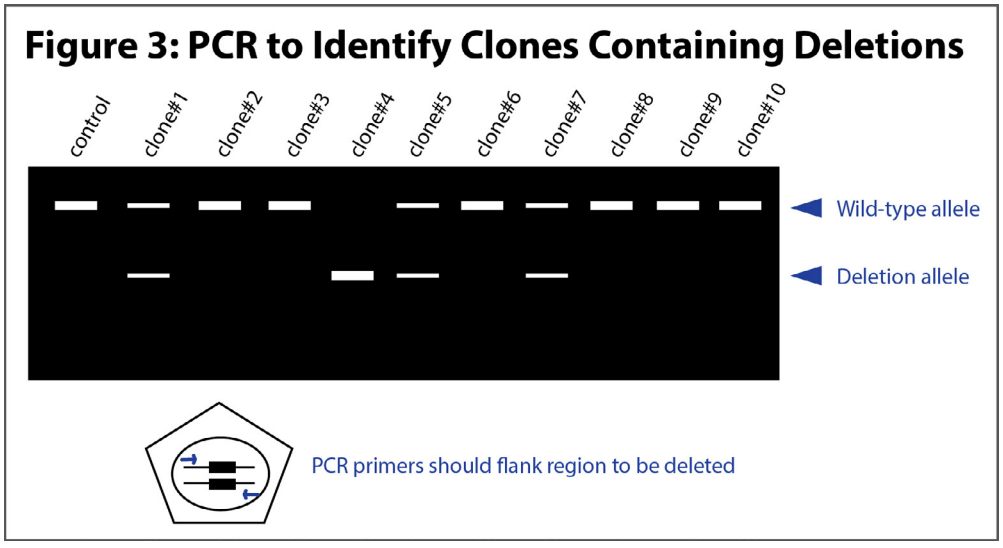


Figure 3: Using PCR to detect homozygous vs. heterozygous clones for the deletion.

Validating Your Genome Edit (CONT'D)

Next Generation Sequencing to validate edits and detect off-target effects

If your lab has the resources, you can quantitatively assess genome edits in your target sequence and other regions of the genome using next generation sequencing (NGS). NGS is a good option if you have a large number of samples and/or want to simultaneously look at off-target changes. When using this method, it is important to keep a set of control cells as you will need to compare the sequencing reads from your edited sample to this untreated population. Software such as CRISPResso can help with the data analysis.

But what if your lab doesn't have the resources for NGS? You still have options for detecting off-target effects. You could use Sanger sequencing, provided that you have some idea of where in the genome you might expect off target editing to occur. Using *in silico* prediction analyses like CRISPRitz or CRISPOR can help narrow down the list of sites where off target editing may occur, allowing you to Sanger sequence a manageable number of the higher probability locations.

If you're worried about off-target effects, you can also think about using a version of Cas9 that has been specifically engineered for reduced off-target editing, such as SpCas9-HF1, eSpCas9(1.1), or HypaCas9.

The techniques described in this post are not CRISPR-specific and can also be used for assessing genome edits created by TALEN or Zinc Finger Nucleases. Regardless of what method you use, validating your edit is time well spent as you prepare for your future experiments.

References

Cancellieri S, Canver MC, Bombieri N, Giugno R, Pinello L (2019) CRISPRitz: rapid, high-throughput and variant-aware *in silico* off-target site identification for CRISPR genome editing. Bioinformatics 36:2001–2008 . https://doi.org/10.1093/bioinformatics/btz867

Haeussler M, Schönig K, Eckert H, Eschstruth A, Mianné J, Renaud J-B, Schneider-Maunoury S, Shkumatava A, Teboul L, Kent J, Joly J-S, Concordet J-P (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol 17: . https://doi.org/10.1186/s13059-016-1012-2

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8:2281–2308 . https://doi.org/10.1038/nprot.2013.143

Validating Your Genome Edit (CONT'D)

Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, Gerard GF (2004) Mutation detection using Surveyor™ nuclease. BioTechniques 36:702–707 . https://doi.org/10.2144/04364pf01

Sequencing Options for CRISPR Genotyping

By Soren Hough | Oct 4, 2016

One of the most important steps in the CRISPR experimental process is validating edits. Regardless of which CRISPR genome editing system you use, there remains a chance that the observed phenotype was caused by an off-target mutation and not an edit in the target gene.

The validation process, also known as CRISPR genotyping, is critical to demonstrating causal relationships between genotype and assayed phenotype. Verifying these connections can help alleviate the reproducibility crisis in biology. It is key to address these concerns as CRISPR use grows across the life sciences and to establish standardized validation techniques for academia, industry, and especially the clinic.

Popular validation assays are insufficient

As discussed in Validating Your Genome Edit, there are a variety of options for CRISPR genotyping. The most common options include mismatch cleavage assays, such as Surveyor™, T7E1, and Sanger sequencing. However, recent studies suggest that both Surveyor™ and Sanger may not be adequate standards for validating edits.

Mismatch cleavage assays rely on pairing between the edited strand and wild-type strand of the host DNA. When these strands hybridize, the nuclease can detect strands with mismatches and cleave them. The results are then visualized using gel electrophoresis.

Surveyor™ and T7E1 have been widely adopted due to their relative simplicity and low cost. The problem with these assays is that they do not provide sequence-level data. They also have a limit of detection of ~5%. This means they do not reliably detect editing events that occur in less that 5% of the population (Fu et al. 2013, Vouillot et al. 2015).

Meanwhile, Sanger sequencing is laborious, time consuming and cannot be applied to heterogenous populations (Bell et al. 2014). Further, Sanger sequencing has a lower detection limit of 50-20% (although this has been improved in some studies) (Davidson et al. 2012, Tsiatis et al. 2010). As the field moves toward standardized thresholds for validating CRISPR experiments, many are turning to next-generation sequencing options over older assays.

Biased sequencing methods

There are two primary methods of off-target detection: biased and unbiased. Biased techniques only sequence certain sites in the genome predicted to contain off-target cleavage events. Unbiased techniques search the whole genome for off-target sites irrespective of *in silico* prediction.

Sequencing Options for CRISPR Genotyping (CONT'D)

These techniques differ in important ways, but can also complement one another by providing both broad and specific details on genome sequencing. Used in concert, these approaches can provide the researcher with a reasonable level of certainly that the effects they see are not due to off-targets. This is a valuable step toward enhancing confidence and reproducibility of a study's findings.

Prediction algorithms: A good place to start for biased validation

At the moment, many software tools predict off-target effects of sgRNAs using computational methods. They identify possible off-target sites across the genome and pinpoint the location of mismatches based on the sequences of the genome and sgRNA. This is a good starting point for most researchers as it provides a list of putative off-target sites that they can later sequence for mutations.

One method a researcher can use to test predicted off-target sites following a CRISPR experiment is targeted amplicon sequencing. The information from targeted amplicon sequencing is highly sensitive with detection levels as low as .01% (Hendel et al. 2015). Low detection rates mean the investigator can be relatively certain that their samples don't have off-target mutations if they remain undetected using these techniques.

Frequencies of off-target mutations are essential data points for investigators looking to definitively link genotype and phenotype. It is also key to perform these validations as translational researchers begin to use CRISPR as a therapy. Low frequency off-target effects may generate irreproducible data in a research setting, but these events could have disastrous health effects in the clinic. NGS-based methods provide the most complete information profile regarding putative off-target sites including both the edit rate and the repair product sequence.

Targeted amplicon sequencing doesn't tell the whole story

Even though progress has been made with off-target prediction algorithms, their genome-wide search criteria are not exhaustive. Mismatch tolerance settings are often limited to off-target sites of <4 bp. The off-target list is also generally weighted by the position of the mismatch along the length of the gRNA given the stricter sequence requirement at the terminal 3' PAM site (Fu et al. 2013; Pattanayak et al. 2013).

This approach misses larger mismatches (e.g. six nucleotides) that may still lead to off-target double-stranded breaks (Tsai et al. 2015). Additionally, current algorithms do not take into account other elements, including those relating to DNA structure (e.g. epigenetic modification, bulges) that may also impact off-target edits. As a result, only sequencing sites predicted by conventional algorithms may not provide a full picture of the impact of CRISPR editing in the model cell line or organism.

Sequencing Options for CRISPR Genotyping (CONT'D)

Several options exist for unbiased off-target detection, including Digenome-seq (Kim et al. 2015) for *in vitro* analysis, IDLV for *in vivo* detection (Gabriel et al. 2011, Wang et al. 2015, Osborn et al. 2016) and HTGTS (Frock et al. 2015) for cell-based experiments. These strategies can be used in concert with *in silico* prediction to create a more comprehensive list of off-target editing events. Two of the most common cell-based methods are genome-wide, unbiased identification of double-strand breaks (DSBs) evaluated by sequencing (GUIDE-seq) (Tsai et al. 2015) and direct *in situ* breaks labeling, enrichment on streptavidin, and next-generation sequencing (BLESS) (Crosetto et al. 2013).

GUIDE-seq and BLESS detect double-stranded breaks and do not require high sequencing read counts making them fast and viable options for multiplex sequencing in many laboratories. Nevertheless, unbiased detection isn't as sensitive as targeted amplicon sequencing. For example, GUIDE-seq seems to have a minimum detection limit of 0.1% (Tsai et al. 2015). This contrasts with detection frequencies of 0.01% in amplicon sequencing (Hendel et al. 2015), a significant difference as CRISPR experiments move closer to the clinic (Tsai and Joung, 2016).

Table 1: Methods for CRISPR genotyping.

Technique	Detection Limit	Application	Advantages	Disadvantages
GUIDE-Seq	0.1%	Cell-based	Searches the genome for all DSBs, doesn't require high read counts, fast multiplexing	Requires delivery of dsODN (potentially toxic)
Digenome Seq	0.1%	Cell-free (*in vitro*)	Works across all cell types	Must be verified with cell-based method
IDLV	1%	Cell-based	Programmable, can detect DSBs in live cells	Not as senstive as other unbiased methods, high background
BLESS	Not reported	Cell-based (*in vitro*)	Can be used on tissue from whole animal models, no exogenous component required (e.g. dsODN), doesn't require high read counts (fast multiplexing)	Requires large cell population, senstive to time since cell fixing
HTGTS	Not reported	Cell-based	Identifies translocations	Limited by chromatin configuration, produces many false negatives

Sequencing Options for CRISPR Genotyping (CONT'D)

Combining sequencing techniques can ensure validated experiments

Unbiased detection methods are excellent for finding evidence of DSBs throughout the genome. However, their decreased sensitivity means that the best option moving forward may be to integrate both biased and unbiased approaches. As suggested in a review by Tycko et al., 2016, unbiased sequencing and *in silico* prediction should give a broad picture of all possible editing events in the genome; from there, amplicon sequencing can evaluate and validate off-target sites in a highly accurate manner.

Using both of these approaches may not be necessary for every CRISPR experiment. Off-target events due to >3 bp mismatches or that are sequence-independent are rare, but they are detectable using just genome-wide unbiased methods. However, most investigators use single cell clones for *in vitro* CRISPR experiments. The likelihood that a single cell clone derived from the pool contains both the rare off target event and the desired edit is low. Therefore, unbiased sequencing may not be worth the cost and labor when single clones are selected. Conversely, translational research may require the rigor of both forms of off-target analysis in order to meet clinical approval.

It is key to maintain a consistent set of standards as the field seeks to generate reproducible, quality data on the role of genetic networks in biological systems. NGS will also play significantly into the realm of clinical therapeutic development as CRISPR is used not only to study disease, but to treat patients, as well. For more information and a detailed overview of the aforementioned sequencing techniques, please see "Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity" by Tycko et al., 2016 and "Defining and improving the genome-wide specificities of CRISPR–Cas9 nucleases" by Tsai and Joung, 2016.

References

1. Fu, Yanfang, et al. "High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells." Nature biotechnology 31.9 (2013): 822-826. PubMed PMID: 23792628. PubMed Central PMCID: PMC3773023.

2. Vouillot, Léna, Aurore Thélie, and Nicolas Pollet. "Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases." G3: Genes| Genomes| Genetics 5.3 (2015): 407-415. PubMed PMID: 25566793. PubMed Central PMCID: PMC4349094.

3. Bell, Charles C., et al. "A high-throughput screening strategy for detecting CRISPR-Cas9 induced mutations using next-generation sequencing." BMC genomics 15.1 (2014): 1. PubMed PMID: 25409780.

Sequencing Options for CRISPR Genotyping (CONT'D)

PubMed Central PMCID: PMC4246457.

4. Davidson CJ., et al. "Improving the limit of detection for Sanger sequencing: a comparison of methodologies for KRAS variant detection." Biotechniques 53.3 (2012): 182-188. PubMed PMID: 22963480. PubMed Central.

5. Tsiatis, Athanasios C., et al. "Comparison of Sanger sequencing, pyrosequencing, and melting curve analysis for the detection of KRAS mutations: diagnostic and clinical implications." The Journal of Molecular Diagnostics 12.4 (2010): 425-432. PubMed PMID: 20431034. PubMed Central PMCID: PMC2893626.

6. Hendel, Ayal, et al. "Quantifying on-and off-target genome editing." Trends in biotechnology 33.2 (2015): 132-140. PubMed PMID: 25595557. PubMed Central PMCID: PMC4308725.

7. Pattanayak, Vikram, et al. "High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity." Nature biotechnology 31.9 (2013): 839-843. PubMed PMID: 23934178. PubMed Central PMCID: PMC3782611.

8. Tsai, Shengdar Q., et al. "GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases." Nature biotechnology 33.2 (2015): 187-197. PubMed PMID: 25513782. PubMed Central PMCID: PMC4320685.

9. Kim, Daesik, et al. "Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells." Nature methods 12.3 (2015): 237-243. PubMed PMID: 25664545.

10. Gabriel, Richard, et al. "An unbiased genome-wide analysis of zinc-finger nuclease specificity." Nature biotechnology 29.9 (2011): 816-823. PubMed PMID: 21822255.

11. Wang, Xiaoling, et al. "Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors." Nature biotechnology 33.2 (2015): 175-178. PubMed PMID: 25599175.

12. Osborn, Mark J., et al. "Evaluation of TCR gene editing achieved by TALENs, CRISPR/Cas9, and megaTAL nucleases." Molecular Therapy (2015). PubMed PMID: 26502778.

13. Frock, Richard L., et al. "Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases." Nature biotechnology 33.2 (2015): 179-186. PubMed PMID: 25503383. PubMed Central PMCID: PMC4320661.

Sequencing Options for CRISPR Genotyping (CONT'D)

14. Crosetto, Nicola, et al. "Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing." Nature methods 10.4 (2013): 361-365. PubMed PMID: 23503052. PubMed Central PMCID: PMC3651036.

15. Tsai, Shengdar Q., and J. Keith Joung. "Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases." Nature Reviews Genetics17.5 (2016): 300-312. PubMed PMID: 27087594.

16. Tycko, Josh, Vic E. Myer, and Patrick D. Hsu. "Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity." Molecular Cell 63.3 (2016): 355-370. PubMed PMID: 27494557.

The PAM Requirement and Expanding CRISPR Beyond SpCas9

By Joel McDade | Nov 12, 2015, Updated Aug 20, 2020 by Jennifer Tsang

Cas9 can be used to modify any desired genomic target provided that (1) the sequence is unique compared to the rest of the genome and (2) the sequence is located just upstream of a Protospacer Adjacent Motif (PAM sequence). The 3-5 nucleotide PAM sequence serves as a binding signal for Cas9 and this sequence is a strict requirement for Cas9-mediated DNA cleavage.

The need for more PAM sequences

While PAM sequences for the commonly used *S. pyogenes* Cas9 (3'-NGG) are abundant throughout the human genome, they are not always positioned correctly to target a particular gene. Furthermore, a target sequence may have high homology elsewhere in the genome. These off-target sequences may be unintentionally mutated along with the desired target locus.

The PAM sequence is of particular concern when trying to edit a gene using homology directed repair, since HDR-mediated gene editing is most efficient when target sites are located in close proximity to the region to be edited. In this article, we will cover three ways to circumvent this limitation: 1) the use of novel *S. pyogenes* Cas9 variants with varying PAM sequences, 2) the use of Cas9 homologs derived from species other than *S. pyogenes*, and 3) the use of non-Cas9 enzymes.

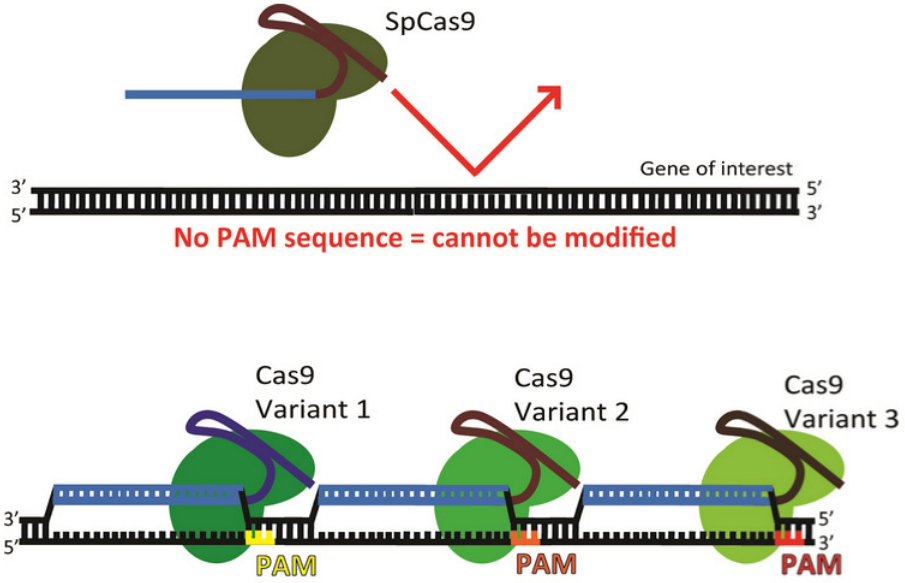


Figure 1: Without a corresponding PAM seqeunce, SpCas9 can't target the region. Cas9 variants with different PAM site requirements can target different DNA regions.

Synthetic *S. pyogenes* Cas9s with novel PAM recognition

In 2015, Keith Joung's lab performed a series of positive selection screens in bacteria to identify mutants of *S. pyogenes* Cas9 that were able to cleave target DNA sequences located upstream of either NGA or NGC PAM sequences (Kleinstiver et al. 2015). From these screens, they identified four novel SpCas9 variants with altered PAM binding specificity:

- D1135E variant (D1135E): NGG
- VQR variant (D1135V, R1335Q and T1337R): NGAN or NGNG

The PAM Requirement and Expanding CRISPR Beyond SpCas9 (CONT'D)

- EQR variant (D1135E, R1335Q and T1337R): NGAG
- VRER variant (D1135V, G1218R, R1335E and T1337R): NGCG

The D1135E variant is far more selective for the canonical *S. pyogenes* PAM sequence NGG compared to wild-type SpCas9, which also displays some cleavage with an NGA PAM. This variant may increase the specificity of genome modifications at DNA targets adjacent to NGG PAM sequences when used in place of wild-type SpCas9. The remaining variants (VQR, EQR, and VRER) recognize novel PAM sequences. The VQR, EQR, and VRER Cas9 variants are capable of cleaving genomic DNA in mammalian cells and zebrafish embryos, and they can be used to modify genomic loci that cannot be modified using wild-type SpCas9. The number of off-target cleavage events for the VQR and VRER variants is similar to wild-type SpCas9, indicating that the variants are likely just as selective as wild-type SpCas9. These VQR, EQR and VRER SpCas9 variants effectively double the targeting range of CRISPR/Cas9 within the human genome.

Another variant, xCas9 3.7 has 7 mutations found in the REC2, REC3, and PAM interacting domains and allows for expanded PAM recognition as well as increased specificity and lower off-target activity (Hu et al., 2018).

In 2020, the Kleinstiver lab reported the development of near-PAMless Cas9 variants. These variants, named SpG and SpRY, can target NGN PAMs and NRN and NYN PAMs, respectively (Walton et al., 2020).

Characterization of Cas9 from additional bacterial species

Many more Cas9 orthologs have been isolated, and our understanding of their PAMs is shown below:

- *Streptococcus pyogenes* (Sp): 3' NGG
- *Staphylococcus aureus* (Sa): NGRRT or NGRRN
- *Neisseria meningitidis* (Nm or Nme): NNNNGATT
- *Campylobacter jejuni* (Cj): NNNNRYAC
- *Streptococcus thermophilus* (St): NNAGAAW
- *Treponema denticola* (Td): NAAAAC

Non-SpCas9's bind a variety of PAM sequences, which makes them useful when no suitable SpCas9 PAM sequence is present. Furthermore, non-SpCas9's may have other characteristics that make them more useful than SpCas9. For example, Cas9 from *Staphylococcus aureus* (SaCas9) is about 1 kilobase smaller than SpCas9, so it can be packaged into adeno-associated virus (AAV).

At 984 amino acids in length, Cas9 from *Campylobacter jejuni* (CjCas9) is even smaller than SaCas9 and is also compatible with AAV delivery. NmCas9, another small ortholog, displays lower off-target

The PAM Requirement and Expanding CRISPR Beyond SpCas9 (CONT'D)

editing than wild-type SpCas9, even when targeting sites that are known to produce off-target editing with SpCas9 (Amrani et al., 2018). When choosing a Cas9, remember to check that your tracrRNA and crRNA (or synthetic gRNA) are derived from the same species.

Expanding the CRISPR toolbox

The isolation of novel CRISPR proteins has and will continue to dramatically increase the number of CRISPR applications. The first non-Cas9 CRISPR protein adapted for genome engineering was Cas12a (formerly Cpf1), a nuclease that generates double strand breaks in target genes resulting in the formation of "sticky ends" rather than the blunt ends created by Cas9. Cas12a displays lower off-target editing than SpCas9 (Kim et al., 2016). The *Acidaminococcus* sp. BV3L6 Cas12a, AsCpf1, uses a 5'-TTN PAM which makes it easier to target AT-rich genomes (Kleinstiver et al., 2016).

Type VI CRISPR systems, which target RNA, offer additional targeting flexibility. Cas13a (formerly C2c2) has been adapted for targeted RNA cleavage in mammalian cells. The REPAIR (RNA Editing for Programmable A to I Replacement) system fuses Cas13b to RNA deaminase ADAR2 to create a specific RNA editor. Cas13 enzymes are advantageous because they do not require a PAM, and RNA targeting is potentially reversible since there is no genomic edit. Some Cas13 enzymes require a single base protospacer flanking sequence (PFS) adjacent to the target, but many do not, showing the flexibility of this system.

References

Abudayyeh OO, Gootenberg JS, Essletzbichler P, Han S, Joung J, Belanto JJ, Verdine V, Cox DBT, Kellner MJ, Regev A, Lander ES, Voytas DF, Ting AY, Zhang F (2017) RNA targeting with CRISPR–Cas13. Nature 550:280–284 . https://doi.org/10.1038/nature24049

Amrani N, Gao XD, Liu P, Edraki A, Mir A, Ibraheim R, Gupta A, Sasaki KE, Wu T, Donohoue PD, Settle AH, Lied AM, McGovern K, Fuller CK, Cameron P, Fazzio TG, Zhu LJ, Wolfe SA, Sontheimer EJ (2018) NmeCas9 is an intrinsically high-fidelity genome-editing platform. Genome Biol 19: . https://doi.org/10.1186/s13059-018-1591-1

Cox DBT, Gootenberg JS, Abudayyeh OO, Franklin B, Kellner MJ, Joung J, Zhang F (2017) RNA editing with CRISPR-Cas13. Science 358:1019–1027 . https://doi.org/10.1126/science.aaq0180

Kim D, Kim J, Hur JK, Been KW, Yoon S, Kim J-S (2016) Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. Nat Biotechnol 34:863–868 . https://doi.org/10.1038/nbt.3609

The PAM Requirement and Expanding CRISPR Beyond SpCas9 (CONT'D)

Kim E, Koo T, Park SW, Kim D, Kim K, Cho H-Y, Song DW, Lee KJ, Jung MH, Kim S, Kim JH, Kim JH, Kim J-S (2017) *In vivo* genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. Nat Commun 8: . https://doi.org/10.1038/ncomms14500

Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales APW, Li Z, Peterson RT, Yeh J-RJ, Aryee MJ, Joung JK (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature 523:481–485 . https://doi.org/10.1038/nature14592

Kleinstiver BP, Tsai SQ, Prew MS, Nguyen NT, Welch MM, Lopez JM, McCaw ZR, Aryee MJ, Joung JK (2016) Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. Nat Biotechnol 34:869–874 . https://doi.org/10.1038/nbt.3620

Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, Zeina CM, Gao X, Rees HA, Lin Z, Liu DR (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature 556:57–63 . https://doi.org/10.1038/nature26155

Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F (2015) *In vivo* genome editing using *Staphylococcus aureus* Cas9. Nature 520:186–191 . https://doi.org/10.1038/nature14299

Walton RT, Christie KA, Whittaker MN, Kleinstiver BP (2020) Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. Science 368:290–296 . https://doi.org/10.1126/science.aba8853

Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F (2015) Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell 163:759–771 . https://doi.org/10.1016/j.cell.2015.09.038

Choosing a CRISPR Nuclease: Site Accessibility, Specificity, and Sensitivity

By Andrew Hempstead | Nov 5, 2019

Considerations when choosing your Cas nuclease

When choosing the Cas9 nuclease for your experiments, there are three important factors to consider: site accessibility, specificity, and sensitivity.

Site accessibility

Site accessibility is a measure of which genomic sequences can be targeted by the Cas9. This is determined by the PAM sequence recognized by the nuclease. For the most frequently used Cas9, spCas9 from *Streptococcus pyogenes*, this sequence is 3'-NGG. But Cas9s found in other species can recognize a variety of different PAM sequences. Researchers have also used experimental evolution to develop Cas9s that have a broader PAM specificity, allowing for the targeting of sites that were previously unavailable.

Specificity

Specificity is the assessment of the off-target activity of Cas9. This can occur when Cas9 cleaves a site that is not a perfect match to the guide RNA (gRNA) used to target the genomic locus of interest. As these off-target effects can be genome-wide and, in some cases, difficult to determine, the potential for off-target cleavage should be an important consideration during any experiment.

Sensitivity

Cas9 sensitivity is often closely related to specificity. Sensitivity is a measure of the on-target activity of the Cas9 nuclease. This can be determined by measuring the indel frequency at a target site. A specificity ratio can be measured by calculating the on-target activity divided by the off-target activity.

Cas9 nucleases that exhibit enhanced site accessibility

In order to generate a genomic alteration at a specific location, it must be accessible by Cas9. Our previous article, The PAM Requirement and Expanding CRISPR Beyond SpCas9, described the PAM specificity of Cas9s isolated from different organisms and engineered Cas9 variants that were able to target different PAM sites (NGG, NGAN, NGNG, NGAG, and NGCG). Since then, the Osamu Nureki Lab described a rationally engineered Cas9, SpCas9-NG, which required only the di-nucleotide PAM NG (Nishimasu et al., 2018). The David Liu Lab similarly described xCas9, which displayed the ability to target multiple PAM sequences (NG, GAA, and GT), greatly increasing the number of genomic loci that can be targeted (Hu et al., 2018).

These altered variants can allow for targeting of your genomic loci of interest that may not be

Choosing a CRISPR Nuclease: Site Accessibility, Specificity, and Sensitivity (CONT'D)

accessible by wild-type Cas9, due to its stringent PAM specificity. This can be especially important for base editing techniques, where Cas9 must be directed to a very specific site.

Specificity vs. sensitivity in choosing the Cas9 to use in your experiments

Another important consideration when choosing a Cas9 is balancing specificity and sensitivity. Wild-type Cas9 exhibits high on-target activity, but also high off-target activity, which may be undesirable in many applications. For example, the Feng Zhang Lab had described eSpCas9, while the Keith Joung Lab had developed SpCas9-HF1. Both of the enzymes showed decreased off-target activity relative to Cas9. Since then, the Joung and Jennifer Doudna Labs used targeted mutagenesis to generate hyper accurate Cas9 (HypaCas9). This nuclease showed decreased or similar off-target activity, relative to eSpCas9 and SpCas9-HF across a number of different sites. Using a combined random and directed mutagenesis approach, the Anna Cereseto Lab created pX-evoCas9, which they showed also increased specificity.

Most recently, the Jungjoon Lee Lab used directed evolution to generate Sniper-Cas9 resulting in a Cas9 with improved specificity. They compared on-site and off-site activity of Sniper-Cas9 to different Cas9 variants using gRNAs of different lengths (Lee et al., 2018) and found that Sniper-Cas9 showed a high degree of both specificity and sensitivity across a large number of loci and gRNA lengths.

Recent alternatives to Cas9

While Cas9 remains the most widely used CRISPR nuclease, recent work has shown that alternative enzymes may better suit specific experimental approaches. One example is CasX, characterized by the Oakes and Doudna Labs, which is both small and potentially less likely to elicit a strong immune response, factors that may make it especially suitable for *in vivo* research and potentially treatments. Cas12b, characterized by the Zhang (Strecker et al., 2019) and Li (Teng et al., 2018) Labs also shows promise for *in vivo* research due to its small size and specificity relative to Cas9. As scientists continue to characterize CRISPR nucleases from a variety of sources, it is likely that the toolkit available to researchers will continue to expand.

References

Hu, Johnny H., et al. "Evolved Cas9 variants with broad PAM compatibility and high DNA specificity." Nature 556.7699 (2018): 57. PubMed PMID: 29512652. PubMed Central PMCID: PMC5951633.

Choosing a CRISPR Nuclease: Site Accessibility, Specificity, and Sensitivity (CONT'D)

Lee, Jungjoon K., et al. "Directed evolution of CRISPR-Cas9 to increase its specificity." Nature communications 9.1 (2018): 3048. PubMed PMID: 30082838. PubMed Central PMCID: PMC6078992.

Nishimasu, Hiroshi, et al. "Engineered CRISPR-Cas9 nuclease with expanded targeting space." Science 361.6408 (2018): 1259-1262. PubMed PMID: 30166441. PubMed Central PMCID: PMC6368452.

Strecker, Jonathan, et al. "Engineering of CRISPR-Cas12b for human genome editing." Nature communications 10.1 (2019): 212. PubMed PMID: 30670702. PubMed Central PMCID: PMC6342934.

Teng, Fei, et al. "Repurposing CRISPR-Cas12b for mammalian genome engineering." Cell discovery 4.1 (2018): 63. PubMed PMID: 30510770. PubMed Central PMCID: PMC6255809.

xCas9: Engineering a CRISPR Variant with PAM Flexibility

By Mary Gearing | Mar 28, 2018

In order to bind DNA, Cas9 and other CRISPR enzymes require a short PAM sequence adjacent to the targeted sequence at the locus of interest. SpCas9's 3' NGG PAM occurs frequently in GC-rich genomes, but a PAM is not always available near the locus you'd like to modify. To tackle the PAM problem, researchers have engineered alternative Cas9s binding distinct PAM sequences. Now, David Liu's lab has gone one step further, using directed evolution to create xCas9, an enzyme recognizing a broad range of PAMs like NG, GAA, and GAT, but also displaying increased editing specificity. We're excited to learn more about xCas9 - here's what we know so far!

Why do we need PAM flexibility?

IDT estimates that there is an NGG PAM every 42 bases in the human genome. SpCas9 targeting should be easy, right? Yes and no. If you're just trying to create a knockout, you might be happy targeting anywhere in an exon. But things get more complicated if you'd like to use homology directed repair to make precise edits or knock-ins. Editing efficiency decreases drastically once your cut site is more than 10 bp away from the insertion site, so HDR will be difficult if there's no PAM near your insertion site. Base editing, another method to create point mutations, also has a strict editing window about 15 nt upstream of the PAM. Even if there's a PAM near your insertion site, the corresponding gRNA might be undesirable due to potential off-target sites or low predicted on-target efficiency.

In cases where Cas9 won't work, there are other options. Commonly used Cpf1 variants, with a 5' TTTV PAM, improve targeting in AT-rich loci or genomes. Cas9s from other species, like *S. aureus* and *N. meningitidis* require different PAM sites than SpCas9, in this case, NNGRRT and NNNNGATT, respectively. Researchers also engineered SpCas9 variants with alternative PAMs like NGAG and NGCG and SaCas9 with an NNGRRN PAM by mutagenizing the PAM-interacting domain. However, these alternative PAM sequences are longer and often more complex than NGG, resulting in lower overall genome frequency. Despite the progress we've made, there is still a need for and interest in enzymes with shorter PAM sequences and/or broader PAM flexibility.

Creating xCas9 through directed evolution

Hu et al. used phage-assisted evolution to generate and select for Cas9 mutants with PAM flexibility. In this system, the mutagenesis plasmid (MP6) permits quick, large-scale mutation of the gene of interest. They created a selection phage (SP) to carry catalytically dead SpCas9 (dCas9) fused to the bacterial polymerase subunit ω. This construct is designed to activate transcription from the accessory plasmid (AP) encoding phage gene III. The accessory plasmid library contained all 64 possible NNN PAM sequences adjacent to a protospacer. If ω-dCas9 could recognize a library PAM, it would induce transcription of gene III, permitting phage propagation and keeping that dCas9 sequence in the pool. According to this logic, ω-dCas9 variants recognizing multiple PAMs should have a fitness advantage

xCas9: Engineering a CRISPR Variant with PAM Flexibility (CONT'D)

over proteins with single PAM recognition.

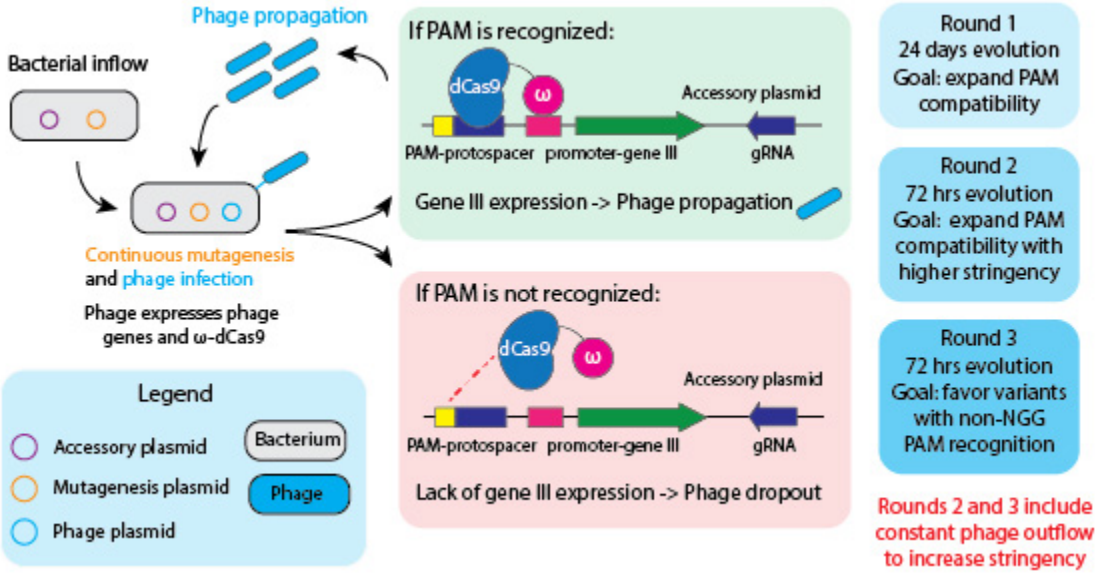


Figure 1: Steps in the directed evolution of xCas9.

After 24 days of phage propagation, Hu et al. isolated xCas9 1.0-1.4 and subjected them to another 72 hours of phage-assisted evolution, generating xCas9 2.0-2.6. This second round of evolution included continuous outflow of phage, increasing stringency and promoting survival of only the broadest-targeting variants. For the final round of evolution, they used a no-G PAM library to enrich for Cas9 variants with activity on non-NGG PAMs, generating xCas9 3.0-3.13.

To test xCas9 3.0-3.13, Hu et al. first restored catalytic residues D10 and H840. To ensure that the clones were capable of cleavage, they tested the clones against an NNN PAM library in bacteria where cleavage leads to the loss of spectinomycin resistance. xCas9 3.7 showed the broadest PAM range, recognizing NG, NNG, GAT, and CAA PAMs, while xCas9 3.6 displayed the second best PAM range.

Characterizing xCas9 in mammalian cells

Hu et al. compared xCas9 activity with wt SpCas9 in mammalian cells, finding xCas9 3.7 displayed about the same editing rate as SpCas9 at an NGG PAM. When they tested a variety of NG, GAT, GAA, and CAA PAMs, they saw that xCas9 3.7 cleaves multiple PAMs at a higher efficiency than SpCas9: NGT (4.5-fold efficiency), NGC (2.1-fold), NGA (1.6-fold) and GAA and GAT (5.2-fold). The low increase in NGA editing efficiency fits the previous observation that NGA is a secondary PAM for SpCas9.

Hu et al. performed GUIDE-seq analysis to characterize genome-wide off-targets, since increasing

xCas9: Engineering a CRISPR Variant with PAM Flexibility (CONT'D)

PAM flexibility might also increase off-target activity. Surprisingly, off-target effects across 5 gRNAs with NGG PAM sequences were actually lower for xCas9 3.6/3.7 than for SpCas9. For some non-promiscuous gRNAs, xCas9 3.7 displayed more than a 100 fold decrease in off-target editing; for known promiscuous gRNAs, xCas9 3.7 still showed a 4.2-9.4-fold improvement in off-target:on-target editing ratio. Importantly, the observed off-target loci displayed a broad range of PAM sequences, indicating they may be less predictable than SpCas9 off-targets.

Employing xCas9 for transcriptional regulation and base editing

Hu et al. also engineered transcriptional activators using the xCas9 scaffold. With an NGG PAM, xCas9 3.7-VPR constructs modestly outperformed SpCas9-VPR in activating transcription. Sites containing NG, GAA, or GAT PAMs averaged 56-91% transcriptional activation compared to the NGG PAM, confirming the flexibility of xCas9 3.7.

Base editor variants with xCas9 are similarly flexible. Like the xCas9 3.7 nuclease, xCas9 3.7 fused to third generation cytidine base editor architecture (BE3) slightly improves NGG PAM editing efficiency over that of SpCas9-BE3 (37 vs. 28%). Cytidine base editing at NGT, NGC, GAA, and GAT PAMs is improved, with a smaller increase in NGA editing. Adenine base editor xCas9 3.7-ABE also follows the same pattern, with improved editing at an NGG PAM (69 vs 48%) compared to SpCas9-containing ABE7.10, as well as improvements at NG and GAT PAMs.

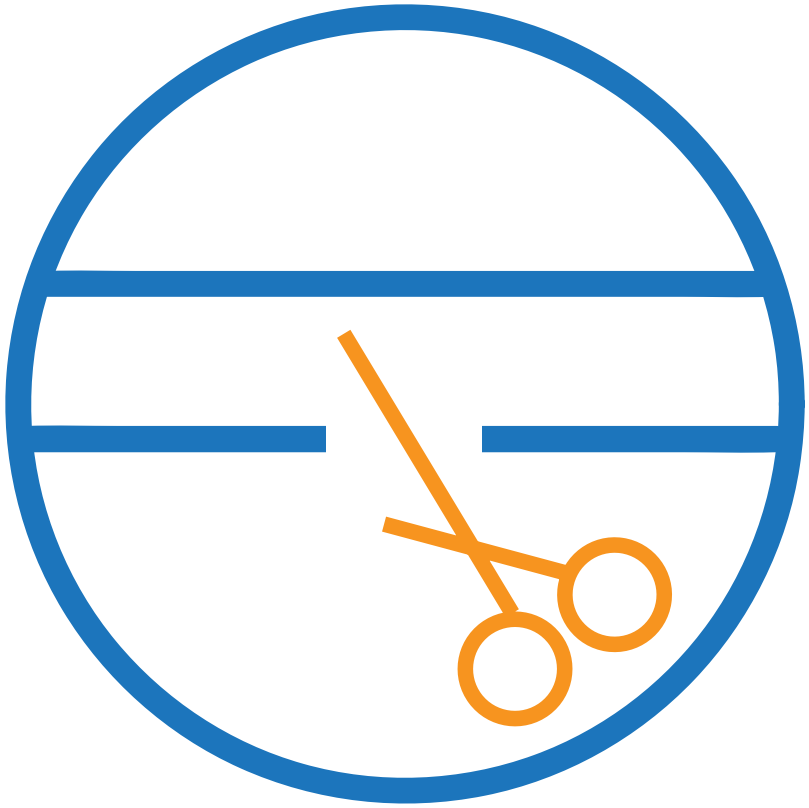
The development of xCas9 3.6 and 3.7 shows that we have only scratched the surface of modifications that can be made to Cas9. It is a pleasant surprise that selecting for one beneficial trait, PAM compatibility, also produced an enzyme with improved on-target and reduced off-target editing rates. This PAM flexibility should expand applications that require very specific targeting, like homology directed repair and base editing, to more loci.

References

Hu, Johnny H., et al. "Evolved Cas9 variants with broad PAM compatibility and high DNA specificity." Nature (2018) doi: 10.1038/nature26155 PubMed PMID: 29512652

Chapter 3

Cas Functions: Nickase



Cas9 Nickase Design and Homology Directed Repair

By Mart Gearing | Mar 15, 2018, Updated Aug 13, 2020 by Leah Schwiesow

By mutating one of two Cas9 nuclease domains, researchers created the CRISPR nickase. Nickases create a single-strand rather than a double-strand break, and when used with two adjacent gRNAs, can lower the probability of off-target editing. In this article, we'll summarize how IDT (Integrated DNA Technologies) first demonstrated how CRISPR nickases improve homology directed repair rates, and share their design rules for your next CRISPR nickase experiment.

Overview of Cas9 nickase

We'll use SpCas9 nickases as examples for this article. The D10A mutation inactivates the RuvC domain, so this nickase cleaves only the target strand. Conversely, the H840A mutation in the HNH domain creates a non-target strand-cleaving nickase. Instead of cutting both strands bluntly with WT Cas9 and one gRNA, you can create a staggered cut using a Cas9 nickase and two gRNAs.

For nickase applications, a common question is: how should the gRNAs be oriented in comparison to each other? The gRNAs must target different strands to create a DSB, but this can be accomplished with either a PAM-in or PAM-out orientation. Just as the names imply, PAM-out designs have the PAM sequences on the extremes of the targeted region, whereas PAM-in designs place the PAMs closer together in the middle of the targeted region.

To identify optimal nickase designs, IDT scientists Mollie Schubert and Shuqi Yan designed experiments to test nickase preferences using HEK293 cells. They started by comparing D10A and H840A total editing efficiency in PAM-in and PAM-out configurations with varied nick distances (40 - 130 nt). From these results, it's clear that editing is much higher in the PAM-out configuration. In addition, D10A editing efficiency is consistently higher than that of H840A, especially with smaller nick distances. Using D10A, they subsequently showed that editing efficiency is very low when the gRNAs are too close (7-23 nt nick distance).

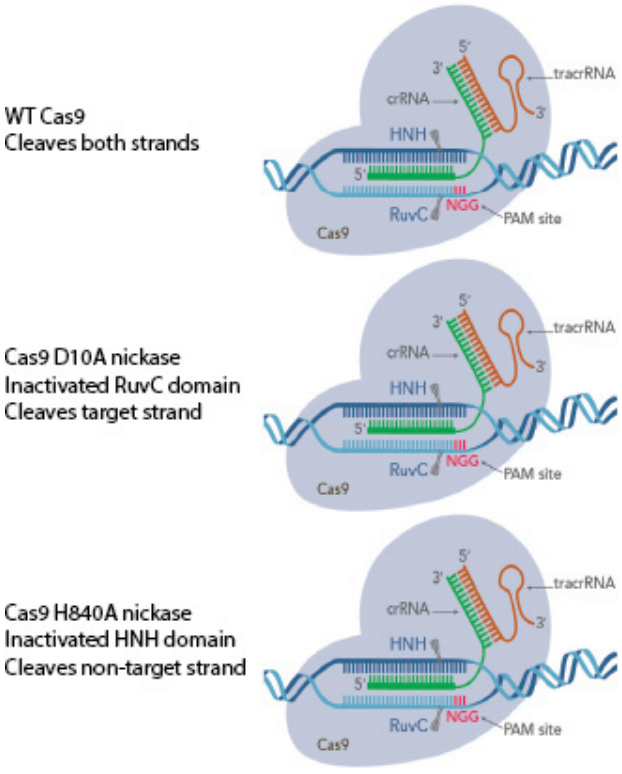


Figure 1: While wild-type Cas9 cleaves both strands, the two Cas9 nickases cleave one strand, either the target strand or the non-target strand.

Both D10A and H840A are potent editors, but their indel profiles vary. In the PAM-out configuration, D10A tends to generate small deletions, while H840A is biased towards large insertions. These profiles are likely due to the distinct overhang patterns generated by the nickases; D10A creates 5' overhangs and H840A

Cas9 Nickase Design and Homology Directed Repair (CONT'D)

creates 3' overhangs in a PAM-out design.

Exploring nickases for homology directed repair

The potential benefit of using nickases for HDR is targeting range: using an individual gRNA with WT Cas9, repair levels decrease rapidly 10 bp from the cut site. So if you can't find a good gRNA that cuts close to your insertion site, you can't obtain high HDR efficiency. Nickases create a staggered cut - could this system mediate repair throughout the entire region between the nicks? To find out, Schubert and Yan designed a PAM-out nickase experiment to insert an EcoRI site using a single-stranded oligonucleotide (ssODN) donor. Using a D10A nickase to insert an EcoRI site in various spots between the nick sites, they observed >20% repair (max 27%) across the 51 nt region using either a top or bottom strand ssODN donor with 40 bp homology arms. The HDR efficiency compared favorably to using WT Cas9 with the left gRNA across the window. In a subsequent experiment, they were also able to introduce a longer insertion (mCherry) using IDT Megamer long ssDNA with 100 nt homology arms.

Schubert and Yan next examined a situation thought to be suboptimal for HDR. At the AAVS1 locus, the two nearest gRNAs had cleavage sites 12 nt and 13 nt away from the target site rather than the optimal <10 nt distance. Employing two gRNA sites that are further away from the target site, they designed a D10A nickase strategy (46 nt nick distance, PAM-out) to see if they could solve this problem. D10A nickase outperformed both WT Cas9 sites tested by a wide margin, exhibiting >20% repair efficiency. Clearly, a nickase strategy can expand the targetable region, which is especially useful if there are no available guides close to the desired mutation site.

Quick tips for nickase design

- Use a PAM-out configuration
- Optimize your spacing
  - D10A: nick sites separated by 37-68 bp
  - H840A: nick sites separated by 51-68 bp
- Use D10A for HDR
- Place intended insert in between nick sites
  - Use 40 nt homology arms for small insertions or tags (ssODN; IDT Ultramer® Oligonucleotides)
  - Use 100 nt homology arms for large insertions (long ssDNA; IDT Megamer® ssDNA Fragments)
  - Test both bottom and top strand ssDNA donors (if possible)

The text and images in this post were adapted from an IDT webinar: Optimized methods to use Cas9 nickases in genome editing. We would like to thank IDT scientists Mollie Schubert and Shuqi Yan for reviewing this article.

Cpf1: A New Tool for CRISPR Genome Editing

By Mary Gearing | Oct 14, 2015, Updated Aug 13, 2020 by Leah Schwiesow

Note: Cpf1 is also called Cas12a.

In 2015, Zetsche et al. added to the CRISPR toolbox with their characterization of two Cpf1 orthologs that display cleavage activity in mammalian cells. Like Cas9 nucleases, Cpf1 family members contain a RuvC-like endonuclease domain, but they lack Cas9's second HNH endonuclease domain. Cpf1 cleaves DNA in a staggered pattern and requires only one RNA rather than the two (tracrRNA and crRNA) needed by Cas9 for cleavage. In certain cases, Cpf1 may be better suited for genome editing than Cas9 - read on to learn more about Cpf1 and check out our CRISPR guide for a refresher on CRISPR/Cas9.

How was Cpf1 found and tested?

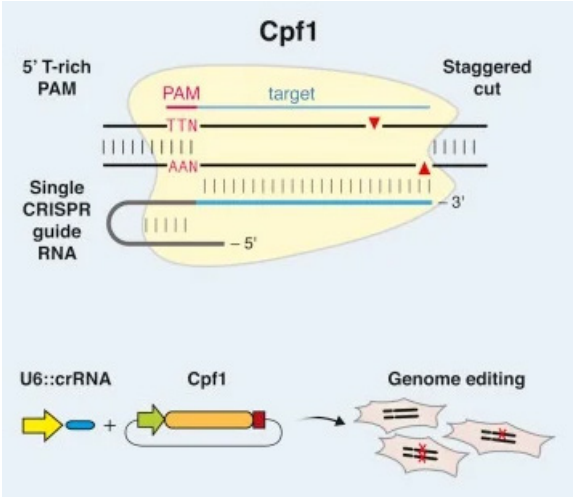


Figure 1: Cpf1 creates a staggered cut.

Class 2 CRISPR systems, including the type II Cas9-based system, require a single-component nuclease to mediate cleavage rather than the multi-subunit complex employed by class 1 systems. A putative new class 2 nuclease, Cpf1 (CRISPR from *Prevotella* and *Francisella*), was annotated in several genomes and is classified as a type V CRISPR system. Like Cas9, Cpf1 contains a RuvC-like endonuclease domain, but it lacks Cas9's other HNH endonuclease domain, indicating that Cpf1 functions differently. Since Cpf1 loci are widely distributed across bacterial species, Zetsche et al. hypothesized that Cpf1 might represent a functional CRISPR nuclease that could be adapted for genome editing. The use of a different nuclease could potentially overcome some of Cas9's shortcomings - namely its blunt double stranded cleavage and G-rich PAM requirement.

Zetsche et al. started from square one to characterize the Cpf1 nucleases. Using *Francisella* Cpf1 (FnCpf1), they employed an *E. coli* plasmid depletion assay to discover FnCpf1's PAM sequence requirements. Cpf1's preferred PAM is 5'-TTN, differing from that of Cas9 (3'-NGG) in both genomic location and GC-content. After sequencing and searching for cellular RNAs important for Cpf1 function, they found that mature crRNAs for Cpf1-mediated cleavage are 42-44 nucleotides in length, about the same size as Cas9's, but with the direct repeat preceding the spacer rather than following it. The Cpf1 crRNA is also much simpler in structure than Cas9's; only a short stem-loop structure in the direct repeat region is necessary for cleavage of a target. Cpf1 also does not require an additional tracrRNA.

Once they had determined the minimal elements of CRISPR-Cpf1, Zetsche et al. turned to characterizing its cleavage pattern. Again, they were in for a surprise! Whereas Cas9 generates blunt ends 3 nt upstream of the PAM site, Cpf1 cleaves in a staggered fashion, creating a 5 nucleotide 5'

Cpf1: A New Tool for CRISPR Genome Editing (CONT'D)

overhang 18-23 bases away from the PAM. With this information, they turned to a cell culture system to see if any Cpf1 nucleases would exhibit *in vivo* activity in mammalian cells. From 16 diverse Cpf1 candidates, Zetsche et al. found two that display robust cleavage activity similar to that of Cas9. These two nucleases, AsCpf1 and LbCpf1 (1307 and 1228 amino acids long respectively), both cleave in a staggered pattern similar to FnCpf1.

Potential advantages of Cpf1 over Cas9

Type II CRISPR systems based on Cas9 were thought to be the simplest CRISPR systems and the easiest to adapt to genome editing, but the introduction of type V Cpf1-driven systems has added another option to the CRISPR toolbox. Cpf1's staggered cleavage pattern opens up the possibility of directional gene transfer, analogous to traditional restriction enzyme cloning. Sticky-end mediated gene transfer would be particularly helpful for targeting non-dividing cells, which are difficult to modify through homology-directed repair (HDR). Cpf1 also expands the number of sites that can be targeted by CRISPR to AT-rich regions or AT-rich genomes that lack the 3'-NGG PAM sites favored by SpCas9.

Since Cpf1 doesn't require a tracrRNA, crRNA guides are only ~42 nt long. Direct synthesis of these crRNAs should be significantly cheaper than that of the ~100 nt crRNA/tracrRNA hybrid guides needed for Cas9 function. Since both Cpf1 and its guide RNAs are smaller than their SpCas9 counterparts, they will also be easier to deliver in low-capacity vectors, such as adeno-associated viral (AAV) vectors. In 2017, Zetsche et al. developed a Cpf1 multiplexing approach using a single crRNA array to express up to 4 crRNAs.

Zetsche et al. also suggest that Cpf1 may improve the frequency of HDR over non-homologous end joining (NHEJ). Cas9-mediated NHEJ usually destroys the PAM site due to its proximity to the cleavage site, preventing future edits. In contrast, since Cpf1 cleaves relatively far away from the PAM, NHEJ might retain the PAM site. Therefore, if HDR does not initially occur after Cpf1-mediated cleavage, the continued presence of the PAM may give Cpf1 the ability to cleave again and possibly mediate HDR. This "second chance" mechanism might improve the frequency of desired HDR edits, but the possibility has not yet been experimentally confirmed. To prevent new editing post-HDR, repair templates should remove the PAM sequence.

Comparing Cpf1 and Cas9 on-target and off-target efficiency

When Cpf1 was first identified, we didn't know much about its on-target and off-target editing efficiency. Kim et al. and Kleinstiver et al. characterized genome-wide editing efficiency of two Cpf1 orthologs known to be active in mammalian cells, LbCpf1 and AsCpf1. In both reports, on target editing efficiency for the Cpf1 orthologs was only slightly lower than that of the widely used SpCas9 and comparable to

Cpf1: A New Tool for CRISPR Genome Editing (CONT'D)

SaCas9. As seen with Cas9 orthologs, Cpf1 efficiency varies widely with gRNA sequence.

Both groups used multiple methods to examine Cpf1 off-target editing. First, they designed gRNAs with single and double mismatches throughout the 23-base sequence. Double mismatches ablated Cpf1 activity, except when they were present in the 3' end of the target sequence (bases 19-23). Cpf1 is also sensitive to single mismatches, but variably so, with Kleinstiver et al. reporting that Cpf1 can tolerate mismatches at gRNA positions 1, 8, 9, and 19-23. Accordingly, the 3' end of the gRNA target sequence does not have an essential function in Cpf1-mediated editing, as Kleinstiver et al. saw no decrease in Cpf1 activity with 4-6 base deletions at the 3' end of the target sequence.

Cpf1's strength may lie in its low off-target editing rates, determined using sophisticated genome-wide analysis. At many of its computationally predicted off-target sites, Cpf1 does not mediate detectable off-target cleavage. Most gRNAs directed low-frequency Cpf1 cleavage at 1-12 off-target sites; in contrast, SpCas9 may cleave at ~90 sites, according to Kim et al. Kim et al. also compared the ratio of total off-target to on-target modification for AsCpf1 and LbCpf1, and found that both orthologs show lower off-target activity than that previously observed with SpCas9. Kleinstiver et al. suggest that AsCpf1's off-target rate is similar to that of high fidelity Cas9s eSpCas9 and SpCas9-HF1. Both AsCpf1 and LbCpf1 ribonucleoproteins (RNPs) failed to induce off-target editing in a cell culture model.

The application of Cpf1 to genome editing is exciting both in terms of basic science and translational applications. This discovery of this type V CRISPR system proved we had a lot more to learn about CRISPR biology. Later work, like the adaptation of Cas13 to RNA targeting and RNA editing, has further shown the diversity of CRISPR-based systems.

References

Zetsche, Bernd, et al. "Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system." Cell (2015). PubMed PMID: 26422227.

Zetsche, Bernd, et al. "Multiplex Gene Editing by CRISPR–Cpf1 Using a Single crRNA Array." Nat Biotechnol. 35.1 (2016): 31-34. PubMed PMID: 27918548.

Kim, Daesik, et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. Nat Biotechnol. (2016). PubMed PMID: 27272384

Kleinstiver, Benjamin P., et al. "Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells." Nat Biotechnol. (2016). PubMed PMID: 27347757.

Makarova, Kira S., et al. "An updated evolutionary classification of CRISPR-Cas systems." Nat Rev Microbiol. (2015). PubMed PMID: 26411297.

Targeted Mutagenesis with EvolvR

By Beth Kenkel | Feb 21, 2019

Mutagenesis is a tool that both evolution and molecular biologists use to tinker with DNA. Making changes to a DNA sequence can help scientists identify and/or facilitate the evolution of new phenotypes, and forward genetics harnesses this at a large scale by screening diverse libraries of genetic variants. Several methods for generating mutant libraries exist, but none provide a means to continuously diversify all nucleotides within a user-defined genomic region. EvolvR, a CRISPR-Cas9 based targeted mutagenesis method developed by the Dueber Lab at Berkeley, provides a new approach for generating novel genetic variants in bacteria. Read on to learn about the key components of EvolvR and its potential applications.

Challenges of existing mutagenesis methods

Several continuous diversification methods exist, but there are a few drawbacks with these approaches that makes it challenging for users to achieve targeted diversification of all nucleotides within a defined region of a genome.

- Require stringent bacterial growth conditions. Expression of an error prone polymerase I (PolI3M) is capable of mutagenesis in *E. coli*, but this approach requires special bacterial growth conditions in order to maximize the mutation rate. For example the Phage Assisted Continuous Evolution (PACE) method requires constant turbidity of the culture, so bacteria must be grown in a turbidostat.
- Only mutate particular types of nucleotides. Both the Bassik Lab and the Chang Lab have created methods that introduce diversification by targeting a site of interest with dCas9 fused to the Activation Induced Cytidine Deaminase (AID) enzyme. AID, however, only changes cytidines or guanines to the other three bases, which limits the level of variation that can be introduced.
- Diversification is limited due to the integration of oligonucleotide libraries of a discrete size at the target site. The use of such libraries means that the diversity introduced by lambda red recombineering and Multiplex Automated Genome Engineering (MAGE) will always be limited by the size of the library used.

EvolvR: key components

The EvolvR system has two key components: the error prone PolI3M that's fused to a nicking Cas9 (nCas9). Similar to mutagenesis with error prone PCR, EvolvR uses an error prone polymerase to introduce mutations. However, by using nCas9 to direct PolI3m to a particular genomic loci, EvolvR allows for targeted mutagenesis of a site of interest. A gRNA is used to direct the PolI3M-nCas9 complex to a DNA site of interest, which nCas9 nicks and then dissociates from. PolI3M then binds the nicked DNA, and extends it from the 3' end, while its native endonuclease activity degrades the displaced strand.

Targeted Mutagenesis with EvolvR (CONT'D)

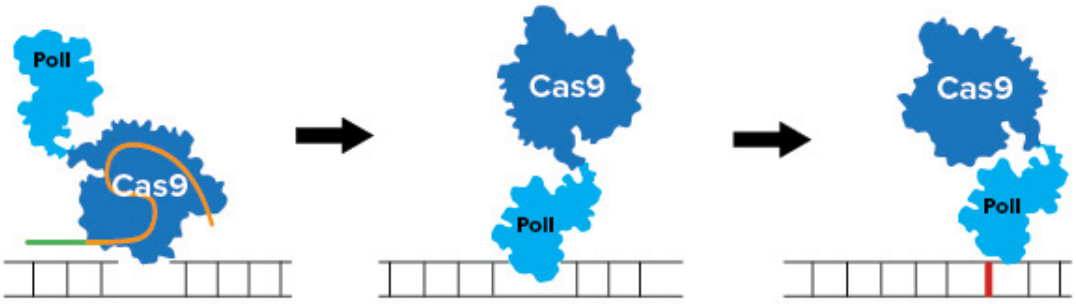


Figure 1: EvolvR consists of an error prone polymerase that's fused to a nicking Cas9. A gRNA directs Cas9 to target site for nicking. Cas9 dissociates and PolI binds and extends from the 3' end of the nick, occasionally introducing mutations.

The initial version of EvolvR has an editing window of ~17 nt from the nick site, which was expected since the processivity window of PolI is 15-20 bp. EvolvR's on-target mutagenesis rate was 2.5 x 10^{-6} mutations per nucleotide per generation vs. 10 x 10^{-10} mutations per nucleotide per generation of wild-type *E. coli*, while only increasing the standard mutation rate seen during DNA replication by 120-fold over background.

Due to its modular nature, the Dueber Lab created a few versions of EvolvR. To increase the mutation rate of EvolvR, three additional amino acid changes were introduced to nCas9. These changes promote nCas9's dissociation from DNA after nicking and yielded an enhance nCas9 (enCas9) EvolvR with a targeted mutagenesis rate ~9-fold higher than the original nCas9, while increasing the standard mutation rate by only ~2-fold over nCas9 levels. Two sets of modifications were also made to PolI3M to increase the mutation rate and extend the editing window of EvolvR: (1) two additional mutations to PolI3M generated PolI5M which increased EvolvR's mutagenesis rate to ~10^{-3} mutations per nucleotide per generation, and (2) the addition of the thioredoxin-binding domain (TBD) from bacteriophage T7 DNA polymerase increased the processivity of PolI3M, and thereby increasing the editing window of EvolvR to ~56 bp. A larger editing window also increased the likelihood that a single gRNA would introduce more than one mutation near the target site.

Applications of EvolvR

The Dueber Lab used EvolvR to discover novel genotypes that result in a spectinomycin resistant phenotype in bacteria. While there are many known spectinomycin resistance mutations, several new mutations were identified by targeting EvolvR to five dispersed regions of the target gene of spectinomycin, *rpsE*. These results not only uncovered novel resistance genotypes, but also provide a

Targeted Mutagenesis with EvolvR (CONT'D)

better understanding of the key residues responsible for this protein-drug interaction.

The Dueber Lab propose that EvolvR could also be used to map protein-protein interactions, investigate non-coding segments of the genome, or engineer microbes to perform tasks.

EvolvR could also be a tool for lineage tracing cells that do not tolerate double-strand DNA breaks by introducing a unique tag into a genomic site of choice via the mutational abilities of PolI3m. This tag would allow researchers to trace the progeny of a single starting cell. Additionally, multiplexing of guides allows EvolvR to target more than one genomic site for mutagenesis, a feature which they used to generate bacteria resistant to both spectinomycin and streptomycin, but could also be used for studying epistatic interactions.

Key takeaways for editing with EvolvR

EvolvR allows for the continuous diversification of all nucleotides at a user-defined loci in bacteria. Its modular nature provides versatility in defining the mutational rate and editing window of the system and the flexibility to address many different types of research questions. Are you ready to start EvolvR-ing your targeted mutagenesis experiments? Find EvolvR plasmids at Addgene!

References

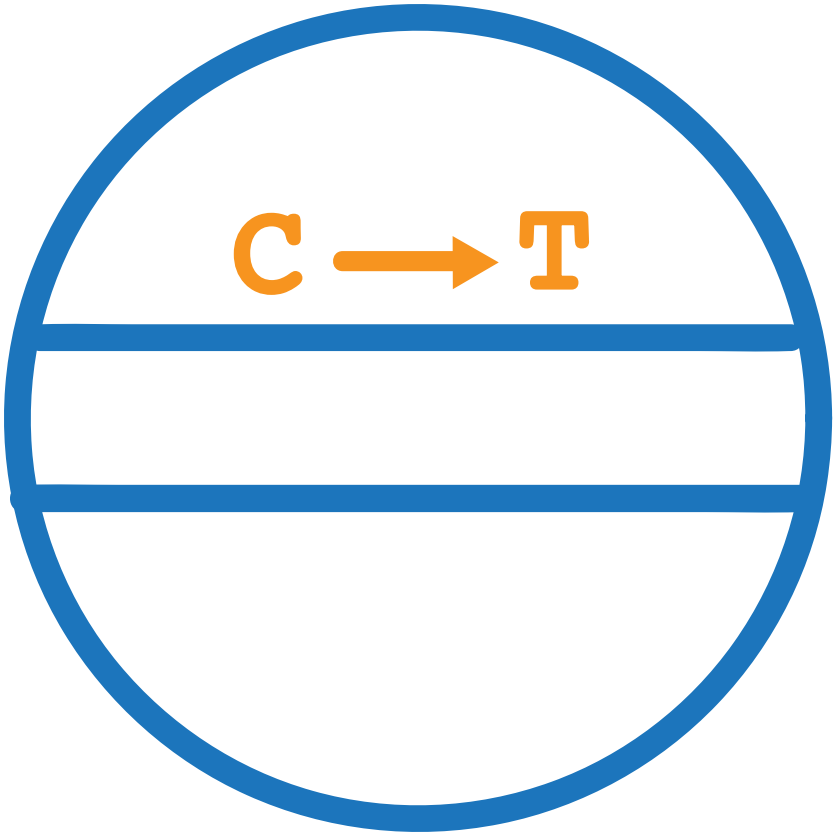
Gaj, Thomas, and Pablo Perez-Pinera. "The continuously evolving CRISPR barcoding toolbox." Genome biology 19.1 (2018): 143. PubMed PMID: 30253789. PubMed Central PMCID: PMC6154929.

Halperin, Shakked O., et al. "CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window." Nature 560.7717 (2018): 248. PubMed PMID: 30069054.

Sadanand, Saheli. "EvolvR-ing to targeted mutagenesis." Nature biotechnology 36.9 (2018): 819.

Chapter 4

Base Edit



Cytosine and Adenine Base Editors

By Mary Gearing | Aug 16, 2016, Updated Aug 6, 2020 by Jennifer Tsang

When we talk about CRISPR applications, one negative often comes up: the low editing efficiency of homology-directed repair (HDR). Compared to non-homologous end joining, HDR occurs at a relatively low frequency, and in nondividing cells, this pathway is further downregulated. Rather than try to improve HDR, scientists have developed two classes of base editors: cytosine base editors (CBEs) and adenine base editors (ABEs).

(There are also RNA base editors, but we'll just be covering DNA base editors here. Learn more about RNA base editors in this section of the eBook.)

What bases can base editors edit?

CBEs mediate a C to T change (or a G to A change on the opposite strand). ABEs make an A to G change (or a T to C change on the opposite strand). This only accounts for four of 12 possible changes.

More recently, the development of prime editing, which uses a different mechanism than CBEs or ABEs, from David Liu's lab has allowed scientists to make all 12 possible base-to-base changes.

How do base editors work?

Base editing requires three elements. Broadly:

1. A Cas nickase or Cas fused to a deaminase that makes the edit
2. A gRNA targeting Cas to a specific locus
3. A target base for editing within the editing window specified by the Cas protein

These elements were the starting point towards the development of the first cytosine base editors from postdocs Alexis Komor and the first adenine base editors from Nicole Gaudelli in David Liu's lab in 2016.

Cytosine base editing

The beginnings of cytosine base editing

Komor created the first cytosine base editor by coupling a cytidine deaminase with the inactive dCas9 (Komor et al., 2016). These fusions convert cytosine to uracil without cutting DNA. Uracil is then subsequently converted to thymine through DNA replication or repair. Fusing an inhibitor of uracil DNA glycosylase (UGI) to dCas9 prevents base excision repair which changes the U back to a C mutation. To increase base editing efficiency, you'd need a way to force the cell to use the deaminated DNA strand

Cytosine and Adenine Base Editors (CONT'D)

as a template. To do so, the lab used a Cas nickase, instead of dCas9. The resulting editor, BE3, nicks the unmodified DNA strand so that it appears "newly synthesized" to the cell. Thus, the cell repairs the DNA using the U-containing strand as a template, copying the base edit.

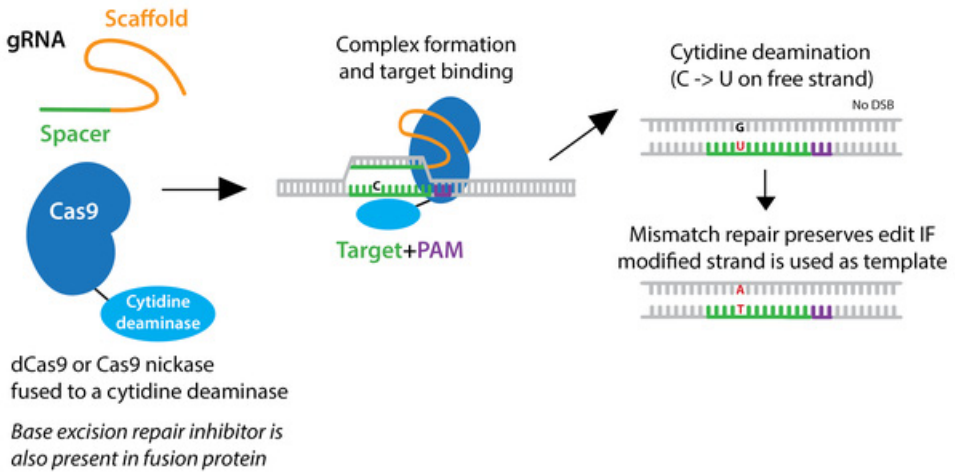


Figure 1: dCas9 or Cas9 nickase fused to a cytidine deaminase binds the target and deaminates the cytidine. Mismatch repair creates the edit if the modified strand is used as a template.

The BE3 system increased editing frequency to above 30% for a variety of targets in human cell lines, with an average indel frequency of only 1.1%. These numbers are a vast improvement over Cas9-mediated HDR for the loci tested where average HDR-mediated editing frequency was only 0.5%, and more indels were observed than point modifications. CRISPR base editing persists through multiple cell divisions, indicating that this method produces stable edits. However, this system is also subject to off-target editing based on Cas9 off-target activity.

Improving cytosine base editing scope and efficiency

Since the development of BE3, many research groups have made improvements to base editors including:

- expanding targeting scope
- improving editing efficiency
- decreasing off-target effects

In 2016, Akihiko Kondo's lab created the Target-AID base editor using a cytidine deaminase from sea lamprey fused to Cas9 nickase (Nishida et al., 2016). Target-AID acts similarly but not identically to BE3, modifying a 3-5 base window 18 bases upstream of the PAM.

Cytosine and Adenine Base Editors (CONT'D)

David Liu's lab generated BE3 variants with other deaminases: AID, CDA1, and APOBEC3G (Komor et al., 2017). CDA1-BE3 and AID-BE3 edited Cs following a G more efficiently than BE3, but APOBEC3G displayed less predictable sequence preferences.

The Liu lab also used natural and engineered Cas9 variants to develop five new base editors with distinct PAM sequences, expanding the number of available target sites for base editing (Kim et al., 2017). For each base editor, they observed editing activity with a minimum efficiency of ~50% and confirmed that the fusion protein retained the PAM properties of the individual Cas9. They also mutagenized the cytidine deaminase portion of the base editor to create SpCas9 base editors with editing windows as small as 1-2 nucleotides.

To reduce off-target effects associated with base editing, the lab Rees et al. created HF-BE3, a base editor containing high fidelity Cas9 variant HF-Cas9 (Rees et al., 2017). HF-BE3 showed 37-fold less off-target editing than BE3, with only a slight reduction in on-target editing efficiency. To further improve specificity, they purified HF-BE3 protein for delivery in ribonucleoprotein particles (RNPs) to both zebrafish embryos and the mouse inner ear.

Fourth-generation base editors

The fourth-generation base editors, BE4, reduce the undesired C->G or C->A conversions that can happen with BE3's. These byproducts likely resulted from excision by uracil N-glycosylase (UNG) during base excision repair. Adding a second copy of the UNG inhibitor, UGI, increases base editing product purity. The lab also extended the APOBEC1-Cas9n and Cas9n-UGI linkers to improve product purity, and these three improvements represent the fourth generation of base editors. Compared to BE3, BE4 offers a 2.3 fold decrease in C->G and C->A products as well as a 2.3 fold decrease in indel formation.

To further decrease indel formation 1.5-2 fold, the team fused bacteriophage protein Gam to the N-terminus of BE4 (Komor et al., 2017). Gam binds the free ends of DSBs, which may lead to cell death rather than NHEJ repair, thus removing these cells from the edited population.

Another way to improve base editing efficiency for mammalian edits is to ensure the editors make it into the nucleus and that they are expressed well. The Liu lab modified the nuclear localization signals and codon usage BE4 to create BE4max and AncBE4max with a 4.2-6-fold improvement in editing efficiency (Koblan et al., 2018).

Cytosine and Adenine Base Editors (CONT'D)

Adenine base editors

The beginning of adenine base editing

Nicole Gaudelli from David LIu's lab created an adenine base editor that would convert adenine to inosine, resulting in an A to G change (Gaudelli et al., 2017). Creating an adenine base editor requires an additional step because there are no known DNA adenine deaminases. They used directed evolution to create one from the RNA adenine deaminase TadA.

After seven rounds of molecular evolution, they obtained four adenine base editors (ABEs). ABE7.10 is the most active editor, displaying an average editing efficiency of 53% with an editing window of target positions 4-7. ABEs 6.3, 7.8, and 7.9 display slightly wider editing windows of position 4-9, although editing efficiency may be lower at positions 8 and 9. While cytosine base editors often produce a mixed population of edits, ABEs do not display significant A to non-G conversion at target loci. The removal of inosine from DNA is likely infrequent, thus preventing the induction of base excision repair.

In terms of off-target effects, ABEs also compare favorably to other methods. In a head to head comparison with Cas9, Cas9 modified 9/12 known off-targets with a 14% indel rate, while ABE7.10 modified only 4/12 off-targets with a frequency of 1.2%. Although the lab did not conduct comprehensive genome-wide studies of ABE specificity, their other experiments suggest that ABEs are robust but specific editors.

Improving adenine base editing

When the Liu lab created BE4max and AncBE4max mentioned above, they also made an adenine base editor with improved nuclear localization and expression. This base editor was named ABEmax.

In 2020, two papers were published describing additional ABEs evolved from ABE7.10 that have improved base editor targeting flexibility and specificity. In the first, the Liu lab used phage-assisted evolution selection systems to generate ABE8e(TadA-8e V106W), which edits ~590-fold faster than the TadA from ABE7.10 without increasing off-target activity (Richter et al., 2020). This is important because ABEs have generally been slow meaning that the Cas9 domain often lets go of the DNA before the edit is made.

Using ABE7.10 as a starting point, Gaudelli evolved the base editor into 40 new ABE8 variants (Gaudelli et al., 2020). Compared to ABE7.10, ABE8s resulted in 1.5-fold more editing at protospacer positions A5-A7 and 3.2-fold more editing at positions A3-A4 and A8-A10 at NGG PAM and 4.2-fold higher editing efficiency at non-NGG PAM variants compared to ABE7.10. ABE8s have an improved base editing capacity, even at sites previously difficult to target. ABE8s can achieve 98-99% target

Cytosine and Adenine Base Editors (CONT'D)

modification in primary T cells making them a promising tool for cell therapy applications.

Dual base editors

What about combining the function of both editors into one? This has recently been done by fusing the adenine and cytosine editing components together.

Keith Joung's lab created a dual-deaminase editor called SPACE (synchronous programmable adenine and cytosine editor) by fusing miniABEmax-V82G and Target-AID to Cas9 (Grünewald et al., 2020). Another lab took a similar approach. Dali Li's lab's A&C-BEmax consists of a fusion of both cytidine and adenosine deaminases with a Cas9 nickase (Zhang et al., 2020). It has increased CBE activity and reduced RNA off-targeting activity compared to ABEmax.

References

Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR (2017) Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature 551:464–471 . https://doi.org/10.1038/nature24644

Gaudelli NM, Lam DK, Rees HA, Solá-Esteves NM, Barrera LA, Born DA, Edwards A, Gehrke JM, Lee S-J, Liquori AJ, Murray R, Packer MS, Rinaldi C, Slaymaker IM, Yen J, Young LE, Ciaramella G (2020) Directed evolution of adenine base editors with increased activity and therapeutic application. Nat Biotechnol 38:892–900 . https://doi.org/10.1038/s41587-020-0491-6

Grünewald J, Zhou R, Lareau CA, Garcia SP, Iyer S, Miller BR, Langner LM, Hsu JY, Aryee MJ, Joung JK (2020) A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing. Nat Biotechnol 38:861–864 . https://doi.org/10.1038/s41587-020-0535-y

Kim YB, Komor AC, Levy JM, Packer MS, Zhao KT, Liu DR (2017) Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nat Biotechnol 35:371–376 . https://doi.org/10.1038/nbt.3803

Koblan LW, Doman JL, Wilson C, Levy JM, Tay T, Newby GA, Maianti JP, Raguram A, Liu DR (2018) Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat Biotechnol 36:843–846 . https://doi.org/10.1038/nbt.4172

Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533:420–424 . https://doi.org/10.1038/

Cytosine and Adenine Base Editors (CONT'D)

nature17946

Komor AC, Zhao KT, Packer MS, Gaudelli NM, Waterbury AL, Koblan LW, Kim YB, Badran AH, Liu DR (2017) Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. Sci Adv 3:eaao4774 . https://doi.org/10.1126/sciadv.aao4774

Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M, Mochizuki M, Miyabe A, Araki M, Hara KY, Shimatani Z, Kondo A (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science 353:aaf8729–aaf8729 . https://doi.org/10.1126/science.aaf8729

Rees HA, Komor AC, Yeh W-H, Caetano-Lopes J, Warman M, Edge ASB, Liu DR (2017) Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. Nat Commun 8: . https://doi.org/10.1038/ncomms15790

Richter MF, Zhao KT, Eton E, Lapinaite A, Newby GA, Thuronyi BW, Wilson C, Koblan LW, Zeng J, Bauer DE, Doudna JA, Liu DR (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat Biotechnol 38:883–891 . https://doi.org/10.1038/s41587-020-0453-z

Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, Wyvekens N, Khayter C, Iafrate AJ, Le LP, Aryee MJ, Joung JK (2014) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol 33:187–197 . https://doi.org/10.1038/nbt.3117

Zhang X, Zhu B, Chen L, Xie L, Yu W, Wang Y, Li L, Yin S, Yang L, Hu H, Han H, Li Y, Wang L, Chen G, Ma X, Geng H, Huang W, Pang X, Yang Z, Wu Y, Siwko S, Kurita R, Nakamura Y, Yang L, Liu M, Li D (2020) Dual base editor catalyzes both cytosine and adenine base conversions in human cells. Nat Biotechnol 38:856–860 . https://doi.org/10.1038/s41587-020-0527-y

Prime Editing

By Jennifer Tsang | Oct 24, 2019, Updated Jun 5, 2020

There are over 75,000 pathogenic genetic variants that have been identified in humans and catalogued in the ClinVar database. Previously developed genome editing methods using nucleases and base editors have the potential to correct only a minority of those variants in most cell types. A new technique from David Liu's lab at the Broad Institute could add more precision and flexibility to the CRISPR editing world.

This new approach, published in Nature in 2019 is called prime editing. It's a "search-and-replace" genome editing technique that mediates targeted insertions, deletions, and all possible base-to-base conversions. And, it can combine different types of edits with one another. All of this is possible without double strand breaks (DSBs) or donor DNA templates. How does this work? First, an engineered prime editing guide RNA (pegRNA) that both specifies the target site and contains the desired edit(s) engages the prime editor protein. This primer editor protein consists of a Cas9 nickase fused to a reverse transcriptase. The Cas9 nickase part of the protein is guided to the DNA target site by the pegRNA. After nicking by Cas9, the reverse transcriptase domain uses the pegRNA to template reverse transcription of the desired edit, directly polymerizing DNA onto the nicked target DNA strand. The edited DNA strand replaces the original DNA strand, creating a heteroduplex containing one edited strand and one unedited strand. Lastly, the editor guides resolution of the heteroduplex to favor copying the edit onto the unedited strand, completing the process.

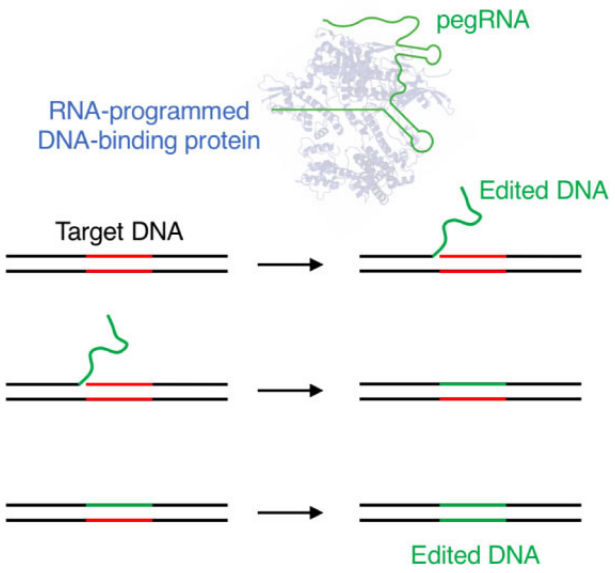


Figure 1: The prime editing guide RNA (pegRNA) specifies the target site and contains the edit. After Cas9 nicking, the reverse transcriptase domain uses pegRNA to reverse transcribe the edit. The edited strand replaces the original DNA strand. Resolution of the heteroduplex to favor copying the edit onto the unedited strand completes the process. Image from David Liu with pernission.

The prime editor: A fusion between Cas9 and reverse transcriptase

To decrease the components prime editing would introduce into the cell, the team fused the M-MLV reverse transcriptase (RT) with the Cas9 H840A nickase to create the prime editor (PE). They found that orientation matters: fusing the RT to the C-terminus of the Cas9 nickase resulted in higher editing efficiency. They called this complex PE1.

Building upon prior reverse transcriptase research, (Baranauskas et al., 2012; Arezi and Hogrefe, 2009), the Liu lab created and evaluated 19 PE1 variants with RT mutations known to increase activity, enhance binding between the template and primer binding site,

Prime Editing (CONT'D)

increase processivity, or improve thermostability. What came out on top? The Cas9 nickase fused to a pentamutant of M-MLV RT. They called this system PE2, which had prime editing efficiencies on average 2.3- to 5.1-fold (though up to 45-fold) higher across different genomic sites compared to PE1.

The pegRNA: A template and guide all in one

The other important component of prime editing is the prime editing guide RNA (pegRNA). The pegRNA is a guide RNA that also encodes the RT template, which includes the desired edit and homology to the genomic DNA locus. Sequence complementary to the nicked genomic DNA strand serves as a primer binding site (PBS). This PBS sequence hybridizes to the target site and serves as the point of initiation for reverse transcription.

To optimize pegRNAs, the team found that extending the pegRNA primer binding site to at least eight nucleotides enabled more efficient prime editing in HEK293T cells.

For designing pegRNAs, the Liu lab created PrimerDesign, a prime editing design tool, with the Pinello lab and the Joung lab.

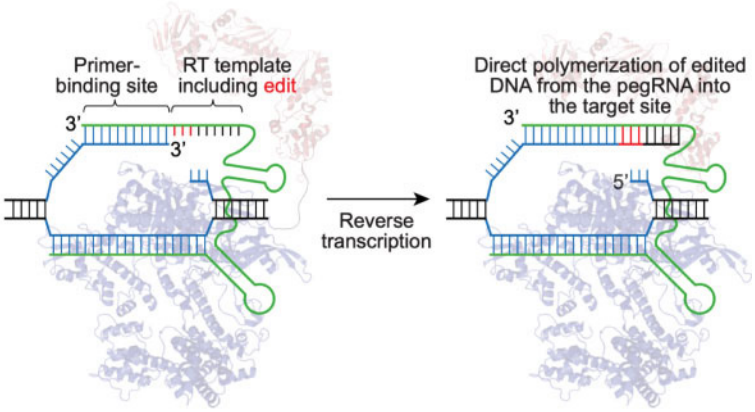


Figure 2: The prime editor with pegRNA. The Cas9 portion of the editor cuts the DNA and the reverse transcriptase portion polymerizes DNA onto the nicked strand based on the pegRNA sequence. Image from David Liu with permission.

Prime Editor 3 (PE3): Resolving mismatched DNA to favor the edit

Once the prime editor incorporates the edit into one strand, there's a mismatch between the original sequence on one strand and the edited sequence on the other strand. To guide heteroduplex resolution to favor the edit, the Liu lab turned to a strategy they previously used when they developed base editing (Komor, et al, 2016). By nicking the non-edited strand, they can cause the cell to remake that strand using the edited strand as the template.

A third prime editing system called PE3 does just this by including an additional sgRNA. Using this sgRNA, the prime editor nicks the unedited strand away from the initial nick site (to avoid creating a double strand break), increasing editing efficiencies 2-3 fold with indel frequencies between 1-10%.

Note: PE2, PE3, and PE3b all use PE2 (pCMV-PE2). PE3 and PE3b add an sgRNA, which can be any

Prime Editing (CONT'D)

guide RNA expression vector or cassette

Advantages of prime editing

Less constrained by PAM sequence location

The prime editor extends the reach of CRISPR genome editing as it can edit near or far from PAM sites making it less constrained by PAM availability like other methods. The PAM-to-edit distance can be over 30 base pairs for prime editing. Since PAM sites occur every ~8 base pairs on either DNA strand, many previously developed base editors (Table 1 from Rees and Liu, 2019) with a <8 base pair editing window cannot edit within what Fyodor Urnov refers to as "PAM deserts" in the genome.

More versatile and precise than base editing (in certain circumstances)

Base editors developed thus far can only create a subset of changes (C->T, G->A, A->G, and T->C). Prime editing allows for all 12 possible base-to-base changes.

Prime editing is also more precise. Base editors, for example, will edit all the C's or A's within the base editing window, while prime editors make a specific edit defined by the pegRNA. In cases when bystander editing is unacceptable, prime editors can be used to avoid this possibility.

However, there are instances where traditional base editors are preferred. For instance, if target nucleotides are positioned within the canonical base editing window, base editing has higher efficiency and fewer indels than prime editing. But for positions that aren't well positioned within the editing window, prime editing is more efficient due to its lower dependence on PAM placement.

Fewer byproducts and more efficient than homology directed repair

Homology directed repair (HDR) stimulated by double strand breaks has been widely used to generate precise changes. However, the efficiency of Cas9 cleavage is relatively high while the efficiency of HDR is relatively low, meaning that most Cas9-induced DSBs are repaired by non-homologous end joining. As a result, Cas9 treatment causes most products to be indels while the efficiency of HDR is typically less than 10%. In contrast, prime editing can offer ~20-50% efficiency in HEK293T cells with 1-10% indels. In other tested cell types, including post-mitotic primary mouse cortical neurons, the authors report lower prime editing efficiencies, but still see much higher ratios of desired edits to indel byproducts than Cas9-initiated HDR.

Prime Editing (CONT'D)

What's next for prime editing?

While prime editing is an exciting step towards more versatile genome editing, it's new at this point and warrants many additional studies. In their paper, the Liu lab points out the need to investigate off-target prime editing in a genome-wide manner, identify any inadvertent effects the prime editors may have on the cells, and assess *in vitro* and *in vivo* delivery strategies.

References

Anzalone, Andrew V., et al. "Search-and-replace genome editing without double-strand breaks or donor DNA." Nature (2019): 1-1. PubMed PMID: 31634902.

Arezi, Bahram, and Holly Hogrefe. "Novel mutations in Moloney Murine Leukemia Virus reverse transcriptase increase thermostability through tighter binding to template-primer." Nucleic acids research 37.2 (2008): 473-481. PubMed PMID: 19056821. PubMed Central PMCID: PMC2632894.

Baranauskas, Aurimas, et al. "Generation and characterization of new highly thermostable and processive M-MuLV reverse transcriptase variants." Protein Engineering, Design & Selection 25.10 (2012): 657-668. PubMed PMID: 22691702.

Komor, Alexis C., et al. "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage." Nature 533.7603 (2016): 420. PubMed PMID: 27096365. PubMed Central PMCID: PMC4873371.

Rees, Holly A., and David R. Liu. "Base editing: precision chemistry on the genome and transcriptome of living cells." Nature reviews genetics 19.12 (2018): 770-788. PubMed PMID: 30323312. PubMed Central PMCID: PMC6535181.

Four Base Editing Reporters to Monitor and Enrich Editing in Real-time

By Alyssa Cecchetelli | Jul 7, 2020

Base editors create specific point mutations in the genome, but they're inefficient compared to CRISPR/Cas9 edits that rely on double strand DNA breaks. Due to this inefficiency it is crucial for scientists to not only easily identify base editing events in real-time but also enrich for base-edited cells in their experiments. In the past few years, scientists have created an array of base editing reporters that can help you do just that.

Monitoring APOBEC and Cas9-mediated editing in real time

The Harris lab created ACE, a reporter that monitors APOBEC (a family of cytidine deaminase proteins) and Cas9-mediated editing in real time (St. Martin et al., 2018). The ACE reporter is a bicistronic construct that consists of a mutated mCherry and a downstream constitutively active eGFP. To create the inactive mCherry, the lab introduced a 43 base-pair insertion that disrupts fluorescence due to a frameshift. Restoration of fluorescence can only occur through editing by a APOBEC-Cas9n-UGI complex. So how exactly does this work?

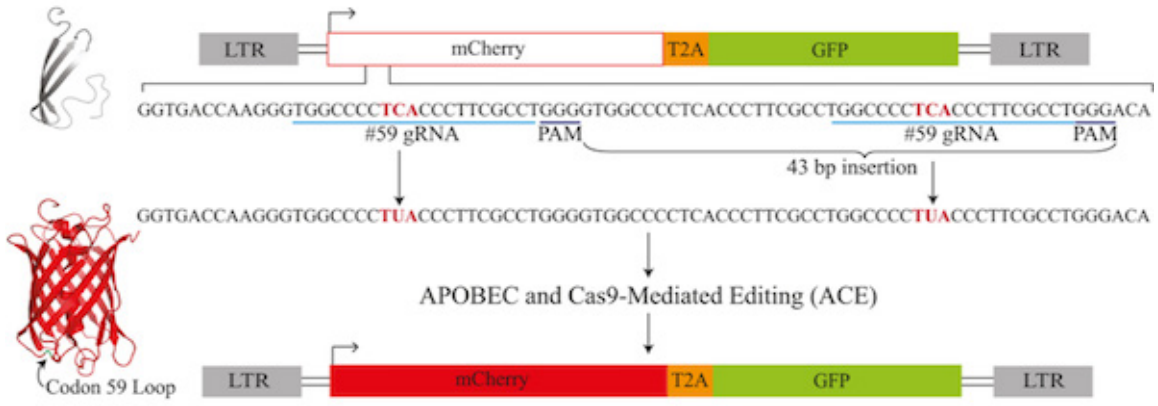


Figure 1: APOBEC- and Cas9-mediated editing (ACE) reporter in action. Image from St. Martin et al., 2018.

APOBEC-Cas9n-UGI targets its preferred trinucleotide motifs (5'-TCA-to-TUA) found within the 43 base-pair mCherry insertion. Editing of these motifs generates lesions that are subject to excision and ssDNA cleavage by canonical base excision repair enzymes. At the same time Cas9n cleaves the opposing DNA strain which results in two DSBs that can be repaired by NHEJ restoring mCherry expression. Cells can then be sorted using Fluorescence-activated cell sorting (FACS) and the editing efficiency can be easily determined by comparing cells expressing only eGFP expression to cells expressing eGFP and mCherry. ACE was used to identify cells that have been successfully edited with APOBEC3A and APOBEC3B (St. Martin et al., 2018).

Four Base Editing Reporters to Monitor and Enrich Editing in Real-time (CONT'D)

eGFP reporters for single base editing by APOBEC-Cas9

Though the ACE reporter quantifies APOBEC-Cas9 efficiency in real-time, it relies on DSBs which is not ideal for monitoring single base edits (St. Martin et al., 2018).

To circumvent sequencing and the need for a DSB in their previous reporter, the Harris Lab created a panel of eGFP reporters to quantify the on-target DNA editing efficiency of APOBEC-Cas9 editisome complexes in real-time (St. Martin et al., 2019). To create the eGFP reporters the Harris lab individually mutated three codons in eGFP to eliminate fluorescence. This created three inactivated eGFP reporters- eGFP L202, eGFP L138, and eGFP L93.

The modified eGFPs were placed downstream of wild-type mCherry and a T2A site and expressed in cells. mCherry is constitutively expressed to identify successfully transfected cells. C-to-T editing by the APOBEC-Cas9 editosome would restore eGFP fluorescence, making it easy to quantify single base editing efficiency by dividing the number of eGFP and mCherry positive cells by the number of only mCherry positive cells (Martin et al., 2019).

The Harris lab used these eGFP reporters to analyze the base editing capability of the seven human APOBEC3 proteins, finding that APOBEC3A and APOBEC3bctd are the most efficient (Martin et al., 2019).

Transient reporter for editing enrichment (TREE)

The Wang and Brafman lab developed a transient reporter for editing enrichment (TREE) to purify single base edited cells without the need for single cell isolation and downstream sequencing (Standage-Beier et al. 2019). TREE is a real-time, fluorescent based identification system for the isolation of base-edited cell populations.

To develop this method the lab was inspired by work that used an integrated GFP reporter that is converted into BFP (blue fluorescent protein) upon CRISPR/Cas9 homology directed repair (HDR) (Glasser et al., 2016). For TREE they engineered a BFP variant that undergoes conversion to GFP after being targeted by a cytidine deaminase base DNA editor. Specifically the BFP mutant (BFPH66) contains a histidine at amino acid 66. This histidine is encoded by a 'CAC' codon that is converted to a 'TAC' or 'TAT' after a C-to-T base editing event. This edit changes the histidine to a tyrosine generating a GFP variant (GFPY66) that has a shifted emission spectra. Thus a successful base editing event would result in GFP fluorescence that can be visualized and sorted via flow cytometry.

The team put TREE to the test by infecting HEK-293 cells with the BFP variant, a BE4am-Cas9 base editor, and sgRNA vectors for the BFP-to-GFP conversion and to target specific loci of interest in the

Four Base Editing Reporters to Monitor and Enrich Editing in Real-time (CONT'D)

genome. The Wang and Brafman lab showed that GFP positive cells isolated with TREE have a significantly higher frequency of genomic base pair edits of interest compared to cells segregated solely by a reporter of transfection, which only report the efficiency of plasmid delivery to a cell. In contrast, TREE provides a read out for plasmid delivery and subsequent base editing efficiency.

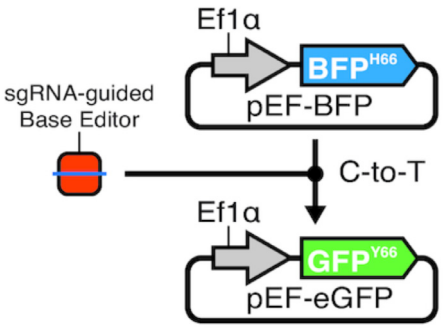


Figure 2: Targeting pEF-BFP with a cytidine deaminase base editor results in shift in emission spectra from BFP to GFP. Image from Standage-Beier et al., 2019.

"Gene On" (GO)- a functional reporter system to identify and enrich for base-editing activity

The reporters mentioned above all rely on GFP fluorescence and cell sorting using FACS which could limit their use across different cell types and systems.

To introduce flexibility, the Dow lab created a base editing reporter that detects and enriches for base editing events *in vivo* without relying solely on GFP (Katti et al., 2020). They named this reporter "Gene On" or GO. GO works by affecting protein translation of different reporter proteins. Protein translation requires the start codon of AUG immediately downstream of a kozak sequence. The Dow lab hypothesized that if they created an ATG codon from an ACG codon using a base-editor (C>T conversion) they could induce the translation of any detectable protein, not just fluorescent proteins.

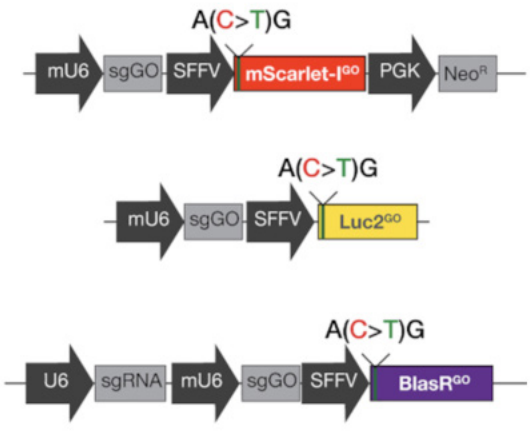


Figure 3: Reporters in Gene On include a mScarlet reporter, luciferase reporter, and neomycin resistance reporter. Image from Katti et al., 2020.

As a proof of concept the Dow lab tested their hypothesis with GFP fluorescence. To do this they generated a silent GFP construct, that contains an ACG start codon mutation, and integrated it into human and mouse cells that expressed either Cas9 or an optimized cytidine base editor. The lab then introduced an sgRNA targeting the GFP ACG to the cells. As expected, only the cells that contained the cytidine base editor exhibited robust induction of GFP fluorescence confirming the efficacy of GO as a base-editing reporter.

As the initiation of protein translation at the start codon ATG is universal, the Dow lab successfully used GO to induce the translation of an array of different reporters including mScarlet-I, luciferase, antibiotic resistance markers, and other enzymes such as Cre-recombinase drastically expanding the base-editing reporter toolbox. The Dow lab also demonstrated

Four Base Editing Reporters to Monitor and Enrich Editing in Real-time (CONT'D)

that GO is an effective reporter for adenine base editors, that make A-to-G edits in the genome (Gaudelli et al., 2017). GO is thus a flexible and adaptable tool to identify and enrich for base-editing events *in vivo*.

References

Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR (2017) Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature 551:464–471 . https://doi.org/10.1038/nature24644

Glaser A, McColl B, Vadolas J (2016) GFP to BFP Conversion: A Versatile Assay for the Quantification of CRISPR/Cas9-mediated Genome Editing. Molecular Therapy - Nucleic Acids 5:e334 . https://doi.org/10.1038/mtna.2016.48

Katti A, Foronda M, Zimmerman J, Diaz B, Zafra MP, Goswami S, Dow LE (2020) GO: a functional reporter system to identify and enrich base editing activity. Nucleic Acids Research 48:2841–2852 . https://doi.org/10.1093/nar/gkaa124

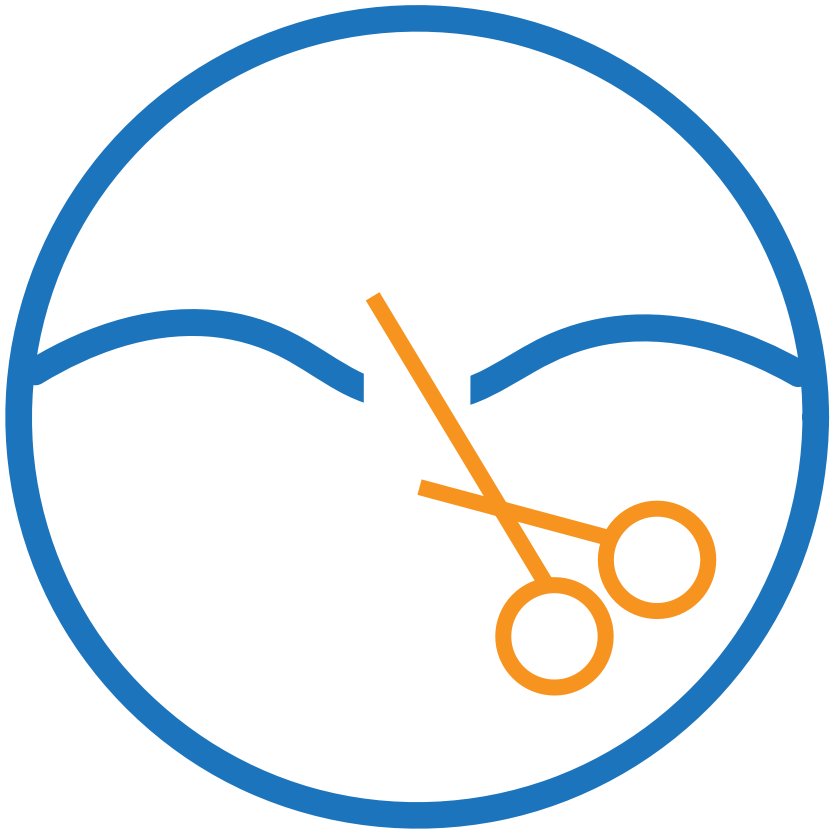
Martin AS, Salamango DJ, Serebrenik AA, Shaban NM, Brown WL, Harris RS (2019) A panel of eGFP reporters for single base editing by APOBEC-Cas9 editosome complexes. Scientific Reports 9: . https://doi.org/10.1038/s41598-018-36739-9

St. Martin A, Salamango D, Serebrenik A, Shaban N, Brown WL, Donati F, Munagala U, Conticello SG, Harris RS (2018) A fluorescent reporter for quantification and enrichment of DNA editing by APOBEC–Cas9 or cleavage by Cas9 in living cells. Nucleic Acids Research 46:e84–e84 . https://doi.org/10.1093/nar/gky332

Standage-Beier K, Tekel SJ, Brookhouser N, Schwarz G, Nguyen T, Wang X, Brafman DA (2019) A transient reporter for editing enrichment (TREE) in human cells. Nucleic Acids Research 47:e120–e120 . https://doi.org/10.1093/nar/gkz713

Chapter 5

Cas Function: RNA Editing



Targeting RNA with Cas13a (C2c2)

By Joel McDade | Sep 21, 2017, Updated Jul 27, 2020 by Cary Valley

CRISPR, and specifically Cas9 from *S. pyogenes* (SpCas9), is truly an exceptional genome engineering tool. It is easy to use, functional in most species, and has many applications. That said, SpCas9 is not the only game in town, and other Cas proteins like SaCas9 and Cpf1 can circumvent the limitations associated with SpCas9. Cas13a (previously referred to as C2c2), has several unique properties that further expand the CRISPR toolbox. We'll cover how Cas13a was identified, the structure and function of Cas13a with a focus on what makes this molecule unique, and the various applications of Cas13a.

The origins of Cas13a: An RNA cleaving CRISPR nuclease

Cas13a was originally identified in 2015 (Shmakov et al., 2015). They were using Cas1, a gene commonly associated with CRISPR arrays and involved in spacer acquisition following infection, as a form of "bait" to identify new CRISPR-associated proteins within the bacterial genome. From this analysis, they identified 53 potential candidate genes that fell into 3 categories based on the architecture of the CRISPR protein in question: C2c1, C2c2, and C2c3 (short for Class 2, candidate 1, 2, or 3). C2c1 and C2c3 are related to Cpf1, but they require both a tracrRNA and crRNA to cleave target DNA (Cpf1 requires only a crRNA for target recognition and cleavage). We will focus on C2c2 (now referred to as Cas13a) since it has a structure and function unique from C2c1 and C2c3.

Perhaps the biggest difference between Cas13a and Cas9 is that Cas13a binds and cleaves RNA rather than DNA substrates. In terms of structure, Cas13a shares no homology to the most commonly used CRISPR enzymes - Cas13a contains two HEPN domains, whereas Cas9 uses HNH and RuvC domains to cleave target DNA. The HEPN domains within Cas13a are essential for RNA cleavage, consistent with known roles for HEPN domains in other proteins. As with Cas9, mutating key residues in the Cas13a molecule results in a "nuclease dead" Cas13a (dCas13a), that is capable of binding target RNA but lacks the ability to cleave the RNA target.

Table 1: Comparison of Cas9, Cpf1, and Cas13a (C2c2)

Name	Enzymatic Domains	Guide RNA	Target	PAM	Cleavage Mechanism
Cas9	HNH, RuvC	TracrRNA, crRNA	DNA	5' NGG	Specific blunt ended DSB in target DNA
Cpf1	RuvC-like	crRNA	DNA	5' TTN	Specific DSB in target DNA with 5' overhangs
Cas13a (C2c2)	2x HEPN	crRNA	RNA	3' A, U, or C (not required by all orthologs)	Specific RNA cleavage (followed by non-specific RNAse activity in bacteria)

Targeting RNA with Cas13a (C2c2) (CONT'D)

Cas13a targeting with a single crRNA

The Feng Zhang lab used Cas13a from *Leptotrichia shahii* (LshCas13a) to begin to characterize this new effector (Abudayyeh et al., 2016). In the endogenous CRISPR/Cas9 system, Cas9 uses both a tracrRNA and crRNA to facilitate binding and cleavage of target DNA respectively. LshCas13a, on the other hand, uses only a short ~24 base crRNA that interacts with the Cas13a molecule through a uracil-rich stem loop and facilitates target binding and cleavage through a series of conformational changes in the Cas13a molecule. Like Cas9, Cas13a tolerates single mismatches between the crRNA and target sequence, however cutting efficiency of Cas13a is reduced when 2 mismatches are present. The protospacer flanking sequence (PFS) for LshCas13a, which is analogous to the PAM sequence for Cas9, is located at the 3' end of the spacer sequence and consists of a single A, U, or C base pair.

In bacteria, once Cas13a has recognized and cleaved its target RNA sequence as specified by the crRNA sequence, it adopts an enzymatically "active" state rather than reverting to an inactive state like Cas9 or Cpf1. As a result, Cas13a will then bind and cleave additional RNAs regardless of homology to the crRNA or presence of a PFS. This is in stark contrast to Cas9, which requires that each DNA target have high sequence identity to the spacer sequence and contain the appropriate PAM sequence. The non-specific cleavage is thought to activate programmed cell death or a dormant state for bacterial cells that have been infected with bacteriophage as to limit the spread of infection throughout the entire population. This property of Cas13a opens up the possibility of using Cas13a as a diagnostic tool, as discussed below.

Applications of Cas13a

What are the potential applications of Cas13a given what we know about its structure and function? For starters, Cas13a can be used to bind and cleave target RNAs, although its usefulness in bacterial cells will likely be limited by the propensity of Cas13a to bind and cleave RNA non-specifically. The ability of dCas13a to bind target RNA without cleaving the molecule (like dCas9) suggests that dCas13a may be useful for isolating specific RNA sequences from a population (either enriching or depleting specific RNAs out of a pool of RNAs) or studying RNA processing in live cells. Indeed, groups have utilized dCas13 fusion proteins for imaging, tracking, modulating splicing, and regulating expression of specifically targeted RNAs.

Finally (and perhaps most interesting), Cas13a's propensity to cleave any RNAs after binding a user-defined RNA sequence could be used to detect single molecules of an RNA species with high specificity (Gootenberg et al., 2017). This system, dubbed SHERLOCK (depicted in figure 1) has been used to differentiate strains of Zika virus, genotype human DNA, identify tumor mutations within cell-

Targeting RNA with Cas13a (C2c2) (CONT'D)

free genomic DNA, and detect COVID-19 (Joung, Ladha et al., 2020).

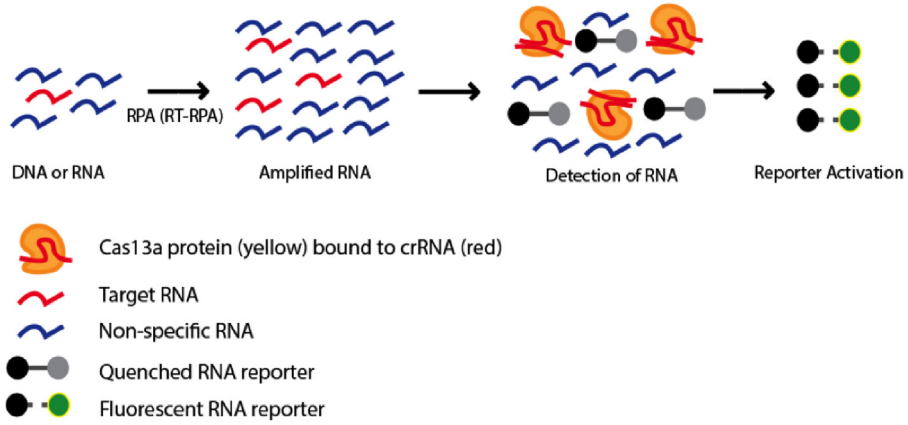


Figure 1: Steps of using Cas13a as a diagnostic tool.

Cas13a in mammalian cells

The Zhang lab made the exciting discovery that some Cas13a orthologs can function in mammalian and plant cells (Abudayyeh et al., 2017). They characterized Cas13a from *Leptotrichia wadei* (LwaCas13a), using a superfolder GFP fusion to stabilize the protein in mammalian cells. LwaCas13a-msfGFP can mediate cleavage with spacers from 20-28 nucleotides in length and does not require a specific PFS, making it a very flexible cleavage system. LwaCas13a is also amenable to multiplexing; expressing 5 gRNAs simultaneously produces comparable gene knockdown to single gRNA expression.

Although the collateral RNA cleavage response is well described in bacterial cells, LwaCas13a does not display nonspecific RNA cleavage activity in eukaryotic cells. They arrived at this finding through analyses of global RNA expression, cell stress, and RNA size distribution. When comparing LwaCas13a to shRNA, they found that LwaCas13a displays similar knockdown efficiency but no significant off-target effects, a distinct contrast to the hundreds of observed shRNA off-targets. They subsequently used catalytically inactive dLwaCas13a as a programmable RNA-binding protein, creating the fluorescent fusion protein dLwaCas13a-NF for *in vivo* RNA imaging.

In short, the Cas13a molecule brings a much anticipated element of diversity to the CRISPR toolbox and will continue to expand the applications of CRISPR. We've also recently seen the new development of Cas13b as a tool for RNA base editing, and we are excited to see this group of enzymes continue to make a splash in the CRISPR field.

Targeting RNA with Cas13a (C2c2) (CONT'D)

References

1. Abudayyeh, Omar O., et al. "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector." Science 353(6299) (2016):aaf5573. PubMed PMID: 27256883. PubMed Central PMCID: PMC5127784.

2. East-Seletsky, Alexandra, et al. "Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection." Nature 538(7624) (2016):270-273. PubMed PMID: 27669025.

3. Gootenberg, Jonathan S., et al. "Nucleic acid detection with CRISPR-Cas13a/C2c2." Science (2017): eaam9321. PubMed PMID: 28408723. PubMed Central PMCID: PMC5526198.

4. East-Seletsky, Alexandra, et al. "RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes. " Mol Cell. 66(3) (2017):373-83. PubMed PMID: 28475872.

5. Abudayyeh, Omar O., et al. "RNA targeting with CRISPR-Cas13." Nature 550(7675) (2017):280-284. PMID: 28976959.

6. Shmakov, Sergey, et al. "Discovery and functional characterization of diverse class 2 CRISPR-Cas systems." Molecular cell 60.3 (2015): 385-397. PubMed PMID: 26593719. PubMed Central PMCID: PMC4660269.

7. Joung J, Ladha A, et al., 'Point-of-care testing for COVID-19 using SHERLOCK diagnostics.' medRxiv 2020 May 8. Pubmed PMID: 32511521. PubMed Central PMCID: PMC7273289.

RNA Editing with Cas13

By Mary Gearing | Nov 30, 2017, Updated Jul 31, 2020 by Cary Valley

Cas13 enzymes are quickly becoming major players in the CRISPR field. Just a year after Feng Zhang's lab identified Cas13a (C2c2) (Abudayyeh et al., 2016) as a RNA-targeting CRISPR enzyme, they adapted Cas13b for precise RNA editing (Cox et al., 2017). This new system, termed REPAIR (RNA editing for programmable A to I (G) replacement) is the first CRISPR tool for RNA editing. Two years after that, the lab published a paper on an RNA editor that allows C to U edits (RESCUE). We'll walk through how these tools were developed and potential ways you can use it in your research.

RNA editing advantages

RNA editing has multiple advantages over more traditional DNA editing systems; first, RNA editing doesn't require homology-directed repair (HDR) machinery, and could thus be used in non-dividing cells. Cas13 enzymes also don't require a PAM sequence at the target locus, making them more flexible than Cas9/Cpf1. Some Cas13 enzymes prefer targets with a given single base protospacer flanking site (PFS) sequence, but orthologs like LwaCas13a do not require a specific PFS. Cas13 enzymes do not contain the RuvC and HNH domains responsible for DNA cleavage, so they cannot directly edit the genome. A Cas13-based RNA editing system would likely be reversible and would avoid genomic off-targets or indels introduced through non-homologous end joining (NHEJ).

Designing an RNA editor

The Zhang lab envisioned a two-component RNA editor: a Cas13 enzyme fused to an RNA adenosine deaminase (ADAR). Such a system would convert adenine to inosine, which the translational machinery treats like guanine. This RNA editor would permit point mutations in RNA which could recapitulate or rescue known pathogenic alleles, or introduce a premature stop codon to render an RNA nonfunctional.

In their quest to build a robust RNA editor, the lab started with the Cas13 scaffold, testing a whopping 21 Cas13a orthologs, 15 Cas13b orthologs, and 7 Cas13c orthologs. They hoped to find a stably folded ortholog that cleaved RNA robustly, unlike LwaCas13a, which must be stabilized by monomeric superfolded GFP and averages only ~50% RNA knockdown. In initial tests, Cas13b from *Prevotella sep. P5-125* (PspCas13b) yielded 62.9% average knockdown, and they chose this enzyme for further studies. PspCas13b does not require a PFS, and it is sensitive to mismatches in target RNA from bases 12-26 of a 30 nt target sequence.

For the editing portion of the protein, they examined ADAR1 and ADAR2, which deaminate adenosine to inosine in RNA, creating a functional A->G change. They fused ADAR deaminase domains (ADARDD) to dPspCas13b, but observed low RNA editing. To increase A->G editing, they employed hyperactive ADAR constructs, like ADAR2DD(E488Q). They also adjusted the structure of the guide RNA by

RNA Editing with Cas13 (CONT'D)

introducing a C opposite the target A to be edited. This change specifies the edit to be made when multiple As are present in the gRNA spacer, as ADAR will preferentially edit an adenine if the template has a cytosine mismatch at that position.

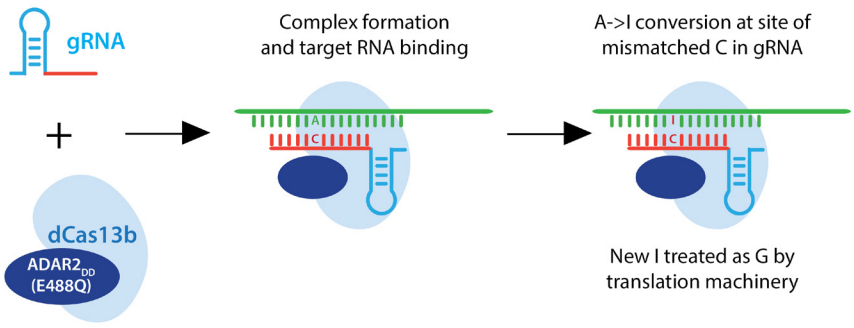


Figure 1: gRNA and dCas13b complex forms and targets the RNA. The A is converted to an I at the site of mismatch. The new I is treated as a G by translation machinery.

PspCas13b-ADAR2DD(E488Q) displayed robust editing with various spacer lengths from 30-84 nucleotides. This system was designated REPAIRv1. Using next-generation sequencing, they confirmed A->I editing and found that 50 nt spacers increase editing efficiency but also increase off-target editing within the target window, possibly due to longer stretches of duplexed RNA. Using an ADAR2DD catalytic mutant, they showed that editing is mediated by ADAR2DD, not PspCas13b.

Testing and improving RNA editing from the Cas13b REPAIR system

To test the robustness of REPAIRv1, the lab targeted 34 pathogenic G->A mutations from the ClinVar database. They successfully edited 33/34 sites with a maximum of 28% editing efficiency in HEK293T cells, as assessed using RNA-seq. Since the REPAIRv1 machinery is too large to fit into adeno-associated viral vectors, they tested truncated ADARDDs to see if they could shrink the construct. They identified ADAR2DD(delta984-1090), which decreases the REPAIRv1 construct size from 4,473 bp to 4,152 bp without decreasing editing efficiency.

Although REPAIRv1 knockdown is more precise than shRNA knockdown, it still displays substantial off-target activity. By mutating residues in ADAR2 that interact with duplex RNA, they hoped to destabilize ADAR-RNA binding to decrease off-targets. Of the tested constructs, ADAR2DD(E488Q/T375G) (REPAIRv2) had the highest on-target efficiency and the lowest amount of off-target editing. In one direct transcriptome-wide comparison, REPAIRv2 reduced off-target sites from the 18,385 observed with REPAIRv1 to merely 20.

RNA Editing with Cas13 (CONT'D)

Expanding RNA Editing Capabilities through RESCUE

Two years later, the Zhang lab expanded the RNA editing toolkit to allow C to U edits using their RESCUE (RNA Editing for Specific C-to-U Exchange) system (Abudayyeh et al., 2019). To avoid some of the inherent drawbacks for RNA editing with natural cytidine deaminases, they evolved the adenine deaminase ADAR2DD to deaminate cytidine and used dRanCas13b (catalytically inactive Cas13) to target the cytosine deaminase. Three rounds of rational mutagenesis on the ADAR2DD residues that contact the RNA substrate resulted in 15% C-to-U editing activity. They then performed sixteen rounds of directed evolution across the full ADAR2DD to further optimize cytidine deaminase activity. While RESCUE can perform C-to-U edits, the authors note that it retains its adenine deaminase ability.

The authors then tested the RESCUE system's ability to edit endogenous transcripts finding up to 42% editing efficiency. They note that RESCUE is most active in a 30 nt guide when C or U base-flips are present across from the target base. However, they saw that in addition to the targeted C-to-U edit, RESCUE had A-to-I off-target edits around the targeted nucleotide. Introduction of guanine-mismatches in the guide across from these off-target edits helped to minimize local off-target editing by RESCUE. To address both C-to-U and A-to-I transcriptome-wide off target effects observed in whole-transcriptome RNA-sequencing, the authors performed rational mutagenesis and identified that a S375A mutation in RESCUE (named RESCUE-S) offered ~76% on-target C-to-U editing while reducing off-target C-to-U edits by ~45% and off-target A-to-I edies by ~94% compared to RESCUE.

Addressing the potential use of RESCUE in therapeutic applications, the authors tested various dRanCas13b truncations that would allow packaging the construct for viral delivery. They found C-terminal truncations allowed the same or improved editing ability in a RESCUE system small enough for viral delivery.

Future directions

Building on previous RNA targeting work with Cas13, REPAIR is the first CRISPR-based system to enable precise RNA editing. The addition of RESCUE further expands the RNA editing toolkit and sequences that can be edited. As discussed above, the ability to edit RNA has multiple advantages, including reversibility and use in non-dividing cells. With the expanded targeting ability using both systems, the authors look forward to the expanded use of RNA editing in research and its potential therapeutic uses.

References

Cox, David B.T., et al. "RNA editing with CRISPR-Cas13." Science (2017):pii: eaaq0180. PMID: 29070703.

RNA Editing with Cas13 (CONT'D)

Abudayyeh, Omar O., et al. "RNA targeting with CRISPR-Cas13." Nature 550(7675) (2017):280-284. PMID: 28976959.

Abudayyeh, Omar O., et al. "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector." Science 353(6299) (2016):aaf5573. PubMed PMID: 27256883. PubMed Central PMCID: PMC5127784.

East-Seletsky, Alexandra, et al. "Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection." Nature 538(7624) (2016):270-273. PubMed PMID: 27669025.

Gootenberg, Jonathan S., et al. "Nucleic acid detection with CRISPR-Cas13a/C2c2." Science (2017): eaam9321. PubMed PMID: 28408723. PubMed Central PMCID: PMC5526198.

East-Seletsky, Alexandra, et al. "RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes." Mol Cell. 66(3) (2017):373-83. PubMed PMID: 28475872.

Abudayyeh, Gootenberg, et al. "A cytosine deaminase for programmable single-base RNA editing." Science 2019 Jul 26;365(6451):382-386. PubMed PMID: 31296651

Cas13d: Small RNA-targeting CRISPR enzymes for transcriptome engineering

By Mary Gearing | May 3, 2018

RNA-editing Cas13 enzymes have taken the CRISPR world by storm. Like RNA interference, these enzymes can knock down RNA without altering the genome, but Cas13s have higher on-target specificity. New work from Konermann et al. and Yan et al. describes new Cas13d enzymes that average only 2.8 kb in size and are easy to package in low-capacity vectors! These small, but mighty type VI-D enzymes are the latest tools in the transcriptome engineering toolbox.

How were Cas13d enzymes discovered?

Microbial CRISPR diversity is impressive, and researchers are just beginning to tap the wealth of CRISPR possibilities. To identify Cas13d, both groups used very general bioinformatic screens that looked for a CRISPR repeat array near a putative effector nuclease. The Cas13d proteins they identified have little sequence similarity to previously identified Cas13a-c orthologs, but they do include HEPN nuclease domains characteristic of the Cas13 superfamily. Yan et al. proceeded to study orthologs from *Eubacterium siraeum* (EsCas13d) and *Ruminococcus* sp. (RspCas13d), while Konermann et al. characterized orthologs from "Anaerobic digester metagenome" (AdmCas13d) and *Ruminococcus flavefaciens* (nicknamed CasRx), as well as EsCas13d.

- Find the Yan et al. plasmids here
- Find the Konermann et al. plasmids here

How do Cas13d enzymes compare to other Cas13s?

Like other Cas13 enzymes, the Cas13d orthologs described in these papers can independently process their own CRISPR arrays into guide RNAs. crRNA cleavage is retained in dCas13d and is thus HEPN-independent. These enzymes also do not require a protospacer flanking sequence, so you can target virtually any RNA sequence! In bacteria, Cas13d-mediated cleavage promotes collateral cleavage of other RNAs. As with other Cas13s, this collateral cleavage does not occur when Cas13d is expressed in a mammalian system.

Since Cas13d is functionally similar to previously discovered Cas13 enzymes - what makes these orthologs so special? The first property is size - Cas13d enzymes have a median length of ~930 amino acid - making them 17-26% smaller than other Cas13s and a whopping 33% smaller than Cas9! Their small size makes then easy to package in low-capacity vectors like AAV, a popular vector due to its low immunogenicity. But these studies also identified other advantages, including Cas13d-specific regulatory proteins and high targeting efficiency.

Cas13d: Small RNA-targeting CRISPR enzymes for transcriptome engineering (CONT'D)

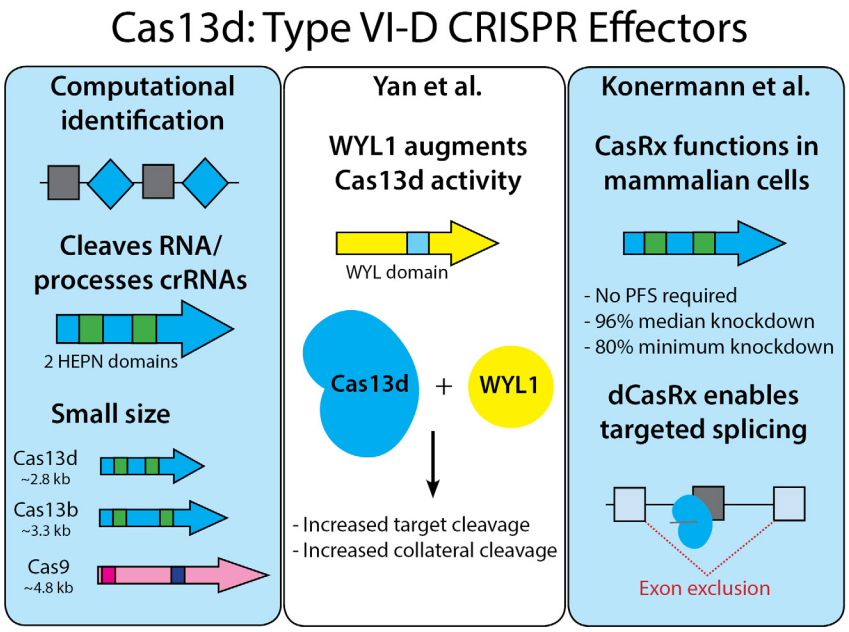


Figure 1: A comparison between Cas13d and CasRx

Yan et al.: WYL-domain-containing accessory protein WYL1 increases RspCas13d and EsCas13d cleavage activity

The majority of Type VI-D loci contain accessory proteins with WYL domains (named for the three conserved amino acids in the domain). Yan et al. from Arbor Biotechnologies found that RspCas13d accessory protein RspWYL1 increases both targeted and collateral RNA degradation by RspCas13d. RspWYL1 also increased EsCas13d activity, indicating that WYL domain-containing proteins may be broader regulators of Cas13d activity. This property makes WYL proteins an intriguing counterpart to anti-CRISPR proteins that negatively modulate the activity of Cas enzymes, some of which are also functional in multiple species (read Arbor Biotechnologies' press release about their Cas13d deposit here).

Konermann et al.: CasRx is a highly potent, specific editor in mammalian cells

Not all Cas13d proteins are functional in mammalian cells, but Konermann et al. saw great results with CasRx and AdmCas13d fused to a nuclear localization signal (NLS). In a HEK293 mCherry reporter assay, CasRx and AdmCas13d produced 92% and 87% mCherry protein knockdown measured by flow cytometry, respectively. Cas13d CRISPR array processing is robust, with CasRx and either an unprocessed or processed gRNA array (22 nt spacer with 30 nt direct repeat) mediating potent knockdown. Multiplexing from the CRISPR array yielded >90% knockdown by CasRx for each of four targets, including two mRNAs and two nuclear long non-coding RNAs.

One interesting twist to Cas13d enzymes is their cleavage pattern: EsCas13d produced very similar cleavage products even when guides were tiled across a target RNA, indicating that this enzyme does not cleave at a predictable distance from the targeted region. Konermann et al. show that EsCas13d

Cas13d: Small RNA-targeting CRISPR enzymes for transcriptome engineering (CONT'D)

favors cleavage at uracils, but a more detailed exploration of this cleavage pattern is necessary.

Konermann et al. compared CasRx to multiple RNA regulating methods: small hairpin RNA interference, dCas9-mediated transcriptional inhibition (CRISPRi), and Cas13a/Cas13b RNA knockdown. CasRx was the clear winner with median knockdown of 96% compared to 65% for shRNA, 53% for CRISPRi, and 66-80% for other Cas13a and Cas13b effectors. Like previously characterized Cas13 enzymes, CasRx also displays very high on-target efficiency; where shRNA treatment produced 500-900 significant off-targets, CasRx displayed zero. Unlike Cas9, for which efficiency varies widely across guide RNAs, each guide tested with CasRx yielded >80% knockdown. It seems that CasRx may make it possible to target essentially any RNA in a cell.

Since catalytically dead dCasRx maintains its RNA-binding properties, Konermann et al. tested its ability to manipulate RNA species through exon skipping. Previous CRISPR exon-skipping approaches used two guide RNAs to remove a given exon from the genome, and showed success in models of muscular dystrophy. In this case, Konermann et al. targeted MAPT, the gene encoding dementia-associated tau, delivering dCasRx and a 3-spacer array targeting the MAPT exon 10 splice acceptor and two putative splice enhancers. After AAV-mediated delivery to iPS-derived cortical neurons, dCasRx-mediated exon skipping improved the ratio of pathogenic to non-pathogenic tau by nearly 50%, showing proof-of-concept for pre-clinical and clinical applications of dCasRx.

The identification of Type VI Cas13d enzymes is another win for bioinformatic data mining. As we continue to harness the natural diversity of CRISPR systems, only time will tell how large the genome and transcriptome engineering toolbox will be. It is, however, certain that the impact of CRISPR scientific sharing will continue to grow, and we at Addgene appreciate our depositors for making their tools available to the broader community.

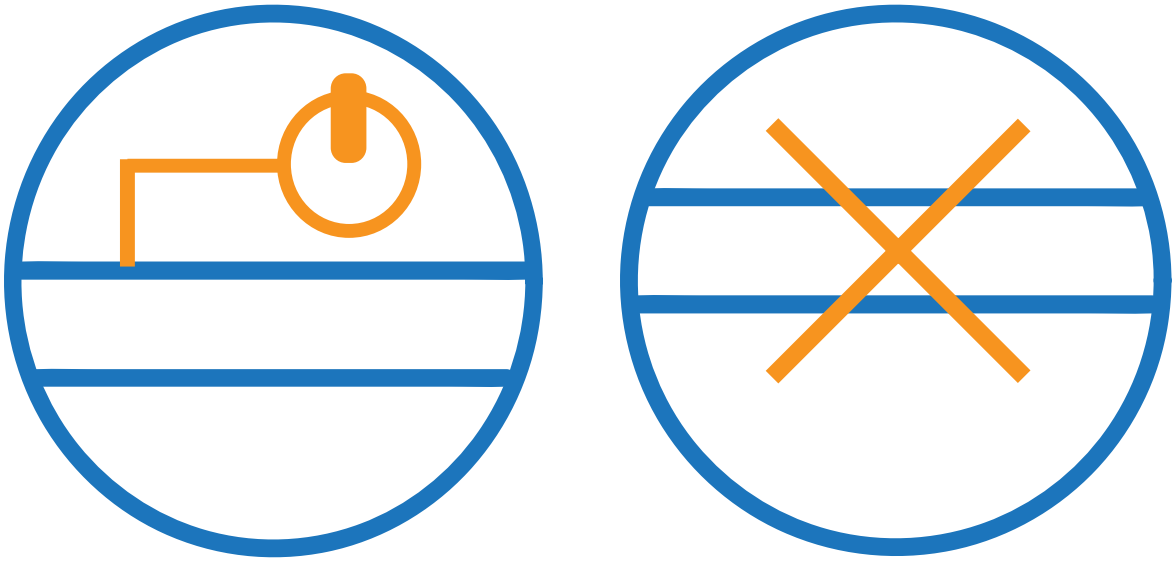
References

Konermann, Silvana, et al. "Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors." Cell (2018) pii: S0092-8674(18)30207-1. PubMed PMID: 29551272

Yan, Winston X., et al. "Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein." Mol Cell. (2018) pii: S1097-2765(18)30173-4. PubMed PMID: 29551514

Chapter 6

Cas Function: Activation/ Repression



CRISPR Activation: A Practical Guide

By Marcelle Tuttle and Alex Chavez | Aug 18, 2016, Updated Oct 7, 2020 by Gabrielle Clouse

CRISPR/Cas9 is an enormously plastic tool and has taken the scientific world by storm. While Cas9 has been most widely used to create specific edits in DNA, there has also been significant work on constructing Cas9 transcriptional activators. These constructs allow for the upregulation of essentially any gene by fusing mutants of Cas9 deficient in DNA cutting activity to a transcriptional activation domain (Fig 1).

Two ways to use CRISPRa for your research

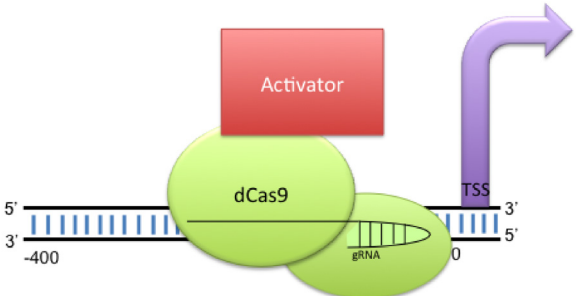


Figure 1: dCas9 binds upstream of a promoter and drives transcription using activators.

CRISPRa for specific gene targets

CRISPR activation for genes with known functions has the potential to be used in therapeutics, as it can increase gene expression at certain loci to create changes in phenotypes. For example, CRISPR activation has been used to differentiate induced pluripotent stem cells (Chavez et al., 2015) into neuronal cells by activating genes that cause differentiation and to reverse HIV viral latency (Bialek et al., 2016).

Additionally, CRISPR-based activation can be made inducible upon environmental triggers. This is done by creating a split dCas9 protein that can reassemble when exposed to specific UV wavelengths or chemicals (Polstein and Gersbach, 2015; Zetsche et al., 2015). These types of CRISPR activation are also reversible: once the environmental trigger leaves, gene expression is no longer activated.

However, there are several factors that may prevent your gene from being up-regulated. For instance, generally, the more highly expressed a gene is under native conditions, the less activation you can achieve using a CRISPRa. Your gene of interest might already be hitting an upper bound of activation that current Cas9 systems cannot help you pass. In addition, activation experiments often require quite a bit of tuning before you know your system is working as expected. Finally, for each gene you want to activate, you should also be ready to test three or four guides directed towards that gene as there can be a large difference in guide potency.

Thus, in cases where users have a single gene they want to activate, we would recommend using a cDNA overexpression vector rather than going through all the troubleshooting required for Cas9-based activation.

CRISPRa for genome-wide studies

One of the best uses for Cas9 activators is in genetic screening. gRNAs targeting every gene in the

CRISPR Activation: A Practical Guide (CONT'D)

human genome, for example, can be made easily and cheaply using oligo library synthesis. Prior to Cas9 activators, similar tools were made using other DNA binding proteins such as zinc fingers (ZF) and TAL effectors (TALE). Unlike these constructs, however, Cas9 allows you to easily change the sequence targeted by the activator by simply providing a new gRNA rather than engineering an entirely new protein. This makes it much cheaper to use Cas9 activators.

This larger scale of activation allows for exploratory-based research as well. For example, CRISPRa can be used to find potential cancer drug targets by finding genes that cause cell death in cancer cell lines upon activation (Gilbert et al., 2014; Behan et al., 2019). Because of its ability to target multiple gene loci simultaneously, CRISPRa can also be used to map molecular pathways and genetic interactions. Additionally, large scale CRISPR-based screens can be used to gain insight into protein structure and function (Pan et al., 2018).

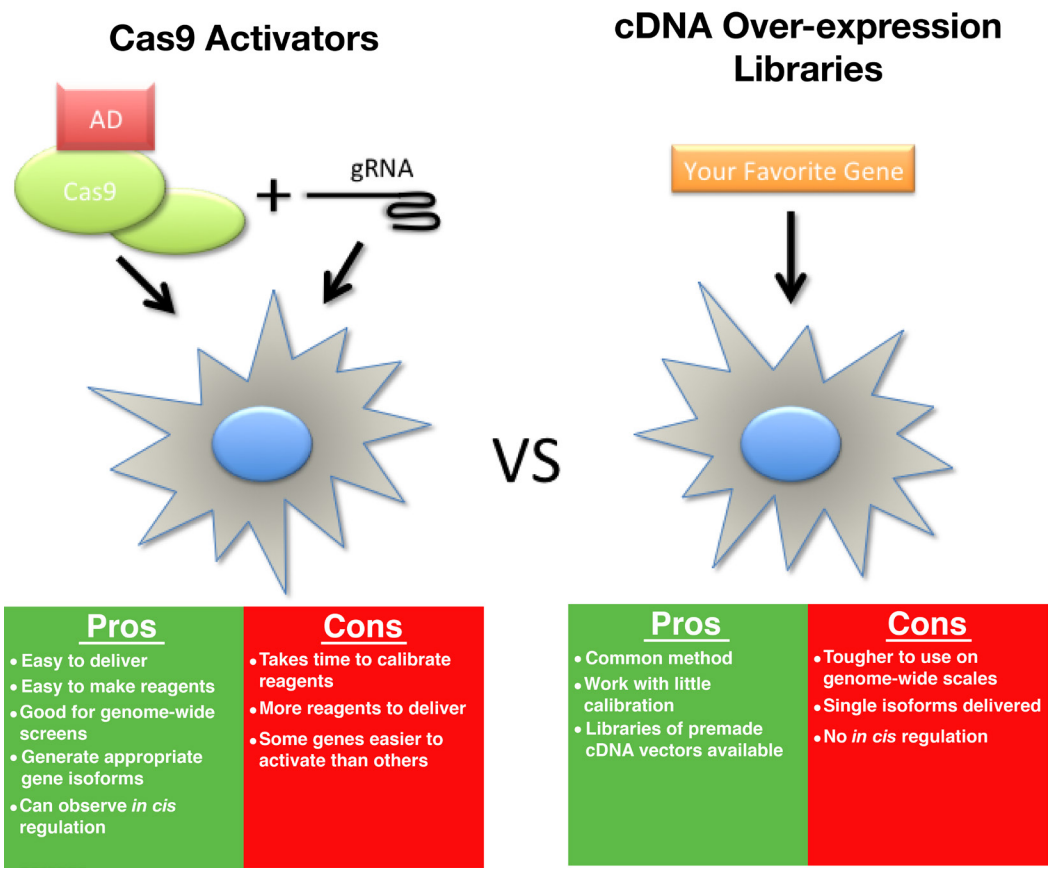


Figure 2: The pros and cons of Cas9 activators and cDNA over-expression libraries

cDNA libraries, which consist of plasmids that over-express coding sequences from a given cell type or organism, have been used in a similar manner to Cas9 activators. However, these can be difficult to construct and deliver when compared to gRNAs. Additionaly, cDNAs cannot be used to study in cis regulation and also suffer from an inability to easily deliver the appropriate isoform(s) of a given gene, as, many times, the isoform(s) common to a particular cell type are unknown or not readily available. By activating from the native context of the gene, Cas9 activators efficiently solve these problems.

CRISPR Activation: A Practical Guide (CONT'D)

Which Cas9 activator should I use?

There are a wide variety of activators you can use for your experiments. We have found that SAM (Konermann et al., 2014), Suntag (Tanenbaum et al., 2014, Gilber et al, 2014), and VPR (Chavez et al., 2015)) are good choices across multiple cell lines (HEK293T, MCF7, U2-OS, Hela, N2A, 3T3) and organisms (Chavez et al., 2016). To learn more about the different types of Cas9 activators, check out the next article.

Worry less about off targets

Unlike Cas9 cutting activity, off target effects are generally not regarded as being a large problem for Cas9 activators. This is believed to be true given the results of previous RNA-seq experiments (Konermann et al., 2014, Chavez et al., 2014, Chavez et al., 2016) along with a belief that the odds are very low that Cas9 would have an off-target that lands in the promoter of another gene, thereby driving aberrant transcription. That being said, we generally pick guides by putting the promoter of the gene into a gRNA finder such as WU-CRISPR (Wong et al., 2015)) or our lab's sgRNA scorer 2.0 (Chari et al., 2015) and picking whichever guides are closest to the transcription start site (TSS). We recommend targeting the guides to a region less than 200 bp upstream of TSS for best results but up to 400 bp works reasonably well.

References

Behan FM, Iorio F, Picco G, Gonçalves E, Beaver CM, Migliardi G, Santos R, Rao Y, Sassi F, Pinnelli M, Ansari R, Harper S, Jackson DA, McRae R, Pooley R, Wilkinson P, van der Meer D, Dow D, Buser-Doepner C, Bertotti A, Trusolino L, Stronach EA, Saez-Rodriguez J, Yusa K, Garnett MJ (2019) Prioritization of cancer therapeutic targets using CRISPR–Cas9 screens. Nature 568:511–516 . https://doi.org/10.1038/s41586-019-1103-9

Bialek JK, Dunay GA, Voges M, Schäfer C, Spohn M, Stucka R, Hauber J, Lange UC (2016) Targeted HIV-1 Latency Reversal Using CRISPR/Cas9-Derived Transcriptional Activator Systems. PLoS ONE 11:e0158294 . https://doi.org/10.1371/journal.pone.0158294

Chari R, Mali P, Moosburner M, Church GM (2015) Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. Nat Methods 12:823–826 . https://doi.org/10.1038/nmeth.3473

Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, P R Iyer E, Lin S, Kiani S, Guzman CD, Wiegand DJ, Ter-Ovanesyan D, Braff JL, Davidsohn N, Housden BE, Perrimon N, Weiss R, Aach J, Collins JJ, Church

CRISPR Activation: A Practical Guide (CONT'D)

GM (2015) Highly efficient Cas9-mediated transcriptional programming. Nat Methods 12:326–328 . https://doi.org/10.1038/nmeth.3312

Chavez A, Tuttle M, Pruitt BW, Ewen-Campen B, Chari R, Ter-Ovanesyan D, Haque SJ, Cecchi RJ, Kowal EJK, Buchthal J, Housden BE, Perrimon N, Collins JJ, Church G (2016) Comparison of Cas9 activators in multiple species. Nat Methods 13:563–567 . https://doi.org/10.1038/nmeth.3871

Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, Qi LS, Kampmann M, Weissman JS (2014) Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell 159:647–661 . https://doi.org/10.1016/j.cell.2014.09.029

Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, Nureki O, Zhang F (2014) Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517:583–588 . https://doi.org/10.1038/nature14136

Pan J, Meyers RM, Michel BC, Mashtalir N, Sizemore AE, Wells JN, Cassel SH, Vazquez F, Weir BA, Hahn WC, Marsh JA, Tsherniak A, Kadoch C (2018) Interrogation of Mammalian Protein Complex Structure, Function, and Membership Using Genome-Scale Fitness Screens. Cell Systems 6:555-568. e7 . https://doi.org/10.1016/j.cels.2018.04.011

Polstein LR, Gersbach CA (2015) A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. Nat Chem Biol 11:198–200 . https://doi.org/10.1038/nchembio.1753

Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD (2014) A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. Cell 159:635–646 . https://doi.org/10.1016/j.cell.2014.09.039

Wong N, Liu W, Wang X (2015) WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. Genome Biol 16:218. https://doi.org/10.1186/s13059-015-0784-0

Zetsche B, Volz SE, Zhang F (2015) A split-Cas9 architecture for inducible genome editing and transcription modulation. Nat Biotechnol 33:139–142 . https://doi.org/10.1038/nbt.3149

A Comparison of the Many CRISPR Activators

By Gabrielle Clouse | Oct 6, 2020

Prior to the discovery of CRISPR/Cas systems, gene activation across multiple loci was an arduous process. When using zinc finger proteins or TALE proteins, proteins had to be re-engineered for each gene, making wide-scale gene activation seem next to impossible. The development of CRISPR/Cas systems, however, greatly improved the simplicity of gene activation: rather than requiring protein engineering for each loci, CRISPR/Cas systems only require changing the programmable guide RNA.

Gene activation by dCas9, also referred to as CRISPRa, was initially published in 2013 (Bikard et al., 2013, Perez-Pinera et al., 2013). In the years that followed, innovative methods greatly improved CRISPRa, expanding its practicality and popularity in research (Tanenbaum et al., 2014, Konermann et al., 2015, Chavez et al., 2015).

Most popular methods for CRISPR activation

CRISPR activation uses dCas9, a CRISPR protein variant lacking its endonuclease ability, to bind to genes without editing the genome (Qi et al., 2013). To target specific sequences, CRISPR/Cas systems rely on a guide RNA complementary to the sequence of interest. Upon binding, CRISPRa systems recruit transcription factors to increase gene expression. CRISPRa methods vary in their transcriptional activators: some methods rely on fusion proteins while others re-engineer components of Cas systems themselves. SunTag, SAM, and VPR have all shown significant improvements upon the initial dCas9-VP64 method, so there are multiple options to choose from when looking to activate genes across diverse cell lines.

dCas9-VP64

Description: CRISPR activation can occur through fusing dCas9 with VP64, a strong transcriptional activation domain. Guided by dCas9, VP64 recruits transcriptional machinery to specific sequences, causing targeted gene regulation. This can be used to activate transcription during either initiation or elongation, depending on which sequence is targeted.

Performance: dCas9-VP64 activation is generally thought of as the "first generation" CRISPR activator. While it requires a relatively simple construct, it exhibits modest levels of gene activation, with some genes experiencing around 2-fold activation levels. Although other methods have been able to achieve much higher activation, dCas9-VP64 is great for experiments that requires modest gene activation.

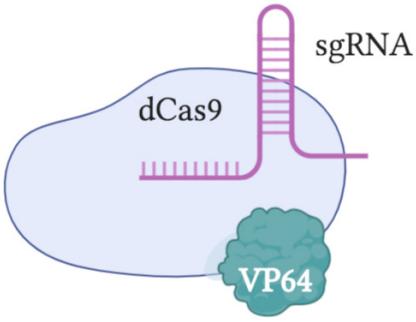


Figure 1: dCas9-VP64. The VP64 domain recruits transcription factors the target DNA location specified by the sgRNA.

A Comparison of the Many CRISPR Activators (CONT'D)

Synergistic Activation Mediator (SAM)

Description: SAM uses specially engineered sgRNAs to increase transcription. This is done through creating a dCas9/VP64 fusion protein engineered with aptamers that bind to MS2 proteins. These MS2 proteins then recruit additional activation domains (HS1 and p65).

Performance: When targeting single genes, SAM consistently shows the highest levels of gene activation compared to other CRISPR activators, making it a popular method for gene activation experiments. In cases of multiplex gene regulation (activating multiple genes at once), however, SAM exhibits activation levels comparable to other popular activation methods (VPR and SunTag) (Chavez et al., 2016).

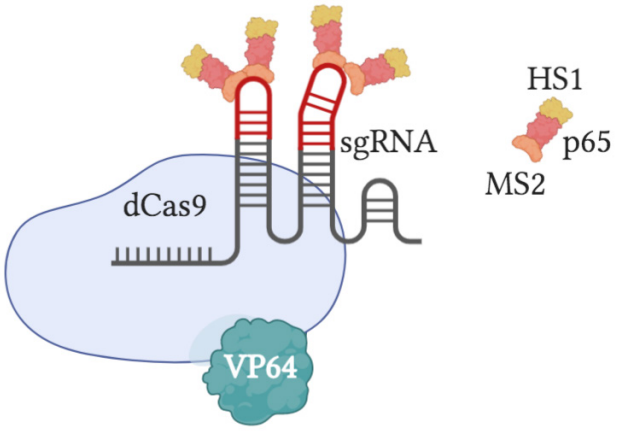


Figure 2: SAM is a dCas9-VP64 fusion with aptamers on the gRNA that bind MS2 proteins.

SunTag

Description: Rather than using a single copy of VP64 per each dCas9, SunTag uses a repeating peptide array to fused with multiple copies of VP64. By having multiple copies of VP64 at each loci of interest, this allows more transcriptional machinery to be recruited per targeted gene.

Performance: SunTag performs better than first generation activators while showing lower activation levels than SAM. One drawback of this method is its construction: it relies on antibody chains, which are relatively large and are not expressed consistently throughout cells.

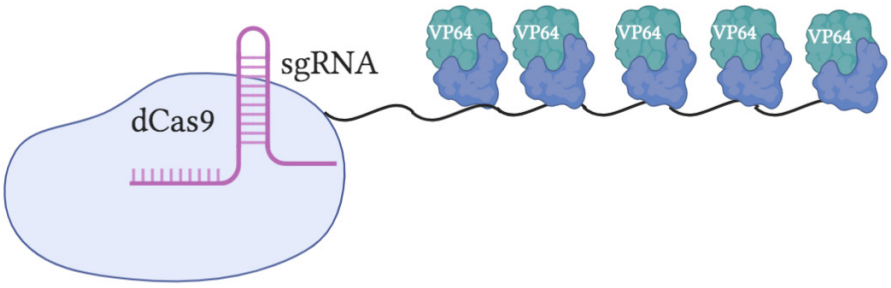


Figure 3: SunTag uses a repeating peptide array fused to multiple copies of VP64.

A Comparison of the Many CRISPR Activators (CONT'D)

VPR

Description: VPR fuses a tripartite complex with dCas9 to activate transcription. This complex consists of the VP64 activator used in other CRISPR activation methods, as well as two other potent transcriptional activators (p65 and Rta). These transcriptional activators work in tandem to recruit transcription factors.

Performance: Generally, VPR is found to have significantly higher activation levels than the initial dCas9-VP64 activator, but lower levels compared to the SAM system. Its activation levels are similar to that of SunTag. An advantage to this method compared to other notable CRISPR activators is that it requires a fusion protein, rather than relying on a two-component system dependent on gRNA design (SAM) or peptide design (SunTag). This streamlines its delivery, making it a common choice for CRISPR activation.

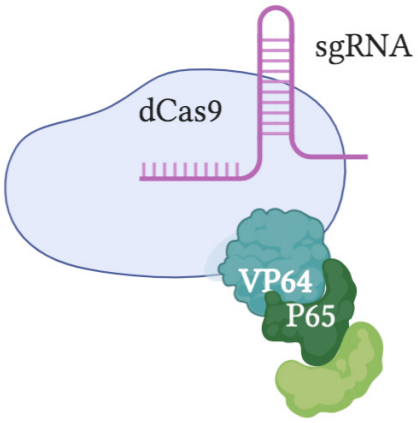


Figure 4: VPR fuses a tripartite complex with dCas9 to activate transcription.

Newer methods for CRISPR activation

While SAM, SunTag, and VPR have become the most popular methods for CRISPR activation, research in the area has continued to develop. These newer methods have not yet been rigorously compared to other CRISPR activators, but their initial results show promise for CRISPR-based gene activation.

dCas9-CBP

In this method, dCas9 is fused to CBP, an histone acetyltransferase domain capable of rearranging chromatin structure. This method has shown higher activation levels than SAM in some loci across Drosophila cell lines (Sajwan et al., 2019).

SPH

SPH combines components of SAM and SunTag activators, fusing the epitope tag of SunTag with the P65-HSF activation domains used in SAM. This method has been shown to activate genes at 2- to 3-fold higher than SAM, SunTag, and VPR (Zhou et al., 2018).

We hope this helps you chose a CRISPR activator for your experiment. Good luck with your experiment!

A Comparison of the Many CRISPR Activators (CONT'D)

References

Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Research 41:7429–7437 . https://doi.org/10.1093/nar/gkt520

Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, P R Iyer E, Lin S, Kiani S, Guzman CD, Wiegand DJ, Ter-Ovanesyan D, Braff JL, Davidsohn N, Housden BE, Perrimon N, Weiss R, Aach J, Collins JJ, Church GM (2015) Highly efficient Cas9-mediated transcriptional programming. Nat Methods 12:326–328 . https://doi.org/10.1038/nmeth.3312

Chavez A, Tuttle M, Pruitt BW, Ewen-Campen B, Chari R, Ter-Ovanesyan D, Haque SJ, Cecchi RJ, Kowal EJK, Buchthal J, Housden BE, Perrimon N, Collins JJ, Church G (2016) Comparison of Cas9 activators in multiple species. Nat Methods 13:563–567 . https://doi.org/10.1038/nmeth.3871

Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, Nureki O, Zhang F (2014) Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517:583–588 . https://doi.org/10.1038/nature14136

Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, Thakore PI, Glass KA, Ousterout DG, Leong KW, Guilak F, Crawford GE, Reddy TE, Gersbach CA (2013) RNA-guided gene activation by CRISPR-Cas9–based transcription factors. Nat Methods 10:973–976 . https://doi.org/10.1038/nmeth.2600

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell 152:1173–1183 . https://doi.org/10.1016/j.cell.2013.02.022

Sajwan S, Mannervik M (2019) Gene activation by dCas9-CBP and the SAM system differ in target preference. Sci Rep 9: 18104. https://doi.org/10.1038/s41598-019-54179-x

Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD (2014) A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. Cell 159:635–646 . https://doi.org/10.1016/j.cell.2014.09.039

Zhou H, Liu J, Zhou C, Gao N, Rao Z, Li H, Hu X, Li C, Yao X, Shen X, Sun Y, Wei Y, Liu F, Ying W, Zhang J, Tang C, Zhang X, Xu H, Shi L, Cheng L, Huang P, Yang H (2018) *In vivo* simultaneous transcriptional activation of multiple genes in the brain using CRISPR–dCas9-activator transgenic mice. Nat Neurosci 21:440–446 . https://doi.org/10.1038/s41593-017-0060-6

Truncated gRNAs for Regulating Gene Expression

By Alissa Lance-Byrne and Alex Chavez | Jan 10, 2017

CRISPR/Cas9 technology has revolutionized the fields of molecular biology and bioengineering, as it has facilitated the development of a simple and scalable means of making targeted genetic edits. Cas9 is a DNA binding protein that can be directed to virtually any genetic locus when complexed with an appropriately designed small RNA, or guide RNA (gRNA). The gRNA conventionally contains a 20-nucleotide sequence that is complementary to the target site, or protospacer, in the genome. Native Cas9 has two catalytic domains, each of which cleaves one strand of DNA upon binding the protospacer. The resulting double strand break (DSB) stimulates DNA repair mechanisms that can be exploited to either inactivate a gene or introduce a desired genetic alteration.

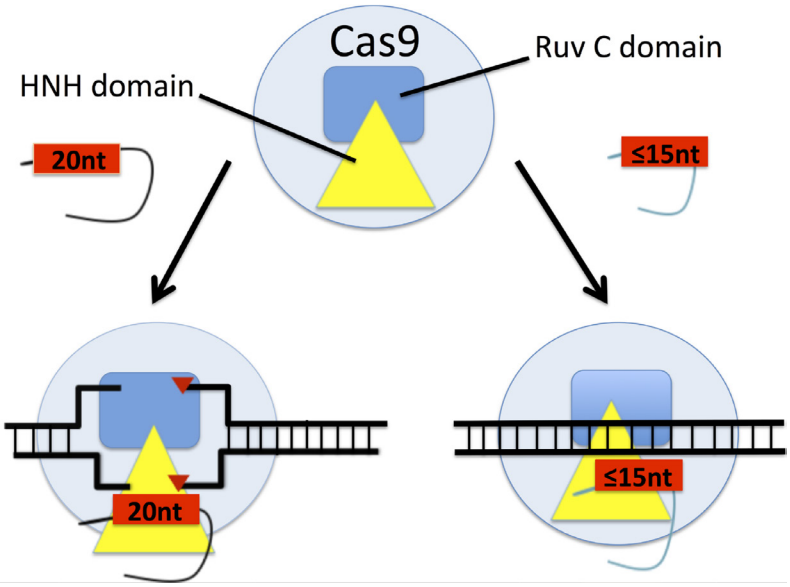


Figure 1: Schematic of native (nuclease-competent) Cas9 interacting with either a full-length gRNA (20 nt complementarity to target site) or truncated gRNA (less than or equal to 15 nt complementarity to target site). When complexed with a full-length gRNA, native Cas9's two catalytic domains, HNH and RuvC, cleave DNA strands complementary and non-complementary to the gRNA, respectively. When complexed with a truncated gRNA, Cas9 binds to the target site, but does not cut either strand of DNA.

Traditional Cas9 regulators

In addition to its utility in making targeted modifications to DNA, Cas9 can be reprogrammed to serve as a regulator of gene expression. Its catalytic domains can be mutated to inactivate the protein's nucleolytic capability, and this nuclease-null or "dead" Cas9 (dCas9) variant can then be fused to effector domains such as transcriptional activators or repressors. These fusion proteins retain the ability to recognize and bind to DNA; when complexed with a gRNA that directs them to the promoter of a gene of interest, they have been shown to dramatically alter levels of expression.

Cas9's ability to alternately serve both as a means of modifying DNA and of modulating gene expression makes it an invaluable tool in the interrogation of gene function. However, there are limitations to a system that relies on two distinct Cas9 variants—nuclease-competent or nuclease-null—in order to effect different perturbations. Consider, for instance, the challenge of attempting to simultaneously and selectively induce both cutting and regulation of expression at different genetic loci within a single cell. One might imagine transfecting cells with both the nuclease-positive and nuclease-null variants of Cas9 along with the necessary gRNAs, but it would not be possible to control which gRNA becomes complexed with which variant of Cas9. To circumvent this problem, previous studies have proposed the concurrent

Truncated gRNAs for Regulating Gene Expression (CONT'D)

use of "orthogonal" Cas9 proteins derived from different bacterial species, each of which interacts with a distinct gRNA that allows users to decide which Cas9 protein is directed to which target site (1). Although this is possible, this strategy suffers from several limitations. Perhaps most importantly, most Cas9 orthologs are less well characterized than the conventionally used SpCas9 (derived from the bacterium *Streptococcus pyogenes*). Consequently, fewer genetic tools for targeted transcriptional and epigenetic regulation have been validated with these proteins. Furthermore, among those Cas9 orthologs that have been scrutinized, the majority have been shown to exhibit more limited activity than SpCas9 (2). In many cases this is due both to lower relative nuclease efficiencies as well as to more stringent targeting rules that result in a decrease in available target sites.

Regulating gene expression with truncated gRNAs

An alternative approach is to modulate Cas9's nuclease activity by modifying the gRNA with which it is complexed rather than the protein itself. When native Cas9 is complexed with a gRNA that has been truncated such that it exhibits 15 or fewer nucleotides of complementarity to a target site, Cas9's DNA binding capability remains intact while its nucleolytic activity is eliminated (Fig. 1) (1, 3). The gRNA can be further modified by incorporation of an RNA hairpin, such as the MS2 hairpin, that is capable of recruiting additional effector domains (4). Taken together, these minor gRNA alterations can be exploited to quickly and inexpensively generate a potent Cas9-based transcriptional regulator without making any functional changes to Cas9. For instance, when native (nuclease-competent) Cas9 interacts with a ≤15nt gRNA that contains an RNA adapter capable of recruiting a transcriptional activator, robust regulation of gene expression is observed in the absence of genome editing (Fig. 2 and 3) (5, 3).

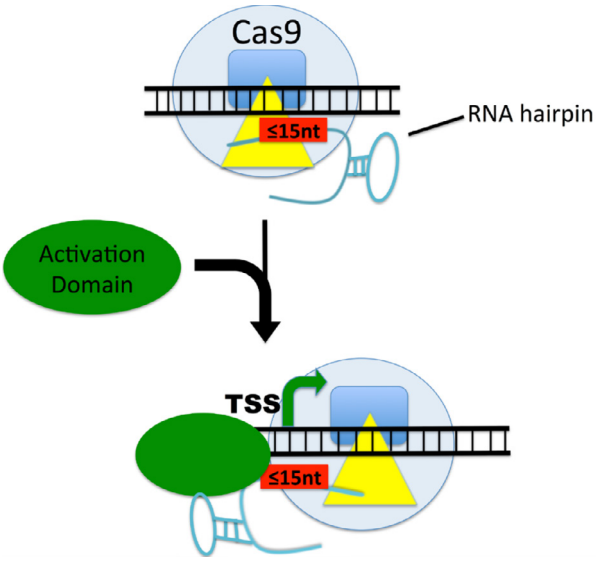


Figure 2: Gene activation with nuclease-competent Cas9. An RNA hairpin incorporated into the truncated gRNA recruits a transcriptional activator to the Cas9-gRNA complex. When this complex is directed to a target upstream of the transcriptional start site (TSS) of a gene of interest, it induces potent gene expression with no genetic alterations observed at the target stite.

Benefits of truncated gRNAs

Importantly, the use of truncated guides has largely been demonstrated to result in decreased mismatch tolerance and, consequently, increased specificity relative to the more commonly employed 20nt gRNAs. It should be noted, however, that on rare occasions truncated gRNAs (≤15nt) have been found to retain some ability to induce Cas9 to make edits (AC, WLC, and JQ, unpublished results). An

Truncated gRNAs for Regulating Gene Expression (CONT'D)

additional benefit of the use of truncated guides is that they can be delivered into systems that already express nuclease-competent Cas9, obviating the need to generate new cell lines or transgenic animals expressing dCas9 as a means of modulating gene expression (6).

Altering the length of gRNAs targeting different sites of interest thus represents a straightforward means of exerting tight control over Cas9 nuclease activity while eliminating the reliance on orthogonal Cas9 species. Within a single cell, full-length guides targeting one set of genetic loci can be introduced together with truncated gRNAs targeting a different set of loci to induce simultaneous cutting and transcriptional activation or repression at the respective sites.

References

1. Esvelt, Kevin M., et al. "Orthogonal Cas9 proteins for RNA-guided gene regulation and editing." Nature methods 10.11 (2013): 1116-1121. PubMed PMID: 24076762. PubMed Central PMCID: PMC3844869.

2. Chari, Raj, et al. "Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach." Nature methods 12.9 (2015): 823-826. PubMed PMID: 26167643.

3. Kiani, Samira, et al. "Cas9 gRNA engineering for genome editing, activation and repression." Nature methods 12.11 (2015): 1051-1054. PubMed PMID: 26344044. PubMed Central PMCID: PMC4666719.

4. Mali, Prashant, et al. "RNA-guided human genome engineering via Cas9." Science 339.6121 (2013): 823-826. PubMed PMID: 23287722. PubMed Central PMCID: PMC3712628.

5. Dahlman, James E., et al. "Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease." Nature biotechnology (2015). PubMed PMID: 26436575. PubMed Central PMCID: PMC4747789.

6. Platt, Randall J., et al. "CRISPR-Cas9 knockin mice for genome editing and cancer modeling." Cell 159.2 (2014): 440-455. PubMed PMID: 25263330. PubMed Central PMCID: PMC4265475.

Epigenetics and Editing the Epigenome

By Mary Gearing | Feb 14, 2017, Updated Jun 24, 2020 by Leah Schwiesow

Epigenetic modifications are an additional layer of control over gene expression that go beyond genomic sequence. Dysregulation of the epigenome (the sum of epigenetic modifications across the genome) has been implicated in disease states, and targeting the epigenome may make certain processes, like cellular reprogramming of iPSCs, more efficient. In general, epigenetic chromatin modifications are correlated with alterations in gene expression, but causality and mechanisms remain unclear. Today, targeted epigenetic modification at specific genomic loci is possible using CRISPR, and Addgene has a number of tools for this purpose.

Epigenetics began as a correlative field in which covalent modifications to DNA or histones, the proteins that help package DNA, were associated with gene expression or silencing. To alter DNA modifications, researchers used blunt tools like histone deacetylases, but targeted epigenetic modification was impossible. With the genome engineering revolution came epigenome-engineering tools - zinc finger nucleases and TALENs fused to epigenetic modifiers enabled epigenetic modifications at a user-specified locus.

Researchers showed that TALE-TET1 constructs, which fused a TALEN to the Tet1 demethylase catalytic domain, could mediate demethylation and induce transcription at CpG regions of various promoters. Other researchers additionally fused a TAL effector to LSD1 histone demethylase to demethylate enhancer regions (Mendenhall et al., 2013). By comparing gene activation when enhancers were active or silent, they could identify the target genes of previously uncharacterized enhancers. The popular TALEN-based LITE system, which uses light to regulate transcription, also includes light-regulated histone methyltransferases and deacetylases.

CRISPR and epigenetics

As with many TALEN-based technologies, the advent of CRISPR has made targeting much easier! With the help of guide RNAs, fusions between Cas proteins and the epigenetic modifier could be targeted to specific DNA sequences. Non-editing CRISPR applications direct catalytically dead dCas9 fused to a variety of epigenetic modifiers to specific loci without inducing double strand breaks. Below you'll find some of the CRISPR-based epigenetic modifiers available from Addgene. For most of these constructs, catalytically dead modifiers are also available as controls. For an up-to-date list of CRISPR epigenetic tools, check out our CRISPR Epigenetics resource.

Transcriptional activation

p300 acetyltransferase: dCas9 fused to the catalytic domain of p300 acetyltransferase increases levels of H3K27ac histone modification at specified loci. Charles Gersbach's lab has deposited mammalian expression constructs including pcDNA-dCas9-p300 Core and pcDNA3.3-Nm-dCas9-p300 Core, as well as pLV-dCas9-p300-P2A-PuroR for lentiviral expression.

Epigenetics and Editing the Epigenome (CONT'D)

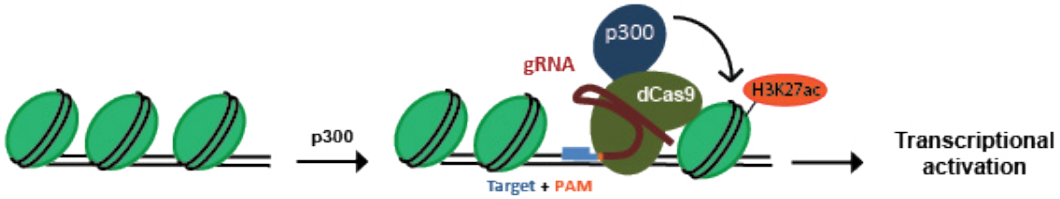


Figure 1. dCas9-p300 adds H3K27ac marks to histones in promoter and enhancer regions. These marks are associated with transcriptional activation.

Tet1 demethylase: Ronggui Hu's lab has created pdCas9-Tet1-CD for targeted cytosine demethylation in mammalian cells. This plasmid is used with pcDNA3.1-MS2-Tet1-CD to decrease methylation and activate transcription. A lentiviral vector with the same modifier, Fuw-dCas9-Tet1CD, is available from Rudolf Jaenisch's lab in plasmid form or as ready-to-use lentivirus. Tet1 initiates cytosine demethylation of DNA. However, several proteins in the DNA oxidation and repair pathways work downstream of Tet1 to restore the DNA after cytosine removal. Recently, Albert Cheng's lab developed Casilio-ME, which is based on their Casilio system. This system allows for targeted delivery of Tet1 alone, or coupled with DNA oxidation and repair factors that allow for increased gene activation at the targeted site compared to other Tet1 delivery systems.

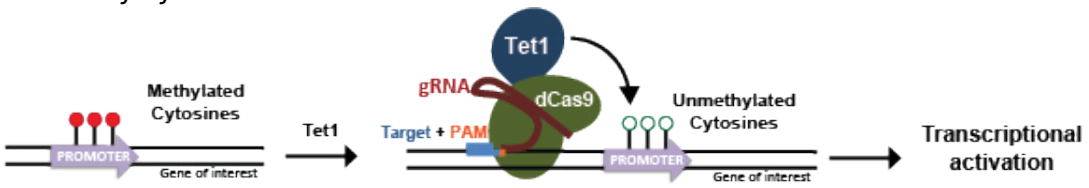


Figure 2. dCas9-Tet1 demethylates cytosines at promoters and enhancers. This targeted demethylation is associated with increased transcription.

Transcriptional repression

DNA Methyltransferase 3 Alpha (DNMT3A): Vlatka Zoldoš' lab has deposited pdCas9-DNMT3A-EGFP and pdCas9-DNMT3A-PuroR for targeted cytosine methylation in mammalian cells. Co-expression markers EGFP and PuroR enable sorting and selection of transduced cells. Grant Challen's lab also created constitutive (pCMV-dCas9-D3A) and Tet-dependent (TetO-dCas9-D3A) constructs. For lentiviral expression, Fuw-dCas9-Dnmt3a and Fuw-dCas9-Dnmt3a-P2A-tagBFP are available from Rudolf Jaenisch's lab, with the former also available as ready-to-use lentivirus.

DNA Methyltransferase MQ1: Margaret Goodell's lab has deposited pcDNA3.1-dCas9-MQ1(Q147L)-EGFP, a fusion of dCas9 to a small DNA methyltransferase from the prokaryote *Mollicutes spiroplasma* (*M. Sss1*) (termed MQ1.) The Q147L mutation improves methylation kinetics such that cytosine

Epigenetics and Editing the Epigenome (CONT'D)

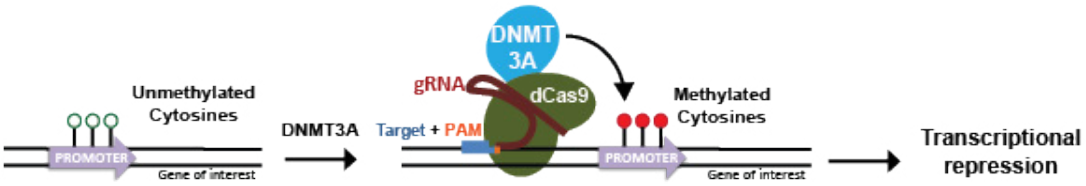


Figure 3. dCas9-DNMT3A methylates cytosines at promoters and enhancers. This targeted methylation is associated with transcriptional repression.

methylation occurs within 24 hours rather than over a period of several days, as seen with other epigenome-editing tools. pLV hUbC-dCas9-MQ1(Q147L)-EGFP is also available for lentiviral transduction.

Lysine-specific Demethylase 1 (LSD1): Tatjana Sauka-Spengler's lab has deposited pX330a dCas9-LSD1 for targeted removal of H3K4me1/2 and H3K9me2 histone modifications. Like the TALE-LSD1 system described above, dCas9-LSD1 inactivates targeted enhancers. Although the lab used this vector in chick embryos, it also functions in mammalian expression systems. Christopher Newgard's Lab has fused LSD1 to the smaller nuclease-dead *Staphylococcus aureus* Cas9 (Sa-Cas) and deposited plasmid IF311: pMAGIC (R4-R3) NLS-Sa dCas9-NLS-LSD1. LSD1 is a particularly large protein, and this fusion allows for packaging into viral vectors and targeted delivery into hard-to-manipulate cell types. They have also fused LSD1 to x-Cas9(3.7). x-Cas9(3.7) is a SpCas9 mutant that exhibits increased PAM flexibility, and this fusion protein allows for LSD1 targeting to a wider set of genomic targets. These two vectors are part of the larger pMVP/pMAGIC cloning system, which contains Cas9 fusions for several epigenetic modifiers, and is available as a kit from Addgene.

Why use epigenetic modifiers?

Epigenetic modification is certainly not the only CRISPR-based technology designed to alter gene expression. Fusing dCas9 to a transcriptional activator like VP64 or VPR activates transcription, whereas dCas9-KRAB fusions repress transcription. Both of these methods also recruit epigenetic machinery - but is there an advantage to using direct epigenetic modifiers?

As with any experiment, your desired outcome will determine the tool that you should use. If you want to study the effects of one particular modification for which a targeted editor, like H3K27ac, is available, an epigenetic tool would be your best bet.

Another potential advantage of CRISPR epigenetic tools is their persistence and inheritance. CRISPR activators and repressors are thought to be reversible once the effector is inactivated/removed from the system. In contrast, epigenetic marks left by targeted epigenetic modifiers may be more frequently inherited by daughter cells. For example, KRAB induced silencing was transient and quickly reversed in

Epigenetics and Editing the Epigenome (CONT'D)

culture. However, DNMT3A-induced methylation persisted throughout a 100 day experimental period, as this mark was faithfully propagated in culture and *in vivo*.

In certain cases, epigenetic modifiers may work better than activators/repressors - dCas9-p300 increased transcriptional activation more than dCas9-VP64, especially when targeting distal enhancers. As the effects of these tools are likely cell type- and context-dependent, it may make sense to try multiple CRISPR tools when setting up your experimental system.

References

Haldeman JM, Conway AE, Arlotto ME, Slentz DH, Muoio DM, Becker TC, Newgard CB (2018) Creation of versatile cloning platforms for transgene expression and dCas9-based epigenome editing. Nucleic Acids Research 47:e23–e23 . https://doi.org/10.1093/nar/gky1286

Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA (2015) Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nature Biotechnology 33:510–517 . https://doi.org/10.1038/nbt.3199

Klann TS, Black JB, Chellappan M, Safi A, Song L, Hilton IB, Crawford GE, Reddy TE, Gersbach CA (2017) CRISPR–Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. Nature Biotechnology 35:561–568 . https://doi.org/10.1038/nbt.3853

Lei Y, Zhang X, Su J, Jeong M, Gundry MC, Huang Y-H, Zhou Y, Li W, Goodell MA (2017) Targeted DNA methylation *in vivo* using an engineered dCas9-MQ1 fusion protein. Nature Communications 8: 16026. https://doi.org/10.1038/ncomms16026

Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, Shu J, Dadon D, Young RA, Jaenisch R (2016) Editing DNA Methylation in the Mammalian Genome. Cell 167:233–247.e17 . https://doi.org/10.1016/j.cell.2016.08.056

Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, Ho QH, Sander JD, Reyon D, Bernstein BE, Costello JF, Wilkinson MF, Joung JK (2013) Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nature Biotechnology 31:1137–1142 . https://doi.org/10.1038/nbt.2726

McDonald JI, Celik H, Rois LE, Fishberger G, Fowler T, Rees R, Kramer A, Martens A, Edwards JR, Challen GA (2016) Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. Biology Open 5:866–874 . https://doi.org/10.1242/bio.019067

Epigenetics and Editing the Epigenome (CONT'D)

Mendenhall EM, Williamson KE, Reyon D, Zou JY, Ram O, Joung JK, Bernstein BE (2013) Locus-specific editing of histone modifications at endogenous enhancers. Nature Biotechnology 31:1133–1136 . https://doi.org/10.1038/nbt.2701

Stolzenburg S, Beltran AS, Swift-Scanlan T, Rivenbark AG, Rashwan R, Blancafort P (2015) Stable oncogenic silencing *in vivo* by programmable and targeted de novo DNA methylation in breast cancer. Oncogene 34:5427–5435 . https://doi.org/10.1038/onc.2014.470

Taghbalout A, Du M, Jillette N, et al (2019) Enhanced CRISPR-based DNA demethylation by Casilio-ME-mediated RNA-guided coupling of methylcytosine oxidation and DNA repair pathways. Nature Communications 10: 4296. https://doi.org/10.1038/s41467-019-12339-7

Vojta A, Dobrinić P, Tadić V, Bočkor L, Korać P, Julg B, Klasić M, Zoldoš V (2016) Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic Acids Research 44:5615–5628 . https://doi.org/10.1093/nar/gkw159

Williams RM, Senanayake U, Artibani M, Taylor G, Wells D, Ahmed AA, Sauka-Spengler T (2017). Genome and epigenome engineering CRISPR toolkit for probing *in vivo* cis-regulatory interactions in the chicken embryo. BioRxiv: https://doi.org/10.1101/135525.

Xu X, Tao Y, Gao X, Zhang L, Li X, Zou W, Ruan K, Wang F, Xu G, Hu R (2016) A CRISPR-based approach for targeted DNA demethylation. Cell Discovery 2: 16009. https://doi.org/10.1038/celldisc.2016.9

Chapter 7

Guide RNAs

How to Design Your gRNA for CRISPR Genome Editing

By John Doench | May 3, 2017, Updated Sep 24, 2020

CRISPR technology has made it easier than ever both to engineer specific DNA edits and to perform functional screens to identify genes involved in a phenotype of interest. This article will discuss differences between these approaches, and provide updates on how best to design gRNAs. You can also find validated gRNAs for your next experiment in Addgene's Validated gRNA Sequence Datatable. A more extended discussion of these subjects can be found in two recent review articles (Doench et al., 2017, and Hanna et al., 2020) and references therein.

Important considerations before you start an experiment with CRISPR

The hammer, the jigsaw, and the wrench are all great tools, but which one you use, of course, depends on what you are trying to do – there's no "best" tool among them. While this seems obvious, it is important to remember that the same is true when designing gRNAs for using CRISPR technology – the "best" gRNA depends an awful lot on what you are trying to do: gene knockout, a specific base edit, or modulation of gene expression.

Location and sequence are important considerations for designing your gRNAs. For indels, it's not so important what location in the gene you target, but it is important that your gRNA sequence is designed to be highly active and reduce off targets. For CRISPRa and CRISPRi, these considerations are of roughly equal importance (target should be near the TSS but you can worry less about sequence optimality because you generally have fewer sequences to choose from). Finally, for HDR, location is much more important because you have to target within ~30 nt of your proposed edit, which means there are so few gRNAs to choose from that sequence preferences must largely be ignored.

The hammer: Gene knockout by NHEJ

Gene knockout with CRISPR technology is usually accomplished by Cas9-mediated dsDNA breaks: following a cut, the error-prone nature of non-homologous end joining (NHEJ) often leads to the generation of indels and thus frameshifts that disrupt the protein-coding capacity of a locus. When using *S. pyogenes* Cas9 (SpCas9), potential target sites are both [5'-20nt-NGG] and [5'-CCN-20nt], as it is equally efficacious to target the coding or non-coding strand of DNA. As a rule of thumb, we avoid target sites that code for amino acids near the N' terminus of the protein, in order to mitigate the ability of the cell to use an alternative ATG downstream of the annotated start codon. Likewise, we avoid target sites that code for amino acids close to the C' terminus of the protein, to maximize the chances of creating a non-functional allele. For a 1 kilobase gene, since potential target sites occur ~1 in every 8 nucleotides, restricting gRNAs to 5 – 65% of the protein coding region will still result in many dozens of gRNAs to choose from. With so many possibilities, picking a gRNA with an optimized sequence is of primary importance (more on this below).

How to Design Your gRNA for CRISPR Genome Editing (CONT'D)

The jigsaw: Editing by HDR, base editing, and prime editing

For a specific edit, such as the insertion of a fluorescent tag or the introduction of a specific mutation, one generally relies on homology directed repair (HDR) to incorporate new information into DNA. This also requires an exogenous DNA template. HDR, however, is a very low-efficiency process, and usually involves the need for single cell cloning and subsequent screening for successful edits. This is a very time consuming process and should not be undertaken lightly! Indeed, truly achieving the gold standard requires not one but two rounds of single cell cloning – as a control, one should revert the edit back to the original in order to prove that the phenotype was really due to the intended edit rather than some passenger variant that came along with the single cell clone (although this is rarely done).

When targeting a dsDNA break for HDR, the choice of target site is far more constrained by the desired location of edit; efficiency decreases dramatically when the cut site is >30nt from the proximal ends of the repair template (Yang et al., 2013). This means that, for gene editing, there are usually few potential gRNAs. While SpCas9, with a PAM preference of NGG, is still the most widely-used Cas enzyme, the development of SaCas9, NmeCas9, Cas12a enzymes, and engineered variants thereof offers additional PAM options that can greatly expand gRNA options.

Two newer technologies offer an alternative to HDR for introducing edits. The same locational constraints are even more exquisite for the so-called base editor Cas9, which makes DNA changes in the absence of dsDNA breaks (Rees et al., 2018). For C>T and A>G base editors, the intended edit must be in a 5 - 10 nt window relative to the PAM, and bystander edits are possible if there is another target C or A in the window. Another technology, prime editing (reviewed in Anzalone et al., 2020). is not limited to single nucleotide transitions but still requires a nearby PAM, although these are still early days for this technology, and the user may need to optimize numerous parameters to generate the desired edit.

The wrench: Gene activation and inhibition by CRISPRa and CRISPRi

Finally, for modulating gene expression at the level of transcription – CRISPRa (activation) and CRISPRi (inhibition) technologies – a nuclease-dead Cas9 (dCas9) is directed near the promoter of a target gene. Here, the target window is not quite as broad as for knockout via CRISPR cutting. For CRISPRa, it is most-efficacious to target a ~100nt window upstream of the transcription start site (TSS), while for CRISPRi, a ~100nt window downstream of the TSS gives the most activity. Thus, a given gene will only have a dozen or so gRNAs to choose from in the optimal location. It is also important to have good information on the exact location of the TSS. Different databases annotate the TSS in different ways, and it has been shown that the FANTOM database, which relies on CAGE-seq to directly capture the mRNA cap, provides the most accurate mapping (Radzisheuskaya et al., 2016). In this case, location

How to Design Your gRNA for CRISPR Genome Editing (CONT'D)

and sequence are of approximately equal importance in design – an optimized sequence will do little if it is in the wrong place, but because the target window is more-narrow, there are fewer gRNA to choose from, and thus an optimal sequence may not be available.

Predicting gRNA efficacy

We and others have examined the ability to use sequence-based and other features to nominate gRNAs that are likely to be active, not only for SpCas9 but also for some other Cas enzymes. It seems to be the case that there is no universal scoring system for selecting a gRNA, as the method of producing the guide (synthetic, *in vitro* transcription, or lentiviral delivery) can affect the accuracy of a predictive score, as well as dynamic aspects of the target (e.g. accessibility due to chromatin status). No computational prediction is ever perfect, but this can decrease the number of guides one needs to test in the lab.

Importantly, for any modification of interest, it would be unwise to make conclusions on the basis of the activity of a single gRNA, and thus diversity of gRNAs across a gene should be examined whenever possible when using knockout or transcriptional modulation approaches.

Avoiding off-target effects

The off-target activity of gRNAs is important to consider. While the basic landscape of mismatches that can nevertheless still lead to activity has been established, and can be used to identify sites that are likely to give rise to an off-target effect, there's not enough data to fully predict which sites will and will not show appreciable levels of modification. Whole-genome sequencing of cells modified by CRISPR indicates that the consequences of off-target activity, at least for the experimental conditions used, led to no detectable mutations (Veres et al., 2014). When working with single-cell clones, the authors note that "clonal heterogeneity may represent a more serious obstacle to the generation of truly isogenic cell lines than nuclease-mediated off-target effects." Further, large-scale datasets of hundreds of genetic screens using genome-wide libraries have shown high concordance between different sequences targeting the same gene, suggesting that off-target effects did not overwhelm true signal in these assays (Dempster et al., 2019). Again, the experimental strategy is clear: for any gene of interest, one should require that multiple gRNAs of different sequences give rise to the same phenotype in order to conclude that the phenotype is due to an on-target effect.

Conclusions

Selection of gRNAs for an experiment needs to balance maximizing on-target activity while minimizing off-target activity, which sounds obvious but can often require difficult decisions. For example, would

How to Design Your gRNA for CRISPR Genome Editing (CONT'D)

it be better to use a less-active gRNA that targets a truly unique site in the genome, or a more-active gRNA with one additional target site in a region of the genome with no known function? For the creation of stable cell models that are to be used for long-term study, the former may be the better choice. For a genome-wide library to conduct genetic screens, however, a library composed of the latter would likely be more effective, so long as care is taken in the interpretation of results by requiring multiple sequences targeting a gene to score in order to call that gene a hit.

This is an exciting time for functional genomics, with an ever-expanding list of tools to probe gene function. The best tools are only as good as the person using them, and the proper use of CRISPR technology will always depend on careful experimental design, execution, and analysis.

References

Anzalone AV, Koblan LW, Liu DR (2020) Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. Nat Biotechnol 38:824–844 . https://doi.org/10.1038/s41587-020-0561-9

Dempster JM, Pacini C, Pantel S, Behan FM, Green T, Krill-Burger J, Beaver CM, Younger ST, Zhivich V, Najgebauer H, Allen F, Gonçalves E, Shepherd R, Doench JG, Yusa K, Vazquez F, Parts L, Boehm JS, Golub TR, Hahn WC, Root DE, Garnett MJ, Tsherniak A, Iorio F (2019) Agreement between two large pan-cancer CRISPR-Cas9 gene dependency data sets. Nat Commun 10: . https://doi.org/10.1038/s41467-019-13805-y

Doench JG (2017) Am I ready for CRISPR? A user's guide to genetic screens. Nat Rev Genet 19:67–80 . https://doi.org/10.1038/nrg.2017.97

Hanna RE, Doench JG (2020) Design and analysis of CRISPR–Cas experiments. Nat Biotechnol 38:813–823 . https://doi.org/10.1038/s41587-020-0490-7

Radzisheuskaya A, Shlyueva D, Müller I, Helin K (2016) Optimizing sgRNA position markedly improves the efficiency of CRISPR/dCas9-mediated transcriptional repression. Nucleic Acids Res 44:e141–e141 . https://doi.org/10.1093/nar/gkw583

Rees HA, Liu DR (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. Nat Rev Genet 19:770–788 . https://doi.org/10.1038/s41576-018-0059-1

Veres A, Gosis BS, Ding Q, Collins R, Ragavendran A, Brand H, Erdin S, Cowan CA, Talkowski ME, Musunuru K (2014) Low Incidence of Off-Target Mutations in Individual CRISPR-Cas9 and TALEN Targeted Human Stem Cell Clones Detected by Whole-Genome Sequencing. Cell Stem Cell 15:27–30 .

How to Design Your gRNA for CRISPR Genome Editing (CONT'D)

https://doi.org/10.1016/j.stem.2014.04.020

Yang L, Guell M, Byrne S, Yang JL, De Los Angeles A, Mali P, Aach J, Kim-Kiselak C, Briggs AW, Rios X, Huang P-Y, Daley G, Church G (2013) Optimization of scarless human stem cell genome editing. Nucleic Acids Research 41:9049–9061 . https://doi.org/10.1093/nar/gkt555

Multiplex Expression of gRNAs

By Mary Gearing | Jan 28, 2016, Updated Sep 10, 2020 by Jennifer Tsang

CRISPR makes it easy to target multiple loci - a concept called multiplexing. Since CRISPR is such a robust system, editing or labeling efficiency doesn't usually change when you add multiple gRNAs on one plasmid. Sound good? Addgene has many tools to help you multiplex - we'll use mammalian plasmids to introduce you to some of your potential options and cloning methods, but please scroll down for plasmids suitable for other model systems, including *E. coli*, plants, Drosophila, and zebrafish!

Why use multiplexed gRNAs?

By expressing multiple gRNAs on the same plasmid, you'll make sure that each cell that gets the plasmid contains all of the desired gRNAs. This increases the chance that all the edits you want to make will happen.

There are many reasons to multiplex gRNAs, some being:

- Using dual nickases to generate a knockout or edit. This can help reduce off-target activity.
- Deleting a large region of the genome by removing the sequence between two target sites.
- Modifying multiple genes at once. Using multiplexed gRNAs can target multiple locations in the genome to modify, whether it be editing CRISPRi, CRISPRa, or base editing.

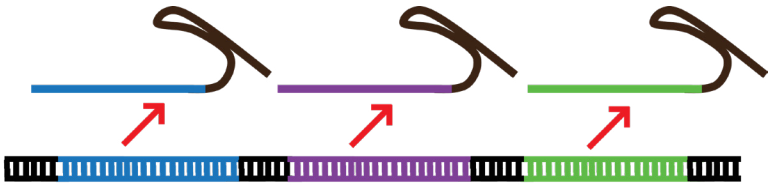


Figure 1: Multiplexing allows researchers to express multiple gRNA from a single construct. DNA and RNA are not to scale.

Multiplexing gRNAs: the basics

One common question Addgene Senior Scientists receive is: can I express more than one gRNA from a single promoter using a plasmid like pX330? Unfortunately, the short answer is no. Unless you use a system for processing a continuous multi-gRNA transcript, each gRNA must be expressed from its own promoter. But that doesn't mean you have to clone and transfect multiple promoter-gRNA constructs in order to target multiple sites! In this post, we'll cover Cas9 multiplexing options, but also check out our blog post about multiplexing with Cpf1.

Let's start with the simplest multiplexing situation: you only need to express two gRNAs at the same time. One system you could use is pX333 from the Ventura lab. pX333, a modification of pX330, contains humanized wtCas9 and two U6 promoters. To use this plasmid, you simply order oligonucleotides for your chosen gRNA target sequences and clone them in just as you would for a single gRNA. You'll clone in the first gRNA using restriction enzyme BbsI and the second gRNA using

Multiplex Expression of gRNAs (CONT'D)

restriction enzyme BsaI. If you're working in Drosophila, a two-gRNA expressing plasmid is available from the Bullock lab, and gRNAs can be inserted using Gibson Assembly or SLIC cloning methods. A BsaI-based *E. coli* multiplexing plasmid is available from the Koffas lab.

Scaling up your multiplexed gRNAs with Golden Gate and Gibson assembly methods

If you want to scale up the number of gRNAs in your plasmid, you'll need to use some assembly methods such as Golden Gate or Gibson assembly.

Golden Gate Assembly of gRNAs

Golden Gate Assembly uses Type IIS restriction enzymes, which cleave outside of their recognition sequence, creating flanking overhangs. These overhangs can be customized to link together multiple fragments, allowing ordered assembly of multiple components into a destination vector. This is illustrated in the schematic.

The first step in CRISPR/Cas9 Golden Gate multiplexing is to clone the oligonucleotides specifying each gRNA target sequence into distinct expression vectors using the enzyme BbsI. These expression vectors each contain Type IIS restriction sites flanking the promoter-gRNA construct, but with different sequences adjacent to the sites. When digested with the appropriate Type IIS enzyme, the unique flanking overhang sequences can link together to allow for ordered assembly into a destination vector that expresses Cas9.

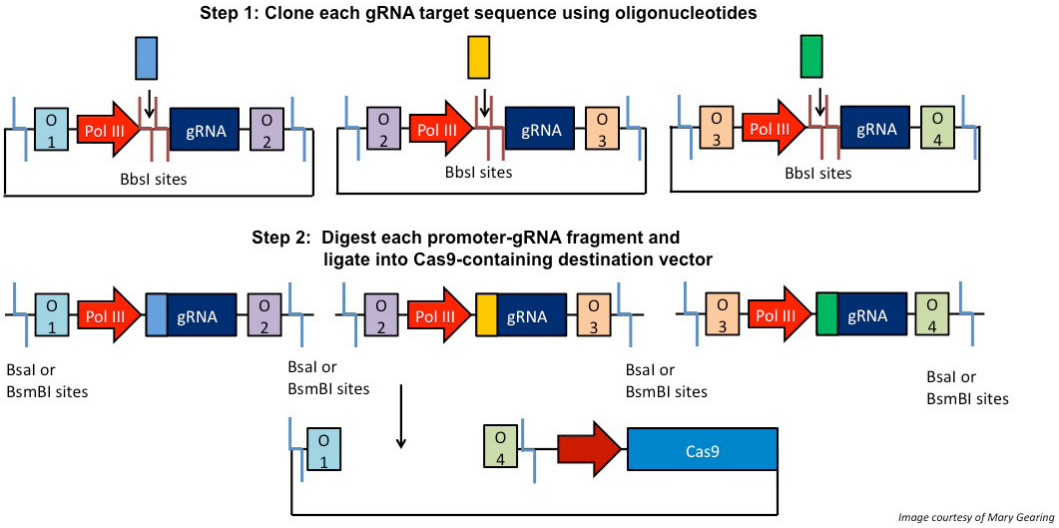


Figure 2: gRNA target sequences are cloned into various plasmids using oligonucleotides. These plasmids contain Type IIS restriction sites that flank the promoter-gRNA constructs. When these plasmids are digested, unique overhangs (here, O1-4) adjacent to the cut sites "link" fragments together and drive ordered assembly into a Cas9-containing destination vector. Note: depending on which method you use, the procedure will vary slightly.

Multiplex Expression of gRNAs (CONT'D)

Two Golden Gate options that are available from Addgene follow these same assembly principles, but they're optimized for different purposes.

Gersbach Lab multiplexing plasmids: This plasmid set allows you to express 2-4 gRNAs, with four being the ideal number. First you generate four unique kanamycin-resistant plasmids, each containing a different gRNA target sequence downstream of the 7SK, human U6, mouse U6, or human H1 promoters. If you express fewer than four gRNAs, you'll clone in a polyT-termination sequence for each unused promoter. This step is necessary to generate all of the overhangs needed for the final ligation step. Plasmids are then digested using BsmBI and ligated into Cas9 or dCas9-containing destination vectors. Destination vector options include humanized wt Cas9, dCas9 (transcriptional repressor), and dCas9-VP64 (transcriptional activator)-containing plasmids. Each destination vector contains GFP, enabling you to select cells with high GFP expression. These cells have the highest levels of Cas9 and gRNA expression, and thus the highest frequency of genome editing events.

Yamamoto Lab Multiplex CRISPR/Cas9 Assembly Kit: This kit is built for serious multiplexing and enables users to express up to 7 gRNAs. The kit contains different destination vectors depending on the total number of gRNAs you wish to clone, from 2-7. For example, if you're expressing 4 gRNAs, you'd use pX330A-1x4; for 6 gRNAs, you'd use pX330A-1x6. This customization means you don't ever need to clone in filler sequences. To build your multiplexing construct, you clone all but one of your gRNAs into spectinomycin-resistant plasmids pX330S-2 to pX330S-(last gRNA number). The 5' most gRNA is cloned into the Cas9-containing destination vector. These constructs are digested using BsaI and assembled to produce a plasmid encoding the gRNAs and Cas9. As with the Gersbach lab plasmids, multiple Cas9 variants are available: wt humanized Cas9, D10A nickase mutant (Cas9n), dCas9 (transcriptional repression), and Fok1-dCas9 (dimeric nuclease).

Gateway assembly method

Frew Lab Multiple Lentiviral Expression Systems (MuLE) Kit: This kit can be used to create lentiviral vectors expressing wt humanized Cas9 and up to three gRNAs. Entry vectors containing the U6 promoter and the gRNA scaffold are provided with the kit. Oligonucleotides specifying the gRNA seed sequence should be compatible with type IIS enzyme BfuAI. Gateway cloning is then used to combine the multiple gRNAs and Cas9 together into a single plasmid. Although only wt hCas9 entry vectors are supplied with the kit, you can clone your own entry vectors containing other Cas9 variants to use with the MuLE system.

Multiplex Expression of gRNAs (CONT'D)

Multiplexing from a single transcript

You can also multiplex gRNAs via a polycistronic transcript. Rather than being transcribed from different promoters, the gRNAs are transcribed together and are flanked by specific sites that allow them to be cleaved and released. These constructs tend to be smaller than constructs with multiple promoter-gRNA cassettes, making them advantageous for small capacity vectors like AAV. In addition to the mammalian option described below, plasmids for making polycistronic gRNAs are also available from the Yang lab for use in plants.

The mammalian multiplex systems use the Csy4 RNA nuclease from *Pseudomonas aeruginosa*. When overexpressed, Csy4 efficiently cleaves gRNAs sandwiched between 28 base Csy4 recognition

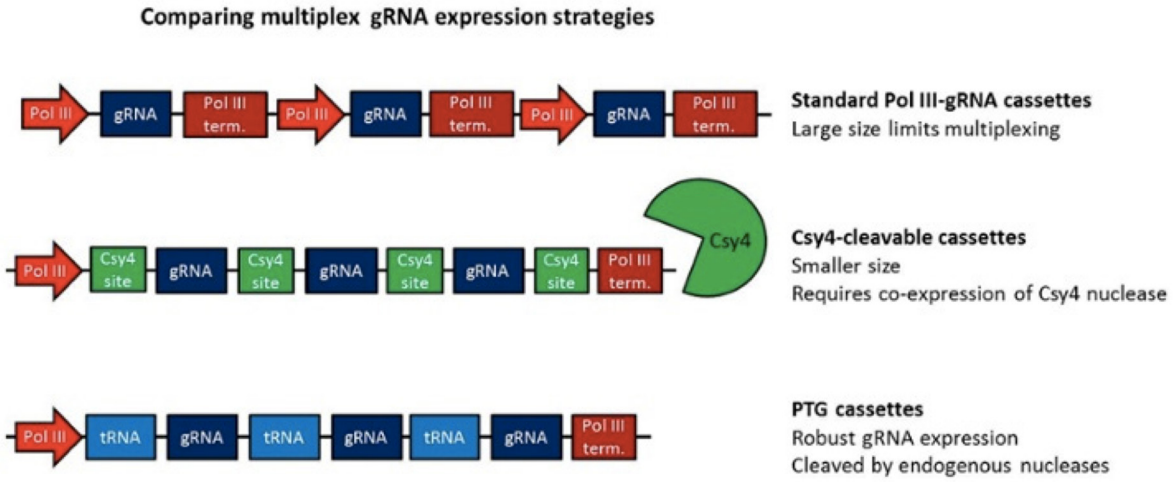


Figure 3: Comparison of Multiplex Strategies including Standard PolIII-gRNA cassette, Csy4-cleavable cassette, and PTG cassette.

sites. If Csy4 is not expressed, the gRNAs cannot be released, adding temporal and/or spatial control to the system. pSQT1313 from the Joung lab allows you to express two gRNAs constructed using oligonucleotide assembly. Unlike some of the plasmids described above, this vector does not contain Cas9, so you'll need to supply it with another plasmid.
Check out Addgene's vectors for multiplex gRNAs!

References

Albers, Joachim, et al. "A versatile modular vector system for rapid combinatorial mammalian genetics." Journal of Clinical Investigation 125(4) (2015):1603-19. PubMed PMID: 25751063. PubMed Central

Multiplex Expression of gRNAs (CONT'D)

PMCID: PMC4396471.

Cress, Brady F., et al. "CRISPathBrick: Modular Combinatorial Assembly of Type II-A CRISPR Arrays for dCas9-Mediated Multiplex Transcriptional Repression in *E. coli*." ACS Synthetic Biology 4(9) (2015):987-1000. PubMed PMID: 25822415.

Kabadi, Ami M., et al. "Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector." Nucleic Acids Research 42(19) (2014): e147. PubMed PMID: 25122746. PubMed Central PMCID: PMC4231726.

Maddalo, Danilo, et al. "In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system." Nature 516(7531) (2014): 423-7. PubMed PMID: 25337876. PubMed Central PMCID: PMC4270925.

Port, Fillip, et al. "Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila." Proceedings of the National Academy of Sciences USA 111(29) (2014): E2967-2976. PubMed PMID: 25002478. PubMed Central PMCID: PMC4115528.

Sakuma, Tetsushi, et al. "Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system." Scientific Reports 4 (2014): 5400. PubMed PMID: 24954249. PubMed Central PMCID: PMC4066266.

Tsai, Shengdar Q., et al. "Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing." Nature Biotechnology 32(6) (2014): 569-576. PubMed PMID: 24770325. PubMed Central PMCID: PMC4090141.

Yin, Linlin, et al. "Multiplex Conditional Mutagenesis Using Transgenic Expression of Cas9 and sgRNAs." Genetics 200(2) (2015): 431-441. PubMed PMID: 25855067.

Chapter 8

Genome Wide Screening

Genome-wide Screening Using CRISPR

By Joel McDade | Aug 18, 2015, Updated Aug 20, 2020 by Alyssa Cecchetelli

What genes are important in your phenotype of interest? Many scientists study diseases for which the underlying genetic cause is not entirely known. Identifying which genes are important for a phenotype can lead to a wealth of additional experiments investigating the role of individual genes or entire pathways in a particular disease process. While CRISPR is certainly not the first means to carry out so-called "forward genetic screening experiments", it is certainly the most robust. In this article, we will discuss how CRISPR libraries are being used to perform genome-wide screens and highlight some of the reagents that have been made publicly available through Addgene.

What makes CRISPR so special?

A major advantage of CRISPR/Cas9 over previous genome editing techniques is its simplicity and versatility. CRISPR/Cas9 consists of two components: the non-specific endonuclease (Cas9) and a single stranded guide RNA (gRNA). The ~20 nucleotide targeting sequence within the gRNA is defined by the user, and it can be easily modified to target Cas9 to virtually any genomic locus, provided the target is unique compared to the rest of the genome and located immediately 5' to a protospacer adjacent motif (PAM) sequence. Co-delivery of wild-type Cas9 and a gRNA generates a double-strand break in the target DNA, which, when repaired through error-prone non-homologous end joining (NHEJ), usually results in a loss-of-function mutation within the target gene. CRISPR can also be used to activate or repress target genes without permanently modifying the genome. If you'd like to brush up on the various CRISPR technologies, check out Addgene's CRISPR Guide.

What CRISPR-based reagents are available for genome-wide screens?

The goal of a genome-wide screening experiment is to generate and screen a population of mutant cells to identify genes involved in a particular phenotype. CRISPR can be readily scaled up for genome-wide screening due to the broad range of potential target sequences and ease of generating gRNA-containing plasmids. CRISPR genome wide-screening experiments commonly use lentivirus to deliver a pooled population of gRNAs to Cas9 expressing cells. Pooled lentiviral CRISPR libraries (referred to simply as "CRISPR libraries") consist of a heterogeneous population of gRNA-containing lentiviral transfer vectors, each targeting a specific gene within the genome. Individual gRNAs are designed *in silico* using publicly available gRNA design software and synthesized. Pooled gRNAs are then cloned into lentiviral transfer vectors to create the CRISPR library.

CRISPR libraries have been created to knock out, activate or repress target genes by combining a gRNA library with the aforementioned derivatives of Cas9. Several CRISPR libraries are publicly available through Addgene, with more being added all the time. Keep in mind that libraries are only as good as the experiments you use them for! A well-developed biological question and experimental

Genome-wide Screening Using CRISPR (CONT'D)

system are absolutely necessary to ensure that you select the correct CRISPR library.

Choosing the CRISPR library that is right for you

There are several factors to consider when selecting a CRISPR library for your experiments.

1. What species are your cells derived from? Currently, Addgene carries CRISPR libraries that target mouse, human, fly, *E. coli*, and *T. gondii* genes.
2. What genetic modification are you trying to make? Addgene carries CRISPR libraries for gene knockout, activation, repression, and barcoding.
3. Are you trying to target every gene in the genome, or a specific class of genes? Addgene currently carries several genome-wide CRISPR libraries and a selection of sub-libraries targeting specific classes of human genes.

What are the steps involved in a CRISPR screen?

Performing a forward genetic screen using CRISPR libraries is a multi-step process (see Fig. 1). In most cases, CRISPR libraries are provided at a concentration that is too low for experimental use. Thus, the first step is to amplify your library to a concentration that is sufficient to generate lentivirus. Be sure to use next-generation sequencing to check the quality of your amplification. If you've obtained a ready-to-use lentiviral preparation from Addgene, you can skip the steps above!

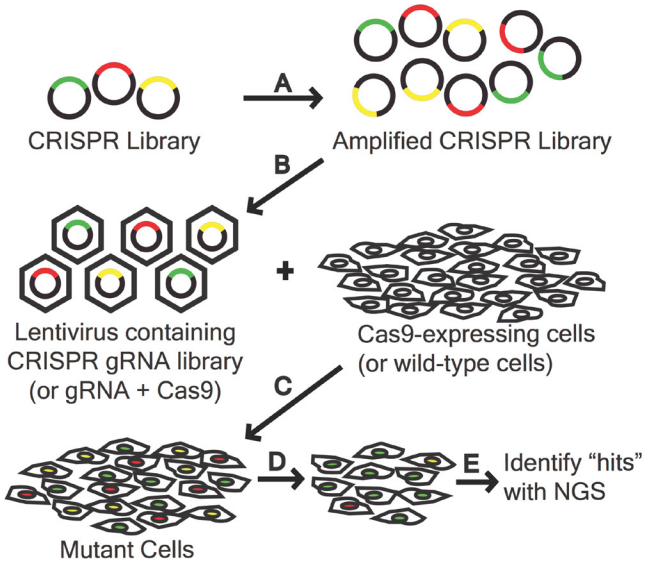


Figure 1: The CRISPR library must be amplified to be used to generate lentivirus. Cas9-expressing cells or wild-type cells are treated with lentivirus containing the gRNA library or gRNA library with Cas9 to generate mutant cells, respectively. Mutant cells are screened and hits are identified using NGS.

Cells are then transduced with lentivirus containing the CRISPR library to generate a heterogeneous population of mutant cells, with each cell or set of cells containing a mutation in a different gene. Libraries may be available in a 1-plasmid system, in which Cas9 is included on the gRNA-containing plasmid, or a 2-plasmid system in which Cas9 must be delivered separately.

Mutant cells are enriched using either drug selection or fluorescence-based cell sorting and screened for a particular phenotype. For example, mutant cells can be used in drug screens to identify genes that confer drug-resistance. Mutant cells are treated with

Genome-wide Screening Using CRISPR (CONT'D)

a drug of interest and gRNA distribution is analyzed in the drug-resistant population compared to a non-treated control group. In this scenario, gRNAs that are "enriched" correspond to genes that confer drug resistance when mutated. Findings from this type of experiment can shed light on the mechanism by which cells gain resistance to drugs and can identify future therapeutic targets for diseases causing uncontrolled cell growth, such as cancer.

Considerations and tips for successful screens

Next-generation sequencing - CRISPR libraries contain thousands of gRNA plasmids, discerned only by a unique barcode on each plasmid. As such, sequencing CRISPR libraries after amplification and after a screen requires the use of next-generation sequencing.

Representation - Most libraries contain 3-6 gRNAs per target gene, and maintaining the distribution of each gRNA within the population is key. Loss of representation due to enrichment or depletion of specific gRNAs can lead to skewed results.

Selecting a cell type - Theoretically, any cell type can be used in a CRISPR screen. However, maintaining sufficient representation within your mutant population requires a massive amount of cells as starting material. Therefore, cell types that are of low abundance are not particularly well suited for genome-wide screening.

Avoid false positives and false negatives – As with any experiment, the use of appropriate controls, multiple replicates and several cell types can strengthen your results. Enrichment or depletion of multiple gRNAs targeting the same gene can be strong evidence that a particular gene is actually important for a given phenotype. Each hit from the screen should be independently validated to ensure that the desired modification produces the phenotype you screened for in the first place.

With the proper experimental design and validation practices, CRISPR libraries can help you learn a lot about your phenotype of interest. For more detailed information on CRISPR screens check out Addgene's 2016 publication "Practical Considerations for Using Pooled Lentiviral CRISPR Libraries" (McDade et al., 2016). To learn a bit about how CRISPR/Cas9 can be used in other types of experiments, check out our CRISPR Plasmids and Resources Page.

References

McDade JR, Waxmonsky NC, Swanson LE, Fan M (2016) Practical Considerations for Using Pooled Lentiviral CRISPR Libraries. Current Protocols in Molecular Biology 115: . https://doi.org/10.1002/cpmb.8

Controlling for Off-target Effects with a New Genome-wide CRISPR Screen Design

By Beth Kenkel | Sep 13, 2018

Genome-wide CRISPR/Cas9 screens are a high-throughput systematic approach for identifying genes involved in a biological process. These screens provide an alternative to genome-wide RNAi screens, which although highly effective, are affected by low on-target efficacy, non-specific toxicity, and off-target effects. The flaws of RNAi screens are well characterized and strategies exist to control for these faults. However, it's still unclear if similar pitfalls exist for CRISPR screens and how best to design these screens to controls for flaws. Recently the Bassik Lab at Stanford developed a new genome-wide CRISPR knockout screen to analyze the following unanswered questions about CRISPR screen design.

Unanswered questions about genome-wide CRISPR/Cas9 screen design

- **Are non-targeting guides an appropriate control?** Many CRISPR screens use non-targeting guides as negative controls. Non-targeting guides do not target any site in the genome. This means that while they control for many of the effects of gRNA and Cas9 expression, they fail to account for the effects of Cas9-induced dsDNA breaks.
- **Are off-target cutting patterns in genome-wide screens similar to patterns seen in single-guide studies?** Many strategies exist for reducing Cas9 off-target effects, but most studies have focused on the off-target cutting of only a handful of guides. It's unclear if off-target cuts confound the results of high-throughput CRISPR screens.
- **Do short guides (17-18 bp) have less off-target cutting than full-length guides (19-20bp)?** Small-scale studies (Tsai et al., Fu et al.) suggest 17-19 bp long guides have reduced off-target cutting without a reduction of on-target activity, but it's unclear if this holds true in genome-wide screens.

Bassik lab's CRISPR knockout libraries

Overview of Library

To better define the optimal conditions for CRISPR knockout screens, the Bassik lab created a human knockout library that has 10 gRNAs per gene and targets all ~20,500 protein-coding genes. Unique, non-overlapping sites in the genome were targeted and predicted on-target activity was balanced with predicted

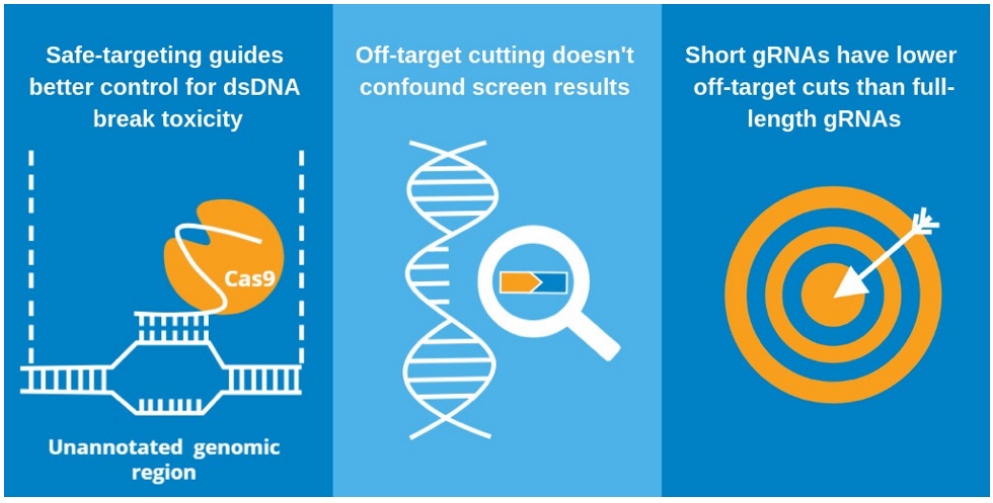


Figure 1: Overview of the Bassik lab's CRISPR library.

Controlling for Off-target Effects with a New Genome-wide CRISPR Screen Design (CONT'D)

off-target activity when selecting guides. Guides were delivered as a pooled library via a lentiviral vector to cell lines that either stably expressed Cas9 or that had been lentiviral infected with Cas9.

Validating the Library

The CRISPR knockout library was validated via a growth screen and a ricin toxicity screen since essential and non-essential genes for both pathways had previously been identified (Hart et al, Bassik et al). The library performed well in both screens, with the growth screen having a 1% false discovery rate and identifying >88% of previously identified essential genes while previous Cas9 and shRNA library screens only identified 60% of the essential genes. The ricin toxicity screen had a 10% false discovery rate and identified 67% of previously identified ricin toxicity genes. The screen also discovered several genes previously not associated with ricin regulation, including almost all genes involved in the production of a cell surface glycan that's required for ricin uptake. The ricin screen did fail to identify some known ricin-regulators, but most of these genes are also essential for growth and would not be expected to be identified in a CRISPR knockout screen.

Improving the Design of CRISPR Knockout Screens

Before diving into the results, it's important to note that in Morgens et al., guides were considered toxic if they were depleted from the screen. This toxicity was used a proxy for detecting Cas9 cutting.

Safe targeting guides better control for non-specific toxicity than non-targeting guides

First, Morgens et al. looked at the effect of safe-targeting guides vs. non-targeting guides on screen analysis and hit calling. 5,644 non-targeting guides and 6,750 safe-targeting guides were included in the library. Safe-targeting guides were designed to target genomic sites with no annotated function, i.e. sites that lack open chromatin marks, DNase hypersensitivity, or are in an enhancer, transcription factor binding site, or a coding region. Targeting these sites should control for the effects of guide expression and dsDNA breaks.

In the growth screen, safe-targeting guides were depleted at greater rates than their non-targeting counterparts, suggesting that safe-targeting guides are more toxic than non-targeting guides. Toxicity is likely due to DNA damage and its subsequent repair. Additionally, the distribution of enrichment and depletion scores for safe-targeting guides more closely mirrors that of gene-targeting guides, suggesting that safe-targeting guides better control for the inherent effects, or background noise, of dsDNA breaks caused by CRISPR.

Safe-targeting guides also alters the discovery of hits from the growth and ricin screen. When safe-targeting guides are used to call hits from the growth screen, p-values were less significant than when

Controlling for Off-target Effects with a New Genome-wide CRISPR Screen Design (CONT'D)

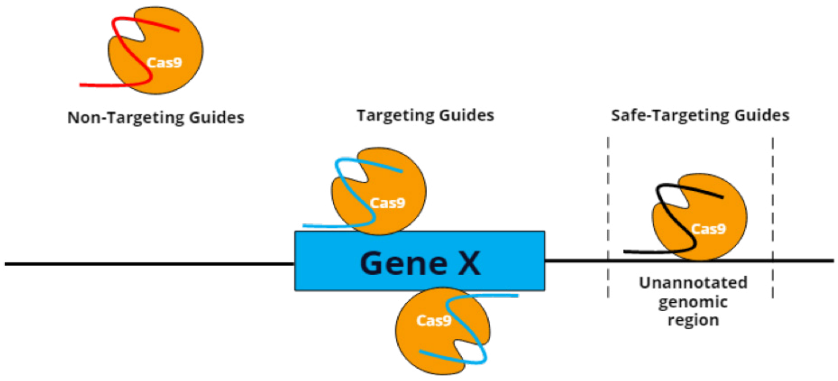


Figure 2: Non-targeting vs. safe-targeting guides. The genome-wide CRISPR screen presented in Morgens et al. contains three types of guides: 1) targeting guides (blue), which target all protein-coding genes present in the genome; 2) non-targeting guides which target no sequence in the genome. Non-targeting guides (red) are commonly used as negative controls for genome-wide CRISPR screens; and 3) safe-targeting guides which target genomic sites with no annotated function.

non-targeting guides are used as controls. Less significant p-values lead to more false-negatives results (i.e. failing to identify a gene that's a hit); but they also lead to fewer false-positives results (i.e. mistakenly calling a gene a hit when really it's not). In practical terms, this means that using safe-targeting guides as a control will result in a smaller list of true positive hits than when using non-targeting guides, when both are analyzed using the same false-discovery rate cutoff.

Adding safe-targeting guides to the CRISPR library resulted in both increased and decreased p-values of known ricin regulation genes compared to results generated using non-targeting guides as controls. While it's not clear how safe-targeting guides will affect phenotype in all non-growth screens, these results suggest that safe-targeting guides may serve as better controls than non-targeting guides and may more accurately determine the significance of hits in non-growth screens.

Off-target cutting does not confound results of CRISPR screen

Next, the toxicity of guides with off-target sites was profiled. Guides with exact or 1-bp mismatch off-targets had greater toxicity than guides that had zero mismatch off-targets. Guides with 2-bp mismatch off-targets were only toxic if they had 5+ off-target sites. Results from the screen also reproduced several previously observed characteristics that influence gRNA off-target activity: 1) mismatches closer to the PAM are less tolerated than mismatches more distal to the PAM, and 2) guides with high GC content have greater off-target activity.

Short gRNAs have lower off-target cutting and similar on-targeting cutting as full-length gRNAs

Lastly, the effects of guide length on off-target cuts was explored using full-length (19-20 bp) or short (17-18 bp) guides. When comparing the toxicity of guides that have 1-bp mismatch off-target sites, short guides had reduced toxicity compared to full-length guides. Reduced toxicity of short guides could be due to lower off-target activity, but it could also be caused by a reduction of on-target activity. However,

Controlling for Off-target Effects with a New Genome-wide CRISPR Screen Design (CONT'D)

there was no significant difference in depletion of ricin sensitizer genes identified with short and long guides. Together these results suggests that short guides have fewer off-target effects than long guides, and do not have a major reduction in on-target efficacy.

Results from Morgens et al. show that genome-wide CRISPR screens are prone to some of the same flaws as RNAi screens, such as non-specific toxicity due to both on- and off-target effects, but also presents a strategy to control for these effects. By testing thousands of guides for off-target activity, Morgens et al drew some generalizable conclusions about off-target activity, demonstrate safe-targeting guides as a potential control, and provide evidence that truncated gRNAs have improved specificity with only a small loss of on-target activity in high-throughput screens. The human and mouse versions of the Bassik CRISPR Knockout Library are available from Addgene!

References

1. Morgens, David W., et al. "Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens." Nature communications 8 (2017): 15178. Pubmed PMID: 28474669. Pubmed Central PMCID: PMC5424143.

2. Tsai, Shengdar Q., et al. "GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases." Nature biotechnology 33.2 (2015): 187. Pubmed PMID: 25513782. Pubmed Central PMCID: PMC4320685.

3. Fu, Yanfang, et al. "Improving CRISPR-Cas nuclease specificity using truncated guide RNAs." Nature biotechnology 32.3 (2014): 279. Pubmed PMID: 24463574. Pubmed Central PMCID: PMC3988262.

4. Hart, Traver, et al. "Measuring error rates in genomic perturbation screens: gold standards for human functional genomics." Molecular systems biology 10.7 (2014): 733. Pubmed PMID: 23394947. Pubmed Central PMCID: 4299491.

5. Bassik, Michael C., et al. "A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility." Cell 152.4 (2013): 909-922. Pubmed PMID: 23394947. Pubmed Central PMCID: 3652613.

Optimized Genome-wide CRISPRko, CRISPRi, and CRISPRa Libraries

By Alyssa Cecchetelli | Oct 4, 2018

CRISPR pooled libraries have allowed scientists to easily perform genome-wide screens to effectively and efficiently investigate gene function. CRISPR libraries can be used to knock out, inhibit or activate target genes by combining specific sgRNAs with Cas9 or Cas9 derivatives.

Unmodified Cas9 causes double stranded DNA breaks that commonly result in a complete loss of function while nuclease-deactivated Cas9 (dCas9) can be tethered to a repressor (CRISPRi) or an activation domain (CRISPRa) to regulate gene expression. Some CRISPR libraries contain many sgRNAs per gene to increase confidence in the genetic screen. However, there are benefits to optimized libraries that utilize only a few sgRNAs per gene. These libraries provide more information while using less resources and are particularly useful when trying to screen a limited number of cells like primary cells or *in vivo*.

Recently, the labs of David Root and John Doench at the Broad Institute have developed human genome-wide CRISPRko (knock out), CRISPRi (interference), and CRISRPa (activation) libraries that do just that (Doench et al., 2016, Sanson et al., 2018). These libraries utilize optimized sgRNAs that were designed based on the rules identified in Doench et al., 2014 to predict sgRNA activity. In this study scientists examined approximately 1,900 sgRNAs and identified sequence features and adjustments that improved sgRNA activity and thus library performance. The Root and Doench lab's optimized genome-wide CRISPRko, CRISPRi, and CRISPRa libraries have been shown to outperform other popular CRISPR libraries.

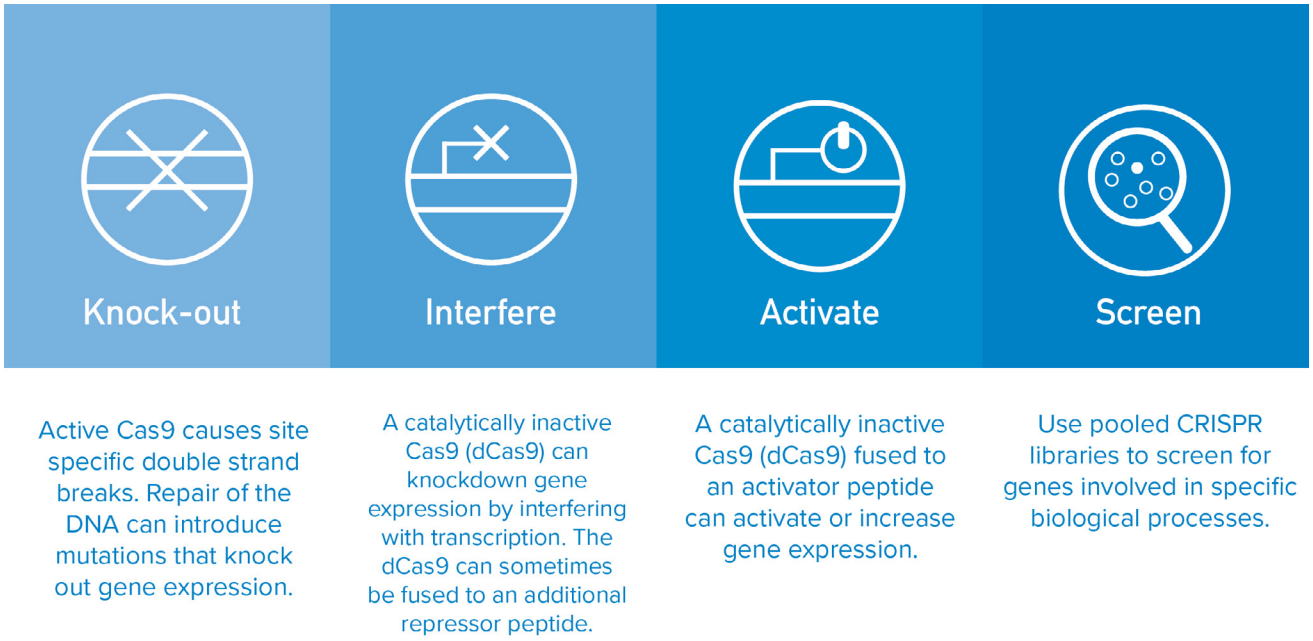


Figure 1: The different ways to use CRISPR libraries.

Optimized Genome-wide CRISPRko, CRISPRi, and CRISPRa Libraries (CONT'D)

Let's take a look at the features and advantages of the optimized human CRISPR libraries.

Brunello: human CRISPRko sgRNA library

The Brunello library consists of 77,441 sgRNAs with an average of 4 sgRNAs per gene and 1,000 non-coding control sgRNAs. This library was designed to improve on-target activity while reducing off-target effects in human genomes (Doench et al., 2016). To test the library, they conducted genome-wide negative selection screens with both A375 (melanoma) and AT29 (colon cancer) cells and assessed the libraries performance via the depletion of a gold standard set of genes (1,580 essential and 927 non-essential genes). The Root and Doench lab developed sgRNA-level metrics to assess library performance in viability screens. Based on these metrics the Brunello library was more effective than other popular CRISPR libraries in the depletion of essential genes. In the subsampling analysis of the sgRNAs, Brunello even with only a single sgRNA outperformed another CRISPR library with 6 sgRNAS, highlighting the benefit of sgRNA design in CRISPR libraries. Thus, the Brunello library proves that a few optimized sgRNAs is enough for efficient library performance, an important criterion for probing a limited number of cells.

Get the Brunello CRISPRko sgRNA library as pooled library or lentiviral prep!

Dolcetto: human CRISPRi sgRNA library

While CRISPRko libraries are highly effective at identifying strong hits in genetic screens, they do have several limitations. CRISPRko results in a complete loss of function meaning that if an essential gene plays a role in the desired screening phenotype, it may be missed because it's essential for viability. In addition, high copy number genes may mimic viability genes as they are subject to an increase in dsDNA breaks that can cause cell death. Thus, CRISPR inhibition (CRISPRi) libraries provide an alternative way to screen for loss of function phenotypes while bypassing these issues.

CRISPRi libraries utilize an engineered Cas9, which contains multiple point mutations in the nuclease domain. These mutations result in an inactive RNA-guided DNA binding protein (dCas9). dCas9 is then tethered to a repressive domain, such as KRAB, to prevent efficient transcription from sgRNA targets. CRISPRi is effective in a narrow range around the transcription start site (TSS). The Root and Doench lab selected sgRNAs in this optimal window and ranked them based on their optimized sgRNA rules and potential off-target effects. The Dolcetto library contains two sets of sgRNAs; A, the top 3 ranked sgRNAs, and B, the next 3 ranked sgRNAs. Dolcetto was able to discriminate between non-essential and essential genes similarly to the Brunello knock-out library while mitigating toxicity induced by high copy number dsDNA cutting. A Dolcetto library containing only 3 sgRNAs (setA) also outperformed CRISPRi libraries with 10 sgRNAs in the sgRNA metrics, confirming again that sgRNA design plays a

Optimized Genome-wide CRISPRko, CRISPRi, and CRISPRa Libraries (CONT'D)

significant role in library efficiency.

Try the Dolcetto library! Now also available as a lentiviral prep.

Calabrese: human CRISPRa sgRNA library

In addition to inhibiting gene expression, dCas9 can also be used for transcriptional activation. Activation of specific genes allows scientists to conduct genome-wide gain of function screens. These screens can help reveal the function of lowly expressed genes or genes that are more drastically affected by gene activation rather than depletion. Multiple strategies have been developed to recruit transcriptional machinery to dCas9 and its RNA directed binding site. These methods include the fusion of an activation domain such as VP16 to dCas9 or the recruitment of multiple copies of VP16 termed the "Sun Tag" method (Tanenbaum et al., 2014). The sgRNA or tracrRNA sequences can also be modified to include RNA domains that recruit additional transcription factors.

The Root and Doench labs utilized tracrRNAs with two PP7 stem loops to recruit additional transcription factors to the targeted region (Sanson et al., 2018) . This tracrRNA was combined with the sgRNA expression vector allowing the screen to be conducted with only two constructs (the other containing dCas9-VP16). Like CRISPRi, CRISPRa is only effective when sgRNAs target a specific location on the gene. For CRISPRa, sgRNAs were targeted between 150-75 nucleotides upstream of the TSS. Similar to the Brunello and Dolcetto library, sgRNAs were selected and ranked based on the previously established sgRNA rules and potential on- and off-target effects. The 6 best sgRNAs were selected per gene and divided into two sets, A and B. Sanson et al. demonstrated that the Calabrese library substantially identified more hits than the SAM library approach in a comparison study screening for vemurafenib resistance in A375 cells. The SAM system (CRISPR/Cas9 Synergistic Activation Mediator) uses an improved activator design that consists of inactive Cas9-VP64, a sgRNA with two MS2 RNA aptamers, and the MS2-P65-HSF1 helper protein (Konermann et al., 2015).

Root and Doench also compared the Calabrese library to screens performed with open reading frame (ORF) overexpression libraries for resistance to MEK inhibition. Although both Calabrese and the ORF overexpression library identified several overlapping genes, most genes were unique to the individual libraries. These unique hits may have been false negatives in the other screen, implying that ORF and CRISPRa libraries can be used in complement to study gene function.

Find the Calabrese pooled library and lentiviral prep here!

Optimized Genome-wide CRISPRko, CRISPRi, and CRISPRa Libraries (CONT'D)

Applications for the Root and Doench CRISPR libraries

Overall the Root and Doench lab's have proven their optimized, human genome wide CRISPRko, CRISPRi, and CRISPRa libraries are highly effective for probing gene function while using fewer sgRNAs. These new compact libraries will allow scientists to conduct genome-wide screens in standard cultured cells as well as more difficult conditions such as model systems and primary cells.

References

Doench, John G., et al. "Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation." Nature biotechnology 32.12 (2014): 1262. PubMed PMID: 25184501. PubMed Central PMCID: MC4262738.

Doench, John G., et al. "Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9." Nature biotechnology 34.2 (2016): 184. PubMed PMID: 26780180. PubMed Central PMCID: PMC4744125.

Konermann, Silvana, et al. "Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex." Nature517.7536 (2015): 583. PubMed PMID: 25494202. PubMed Centra PMCID: MC4420636.

Sanson, Kendall R., et al. "Up, down, and out: next generation libraries for genome-wide CRISPRa, CRISPRi, and CRISPR-Cas9 knockout genetic screens." bioRxiv (2018): 356626.

Tanenbaum, Marvin E., et al. "A protein-tagging system for signal amplification in gene expression and fluorescence imaging." Cell 159.3 (2014): 635-646. PubMed PMID: 25307933. PubMed Central PMCID: PMC4252608.

Chapter 9

Other CRISPR Applications

Isolating Genomic Regions of Interest with the CRISPR System

By Andrew Hempstead | May 30, 2019

While much of CRISPR research has focused on genome editing, numerous discoveries have been made using the Cas9 nuclease in the absence of genomic alterations. These studies utilize a catalytically inactive form of Cas9 known as dCas9 (Jinek et al., 2012). Like Cas9, dCas9 can bind to a specific DNA sequence via a targeting gRNA. But dCas9 does not cleave the DNA. Much of the research using dCas9 has focused on transcriptional activation using a fusion to a transcriptional activator such as VP64 (Gilbert et al., 2013), or repression of transcription through binding a promoter region to inhibit association of transcriptional activators (Qi et al., 2013). However, the fusion of dCas9 with a protein tag allows for the isolation of a genomic region of interest targeted by a gRNA.

CRISPR-mediated purification of a specific genomic region

In 2013, Hodaka Fujii's lab first described a method to purify a specific genomic region using the CRISPR system, consisting of dCas9 and targeting gRNA, to identify molecular interactions at this site (Fujita and Fujii, 2013). They call this method engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) because any engineered DNA-binding molecule, such as zinc-finger or TAL proteins, can be used. The Fujii lab demonstrated this in mammalian cells by expressing dCas9 with an N-terminal 3xFLAG tag (3xFLAG-dCas9) and a gRNA targeting their genomic locus of interest, the promoter region of the transcription factor IRF-1. Immunoprecipitation, using an antibody targeting the 3xFLAG tag, allowed these researchers to isolate dCas9 with the bound genomic locus, which they then confirmed by PCR. Subsequent mass spectroscopy studies allowed for the identification of proteins associated with the IRF-1 locus.

Prior to the development of this method, isolation of a specific genomic region was an intensive process. One method required insertion of a recognition sequence into the genomic DNA that could then be recognized by a DNA-binding molecule (Hoshino and Fujii, 2009). Other forms of enChIP, that do not use dCas9, require Zinc-finger or TAL proteins targeting a specific genomic locus, which can require multiple cloning steps (Fujita et al., 2013). CRISPR-mediated purification of a specific genomic region alleviates these issues as a region of interest can be targeted using a gRNA, which can easily be cloned into a gRNA expression plasmid.

Experimental use

Since the initial report describing CRISPR-mediated purification of a specific genomic region, numerous adaptations of this system have been made in different fields of study. In addition to the identification of DNA binding proteins, this method can characterize nucleic acid-nucleic acid interactions at a genomic locus (Fujita et al., 2017). CRISPR-mediated purification of a specific genomic region has also recently been adapted to isolate genomic loci in bacteria (Fujita et al., 2018). Furthermore, the Fujii Lab generated a transgenic mouse line and retroviral expression system for using this approach in *in vivo*

Isolating Genomic Regions of Interest with the CRISPR System (CONT'D)

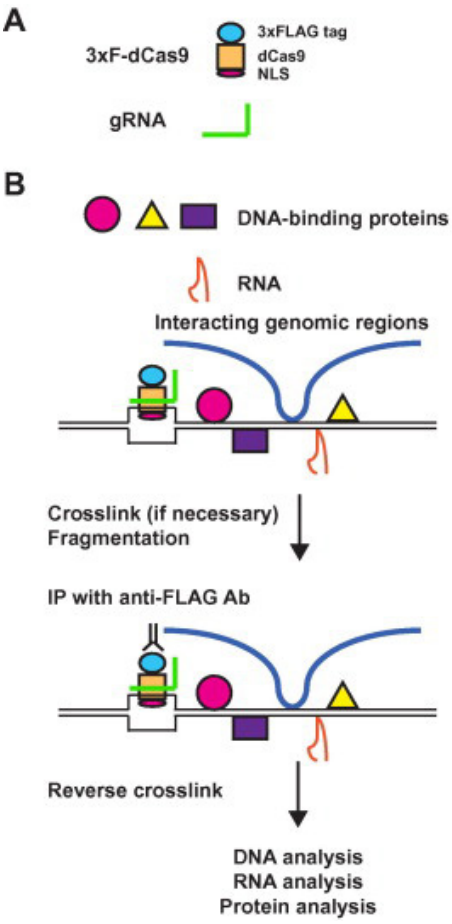


Figure 1: Schematic of enChIP. Image from Fujita and Fuji, 2013.

studies (Fujita et al., 2018).

In addition to the "in cell" form of this method, in which an engineered DNA-binding molecule such as the CRISPR complex is expressed in cells to be analyzed, the Fujii lab also developed another "*in vitro*" form. In this form, the CRISPR complex, consisting of a recombinant dCas9 protein and synthetic gRNA, is incubated with fragmented chromatin or library of purified DNA for purification of a specific chromatin or DNA fragment (Fujita et al., 2016). 3xFLAG-tag fused to dCas9, along with a biotinylated gRNA, were successfully used in this "*in vitro*" form.

Any tag system can be used to isolate genomic regions of interest with dCas9 and a targeting gRNA. One such example by the Xu lab is CRISPR affinity purification *in situ* of regulatory elements (CAPTURE) (Liu et al., 2017). This system utilizes dCas9 fused to a biotinylation tag expressed on a plasmid, allowing for *in vivo* biotinylation through a plasmid-based biotin ligase. Biotinylated dCas9, with bound DNA, can then be purified by streptavidin high-affinity purification. This system is both highly sensitive and specific and allowed the Xu lab to identify protein and nucleic acid binding partners at numerous genomic loci through proteomics and the chromosome conformation capture (3C) technique. In addition to 3xFLAG and biotin tags, use of tag systems including Protein A, 1xFLAG, 2xAM, and HA have been reported, as well as affinity purification with an anti-Cas9 antibody. These different tag systems and affinity purification schemes add additional flavors to this methodology.

Using this technique, researchers have begun to make important discoveries in various biological fields. In 2015, the Fujii Lab identified non-coding RNAs associated with telomeres by combining purification of a specific genomic region and RNA sequencing (Fujita et al., 2015). Using similar affinity purification with an anti-Cas9 antibody, the Li lab found a microRNA (miR483) that binds to the promoter of the growth factor IGF2 open reading frame to relax imprinting (Zhang et al., 2017). This upregulates expression of IGF2, a characteristic of many human malignancies. Additionally, the Tapscott Lab identified regulators of DUX4, a transcriptional regulator that when inappropriately expressed can cause facioscapulohumeral muscular dystrophy (FSHD), a disease for which there is currently no cure (Campbell et al., 2018).

Find plasmids from the Fujii lab!

Isolating Genomic Regions of Interest with the CRISPR System (CONT'D)

Conclusions

CRISPR-mediated purification of a specific genomic region is another example of the many tools based on targeting of a genomic locus through a specific gRNA-dCas9 complex. When combined with next-generation sequencing, this method allows for the recovery of a vast amount of data from a single experimental setup. As we've seen for many other CRISPR techniques, this method also has great potential for improvements and adaptations that can provide an increased understanding of DNA binding and gene regulation.

For additional information on CRISPR-mediated purification of a specific genomic region, other forms of enChIP, and additional research conducted in the Hodaka Fujii Laboratory, please also see our interview with Dr. Fujii.

References

Campbell, Amy E., et al. "NuRD and CAF-1-mediated silencing of the D4Z4 array is modulated by DUX4-induced MBD3L proteins." Elife 7 (2018): e31023. PubMed PMID: 29533181. PubMed Central PMCID: PMC5849414.

Fujita, Toshitsugu, and Hodaka Fujii. "Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR." Biochemical and biophysical research communications 439.1 (2013): 132-136. PubMed PMID: 23942116.

Fujita, Toshitsugu, et al. "Identification of telomere-associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP)." Scientific reports 3 (2013): 3171. PubMed PMID: 24201379. PubMed Central PMCID: PMC3821016.

Fujita, Toshitsugu, et al. "Identification of non-coding RNAs associated with telomeres using a combination of enChIP and RNA sequencing." PLoS One 10.4 (2015): e0123387. PubMed PMID: 25874893. PubMed Central PMCID: PMC4395285.

Fujita, Toshitsugu, Miyuki Yuno, and Hodaka Fujii. "Efficient sequence-specific isolation of DNA fragments and chromatin by *in vitro* enCh IP technology using recombinant CRISPR ribonucleoproteins." Genes to Cells 21.4 (2016): 370-377. PubMed PMID: 26848818.

Isolating Genomic Regions of Interest with the CRISPR System (CONT'D)

Fujita, Toshitsugu, et al. "Identification of physical interactions between genomic regions by enCh IP-Seq." Genes to Cells22.6 (2017): 506-520. PubMed PMID: 28474362.

Fujita, Toshitsugu, Miyuki Yuno, and Hodaka Fujii. "An enChIP system for the analysis of bacterial genome functions." BMC research notes 11.1 (2018): 387. PubMed PMID: 29898790. PubMed Central PMCID: PMC6001023.

Fujita, Toshitsugu, et al. "Transgenic mouse lines expressing the 3x FLAG-dC as9 protein for enCh IP analysis." Genes to cells 23.4 (2018): 318-325. PubMed PMID: 29480524.

Gilbert, Luke A., et al. "CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes." Cell 154.2 (2013): 442-451. PubMed PMID: 23849981. PubMed Central PMCID: PMC3770145.

Hoshino, Akemi, and Hodaka Fujii. "Insertional chromatin immunoprecipitation: a method for isolating specific genomic regions." Journal of bioscience and bioengineering 108.5 (2009): 446-449. PubMed PMID: 19804873.

Jinek, Martin, et al. "A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity." Science337.6096 (2012): 816-821. PubMed PMID: 22745249. PubMed Central PMCID: PMC6286148.

Liu, Xin, et al. "*In situ* capture of chromatin interactions by biotinylated dCas9." Cell 170.5 (2017): 1028-1043. PubMed PMID: 28841410.

Qi, Lei S., et al. "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." Cell 152.5 (2013): 1173-1183. PubMed PMID: 23452860. PubMed Central PMCID: PMC3664290.

Zhang, Yiqun, et al. "CRISPR Cas9-guided chromatin immunoprecipitation identifies miR483 as an epigenetic modulator of IGF2 imprinting in tumors." Oncotarget 8.21 (2017): 34177. PubMed PMID: 27486969. PubMed Central PMCID: PMC5470959.

Optogenetics + CRISPR, Using Light to Control Genome Editing

By Caroline LaManna | Mar 8, 2016, Updated Sep 3, 2020 Nyla Naim

Scientists around the world have been making major improvements to CRISPR technology since its initial applications for genome engineering in 2012. Many of these advances have stemmed from the goal of reducing off-target Cas9 activity. The use of nickases, prime editing, anti-CRISPR proteins, and other techniques all aim to improve targeting specificity or reduce the duration of Cas9 activity.

The field of optogenetics is renowned for enabling precise temporal and spatial control. Optogenetics uses genetically encoded tools, such as microbial opsins, to control cellular activities using light. In 2015, scientists combined CRISPR and optogenetic techniques to develop a variety of photoactivatable CRISPR tools. These tools allow scientists to use light to externally control the location, timing, and reversibility of the genome editing process. Read on to learn about the various light-controlled CRISPR tools available to researchers - some readily found at Addgene.

Shining light on transcriptional activation using dCas9

Initial photoactivatable CRISPR systems published in early 2015 focused on using light to control transcription. Two separate labs, Moritoshi Sato's lab at the University of Tokyo and Charles Gersbach's lab at Duke University developed similar systems based on the light-inducible heterodimerizing cryptochrome 2 (CRY2) and calcium and integrin-binding protein 1 (CIB1) proteins. The goal of both groups was to create a system that would use light to turn on and off gene expression while imparting spatiotemporal control, reversibility, and repeatability.

The system developed by the Sata lab is composed of two fusion proteins: 1) the genomic anchor - an inactive, dead Cas9 protein (dCas9) fused to CIB1; and 2) the activator - the CRY2 photolyase homology region (CRY2PHR) fused to a transcriptional activator domain (VP64 or p65). Upon expression in the cell, the dCas9-CIB1 fusion binds to the target DNA sequence as directed by the guide RNA (gRNA), while the CRY2PHR-activator fusion floats freely, depicted in the figure below. Once triggered by blue light, the CRY2 and CIB1 proteins heterodimerize and move the activator into position to activate gene transcription. The researchers tested a variety of combinations to optimize both fusion proteins. The best performing combination was NLS-dCas9-trCIB1 and NLSx3-CRYPHR-p65 - it had the lowest background activity in the dark state and highest fold induction at 31X. By using a slit pattern during blue light exposure (470nm), the researchers showed that expression of the human ASCL1 gene could be spatially controlled in a reversible and repeatable manner.

With their light-activated CRISPR/Cas9 effector (LACE) system, the Gersbach lab utilized a similar strategy to develop an optimized photoactivatable CRISPR gene activation system, but settled on a different optimal fusion protein combination. The optimized LACE system consisted of: 1) CIBN-dCas9-CIBN, where CIBN is the N-terminal fragment of CIB1 and it was fused to both the N- and C-termini of dCas9; and 2) CRY2FL-VP64, a fusion of full-length CRY2 and the transcriptional activator domain VP64. Using this system in HEK293T cells to induce expression of human IL1RN, the researchers saw

Optogenetics + CRISPR, Using Light to Control Genome Editing (CONT'D)

an 11-fold increase in mRNA levels after 2 hr and a 400-fold increase after 30 hr. The system was also shown to be reversible and repeatable when blue light (450nm) was cycled on-off-on. Using a slit photomask, the researchers also demonstrated the ability to spatially control gene expression.

Both of these techniques take advantage of light-induced dimerization to activate their CRISPR system. This common optogenetic technique can be employed in other split systems where light is required to reconstitute enzymatic activity.

Photoactivatable genome modifications by NHEJ and HDR using split systems

Later in 2015, the Sato lab unveiled a photoactivatable system to cleave a target DNA sequence (Nihongaki, et al., 2015) resulting in a double strand break (DSB) that can be repaired by either non-homologous end joining (NHEJ) or homology directed repair (HDR). This system is unique in that it utilizes a "split" nuclease - the authors fragmented Cas9 into N-terminal (residues 2-713, N713) and C-terminal (residues 714-1368, C714) halves, rendering the Cas9 non-functional when split but regaining functionality when the halves are reassociated. By fusing a photoinducible, heterodimerizing domain to each of the Cas9 fragments, the authors created a photoactive Cas9 tool, as shown in the figure. Their most successful design utilized magnet photoswitchable proteins derived from the fungal photoreceptor, Vivid (VVD, N. crassa) (Kawano et al., 2015). Nicknamed paCas9-1 and consisting of the fusion proteins N713-pMag and nMagHigh1-C714, this new system had both low background and high fold-induction of Cas9 activity (16.4-fold). This paCas9-1 light-inducible system was able to recognize the same PAM and had similar targeting specificity as full-length Cas9 (flCas9). When triggered by blue light (470nm), paCas9-1 induced indel mutations via NHEJ (frequency of 20.5%) and induced modifications by HDR (frequency of 7.2%).

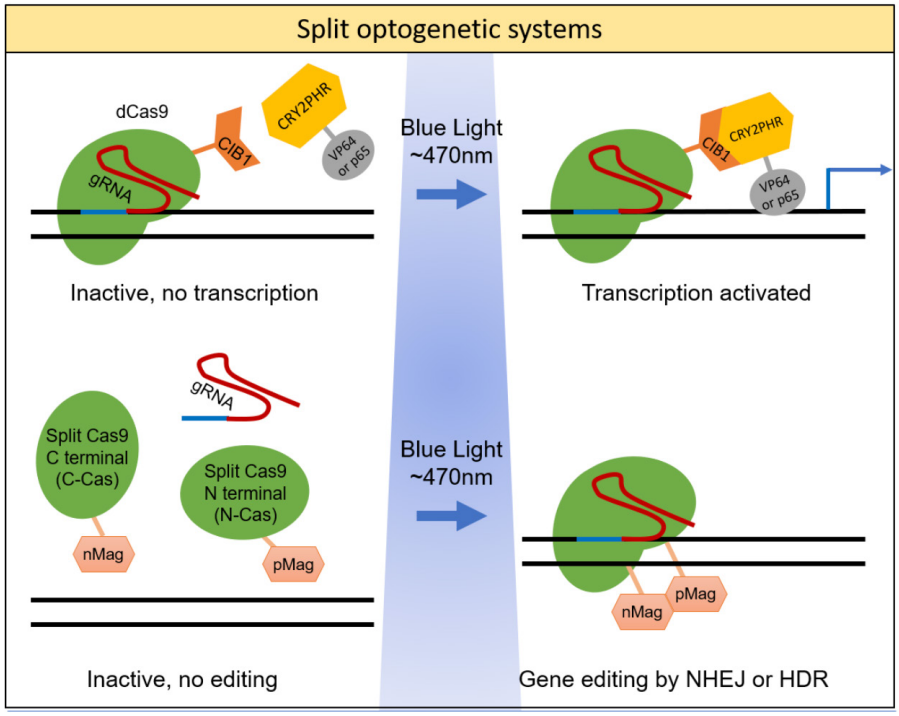


Figure 1: These split optogenetic system is activated by blue light to cause transcriptional activation or editing at target locus.

Optogenetics + CRISPR, Using Light to Control Genome Editing (CONT'D)

The authors additionally showed that they could lower the background activity of the system by modifying paCas9-1 using nMagC714 instead of nMagHigh1-C714, generating paCas9-2. This change did not significantly alter the system's efficiency at generating mutations when activated with light and lowered background DSBs (non detectable). Like their prior work, the Sato lab also showed that the paCas9-2 system could be spatially controlled and reversibly activated by turning blue light on and off.

As one might expect from the modular nature of Cas9, Nihongaki et al. showed that it was possible to swap out the Cas9 domains in their split fusions and generate a photoactivatable nickase and a photoactivatable repressor (dCas9). The activity of all variants was reversible and repeatable.

Since these initial studies published in 2015, more split systems have been developed for use with CRISPR-based gene editing. For example, the far-red light-activated split-Cas9 (FAST) system constitutively expresses the C-terminal fragment of Cas9 while transcription of the N-terminus is light-dependent (Yu et al., 2020). To express the N-Cas9, Haifeng Ye's lab used the bacterial phytochrome BphS which produces c-di-GMP in response to far-red light (730 nm). This nucleotide triggers the translocation of the protein BldD to the nucleus. To couple this light-dependent activity to transcription, they fused BldD to the transcriptional activators p65 and VP64 which in turn transcribes N-Cas9. The N- and C-Cas9 fragments then dimerize to form a functional protein due to being fused with high-affinity interacting proteins Coh2 and DocS, respectively.

The use of far-red light is of particular interest since blue light has limited applications *in vivo*. UV and blue light is readily absorbed and scattered by the skin, so it is not able to reach deeper tissues. Additionally, high intensity exposure can cause DNA damage. Far-red light, however, can penetrate 5 mm past the surface of skin and is less phototoxic. Using the FAST system, Ye's lab could edit a tdTomato fluorescent reporter gene in mice using Minicircle DNA vectors encoding the FAST constructs. This proof-of-concept model opens intriguing possibilities for optogenetic manipulation of CRISPR-based systems in animal models.

Single-chain Cas9 photoswitches

Photoactivation of Cas9 can also be achieved using single-chain fusion proteins by fusing Cas9 to photoreceptors. One prime example was from the lab of Michael Lin, who developed a chimeric protein consisting of Cas9 and a photo-dissociable dimeric green fluorescent protein (pdDronpa) (Zhou et al., 2018). pdDronpa dimers dissociate in response to cyan light (500nm). By fusing the two pdDronpa domains to the sites flanking the DNA-binding cleft of Cas9, the fluorescent domains dimerize and prevent DNA binding in the dark. Upon light stimulation, the pdDronpa moieties move apart, allowing DNA binding and cleavage. This technique was applied to create photoswitchable spCas9, dCas9, and *S. aureus* Cas9 (saCas9) proteins, showing the versatility of this approach. This study generated the first optically active saCas9 protein.

Optogenetics + CRISPR, Using Light to Control Genome Editing (CONT'D)

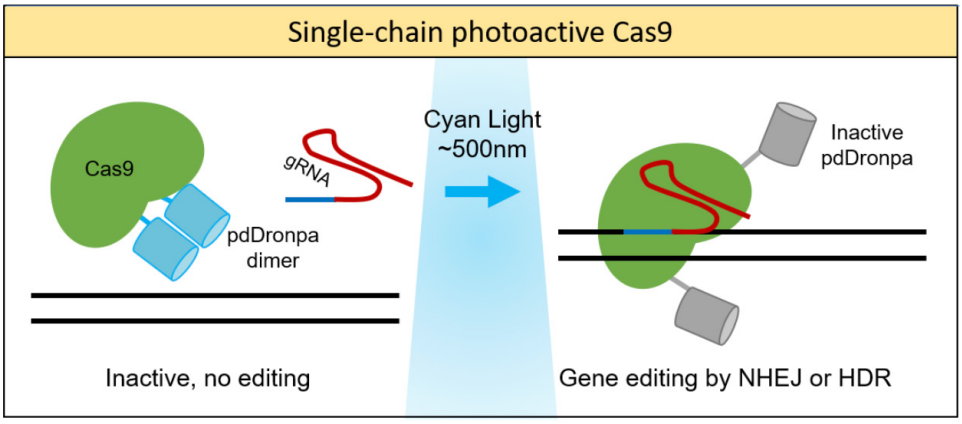


Figure 2: A fusion protein of Cas9 with photo-dissociable dimeric GFP responds to cyan light. Without light, the dimer blocks Cas9. With light, Cas9 can bind DNA.

Stop light! Photoactive anti-CRISPRs halt Cas9 activity

Anti-CRISPR proteins are a highly diverse group of proteins that are able to block CRISPR-Cas systems. They have been used to block genome editing, reduce off-target activity, and prevent cytotoxic effects. For a more detailed review, read the article, Anti-CRISPR Proteins Switch Off CRISPR-Cas Systems.

Simply expressing anti-CRISPR proteins does not provide much control over when Cas9 activity is blocked. Hence, co-opting optogenetic techniques allows greater precision and control over Cas9 activity. Dominik Niopek's lab achieved this by fusing the anti-CRISPR Acr protein, AcrIIA4, to a photo-sensitive LOV2 domain from *A. sativa* (Bubeck et al., 2018). LOV2 has a C-terminal helix that unfolds when stimulated by blue light. When fused to the catalytic site of AcrIIA4, this unfolding conformational shift can interfere with activity. The resulting system was coined as CRISPR–Cas9 activity switching via a novel optogenetic variant of AcrIIA4 (CASANOVA). As shown in the figure below, in the absence of light, Cas9 is bound by AcrIIA4 and rendered inactive. In the presence of blue light, AcrIIA4 is inhibited by the unfolding of LOV, allowing the release of Cas9.

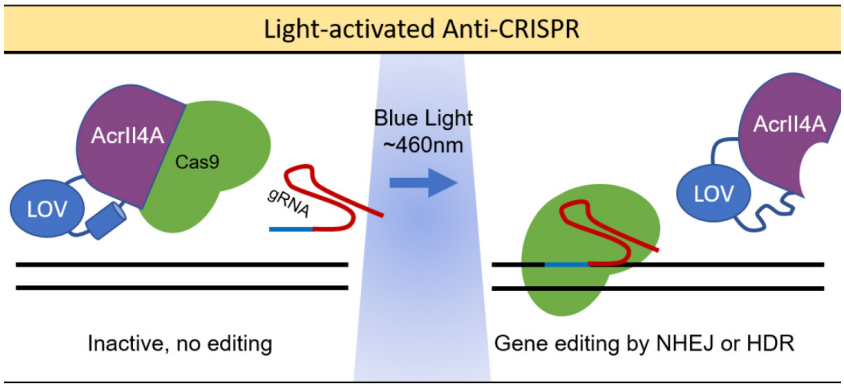


Figure 3: Anti-CRISPR activity controlled with blue light. With blue light, the anti-CRISPR activity is alleviated.

A noteworthy application of the CASANOVA system was its use in light-

Optogenetics + CRISPR, Using Light to Control Genome Editing (CONT'D)

mediated CRISPR labeling to study chromatin architecture (Hoffman et al., 2020). Fluorescent dCas9 proteins can be used to identify specific genomic loci for imaging techniques. Since expression of dCas9 can potentially alter the genome and chromatin architecture, it is useful to regulate the time in which dCas9 binding occurs. Bubeck et al. expressed a telomere-targeting gRNA and a dCas9 fusion protein with three red fluorescent tags in U2OS cells. Upon light activation, dCas9 is released within minutes and red foci formed 20-40 min later. The CASANOVA system is versatile in that it can be adapted to regulate gene editing, turn on transcription, label chromatin, and study kinetics.

Using chemistry to photocage CRISPRs

The aforementioned techniques each employed a photoactive strategy which had been engineered from naturally occurring photoactive proteins (i.e. CRY2 and Vivid) - Alexander Deiters' lab, on the other hand, took a different approach. These researchers used a genetically encoded photocaging technique to insert a light-removable protecting group, specifically a nitrobenzyl photocaged lysine (PCK), on the Cas9 protein (Hemphill et al., 2015). In order to insert the PCK into a specific site on the Cas9, the group used an engineered pyrrolysyl tRNA/tRNA synthetase pair which would add the PCK upon reaching the amber stop codon, TAG. (To learn more about site-specific incorporation of amino acids using pyrrolysl tRNA synthetase, read this article).

The group first tested photocaging various lysines in Cas9 to determine which best deactivated the protein's ability to cleave targeted DNA, settling on photocaging the K866 lysine, as seen in the figure. Next, by using a dual reporter fluorescence assay, Hemphill et al. demonstrated that the Cas9 K866PCK mutant was indeed inactive prior to irradiation with UV light (365nm) and that post-UV exposure it showed cleavage activity similar to the wild-type Cas9. This photocaging technique was also shown to impart spatial control of Cas9 cleavage when using a photomasking technique. Last, Hemphill et al. presented data showing that this genetically encoded, photocaged Cas9 system could silence endogenous gene expression - demonstrating light-induced silencing of transferrin receptor CD71 in HeLa cells.

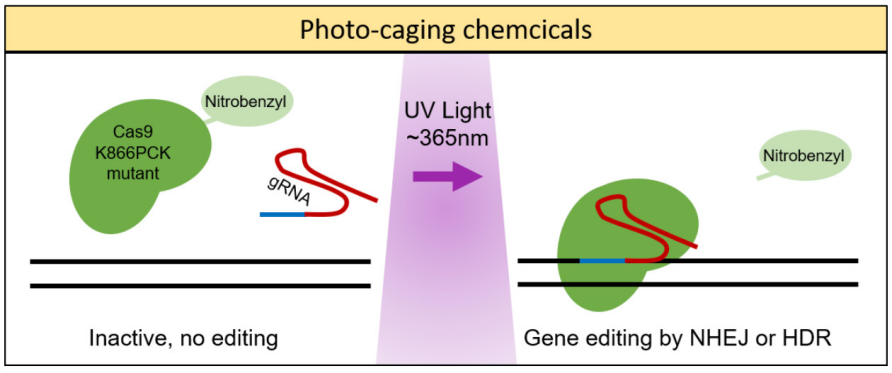


Figure 4: The chemical addition of nitrobenzyl photocaged lysine on the Cas9 protein is removed by light to allow editing.

For more details, check out this comprehensive review comparing optogenetic CRISPR techniques

Optogenetics + CRISPR, Using Light to Control Genome Editing (CONT'D)

(Matony et al., 2020). Whether you are looking to activate, repress, or modify a gene, you now have the tools at your disposal to control your genome editing using light. We look forward to more tools as CRISPR and optogenetics continue to evolve and can't wait to see what cool applications you use these for in the future!

References

Bubeck F, Hoffmann MD, Harteveld Z, Aschenbrenner S, Bietz A, Waldhauer MC, Börner K, Fakhiri J, Schmelas C, Dietz L, Grimm D, Correia BE, Eils R, Niopek D (2018) Engineered anti-CRISPR proteins for optogenetic control of CRISPR–Cas9. Nat Methods 15:924–927 . https://doi.org/10.1038/s41592-018-0178-9

Hemphill J, Borchardt EK, Brown K, Asokan A, Deiters A (2015) Optical Control of CRISPR/Cas9 Gene Editing. J Am Chem Soc 137:5642–5645 . https://doi.org/10.1021/ja512664v

Jain PK, Ramanan V, Schepers AG, Dalvie NS, Panda A, Fleming HE, Bhatia SN (2016) Development of Light-Activated CRISPR Using Guide RNAs with Photocleavable Protectors. Angew Chem Int Ed 55:12440–12444 . https://doi.org/10.1002/anie.201606123

Kim JH, Rege M, Valeri J, Dunagin MC, Metzger A, Titus KR, Gilgenast TG, Gong W, Beagan JA, Raj A, Phillips-Cremins JE (2019) LADL: light-activated dynamic looping for endogenous gene expression control. Nat Methods 16:633–639 . https://doi.org/10.1038/s41592-019-0436-5

Mathony J, Hoffmann MD, Niopek D (2020) Optogenetics and CRISPR: A New Relationship Built to Last. In: Methods in Molecular Biology. Springer US, pp 261–281

Manna D, Maji B, Gangopadhyay SA, Cox KJ, Zhou Q, Law BK, Mazitschek R, Choudhary A (2019) A Singular System with Precise Dosing and Spatiotemporal Control of CRISPR- Cas9. Angew Chem Int Ed 58:6285–6289 . https://doi.org/10.1002/anie.201900788

Nihongaki Y, Furuhata Y, Otabe T, Hasegawa S, Yoshimoto K, Sato M (2017) CRISPR–Cas9-based photoactivatable transcription systems to induce neuronal differentiation. Nat Methods 14:963–966 . https://doi.org/10.1038/nmeth.4430

Nihongaki Y, Kawano F, Nakajima T, Sato M (2015) Photoactivatable CRISPR-Cas9 for optogenetic genome editing. Nat Biotechnol 33:755–760 . https://doi.org/10.1038/nbt.3245

Nihongaki Y, Yamamoto S, Kawano F, Suzuki H, Sato M (2015) CRISPR-Cas9-based Photoactivatable Transcription System. Chemistry & Biology 22:169–174 . https://doi.org/10.1016/j.chembiol.2014.12.011

Optogenetics + CRISPR, Using Light to Control Genome Editing (CONT'D)

Polstein LR, Gersbach CA (2015) A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. Nat Chem Biol 11:198–200 . https://doi.org/10.1038/nchembio.1753

Richter F, Fonfara I, Bouazza B, Schumacher CH, Bratovič M, Charpentier E, Möglich A (2013) Engineering of temperature- and light-switchable Cas9 variants. Nucleic Acids Research 44: . https://doi.org/10.1093/nar/gkw930

Yu Y, Wu X, Guan N, Shao J, Li H, Chen Y, Ping Y, Li D, Ye H (2020) Engineering a far-red light–activated split-Cas9 system for remote-controlled genome editing of internal organs and tumors. Sci Adv 6:eabb1777 . https://doi.org/10.1126/sciadv.abb1777

Zhou W, Brown W, Bardhan A, Delaney M, Ilk AS, Rauen RR, Kahn SI, Tsang M, Deiters A (2020) Spatiotemporal Control of CRISPR/Cas9 Function in Cells and Zebrafish using Light-Activated Guide RNA. Angew Chem Int Ed 59:8998–9003 . https://doi.org/10.1002/anie.201914575

Zhou XX, Zou X, Chung HK, Gao Y, Liu Y, Qi LS, Lin MZ (2017) A Single-Chain Photoswitchable CRISPR-Cas9 Architecture for Light-Inducible Gene Editing and Transcription. ACS Chem Biol 13:443–448 . https://doi.org/10.1021/acschembio.7b00603

Finding nucleic acids with SHERLOCK and DETECTR

By Alyssa Cecchetelli | Aug 30, 2018, Updated Apr 16, 2020 by Jennifer Tsang

Sensitive and specific nucleic acid detection is crucial for clinical diagnostics, genotyping, and biotechnological advancements. Many methods of nucleic acid detection however, either lack the sensitivity or the specificity to detect nucleic acids at low concentrations and/or are too expensive, time-consuming, and complex to use outside of standard laboratories. In the case of the COVID-19 pandemic, qPCR can be used to diagnose the presence of SARS-CoV-2 RNA, but inadequate access to reagents and equipment has become a bottleneck.

In the last few years, scientists have utilized CRISPR-Cas9 protein variants such as Cas13, and Cas12a to develop simple, portable, and inexpensive platforms to reliably detect nucleic acids at the atomolar level.

The Zhang lab has adapted natural RNase activity of the Cas13 protein to develop and optimize the method termed Specific High Sensitivity Enzymatic Reporter UnLOCKING (SHERLOCK and SHERLOCKv2) (Gootenberg et al., 2017 and 2018). Meanwhile, the Doudna lab has used Cas12a's non-specific ssDNA degradation to develop the method termed DNA Endonuclease Targeted CRISPR Trans Reporter (DETECTR) (Chen et al., 2018).

Both SHERLOCK and DETECTR harness the promiscuous cleavage and degradation of neighboring ssRNA and ssDNA by Cas13 and Cas12a, respectively, to cleave and activate a reporter. The detectable signal from this reporter can be measured and quantified to determine the presence and quantity of DNA, RNA or a mutation of interest (Figure 1). Together SHERLOCK and DETECTR demonstrate the power of CRISPR-based diagnostics.

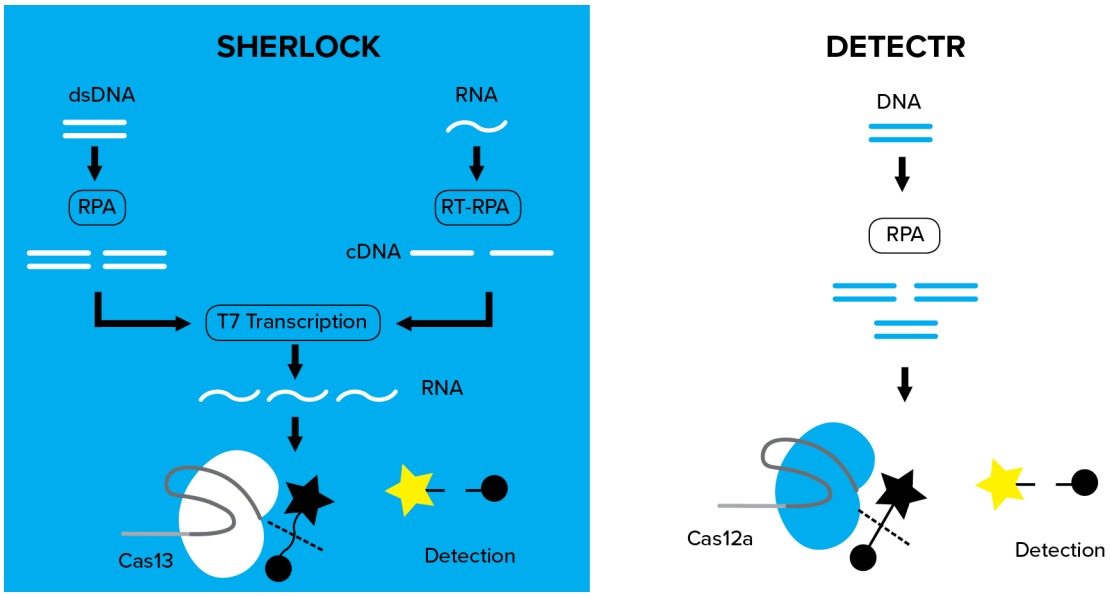


Figure 1: A compaison between SHERLOCK and DETECTR.

Finding nucleic acids with SHERLOCK and DETECTR (CONT'D)

SHERLOCK- Specific High Sensitivity Enzymatic Reporter UnLOCKING

Cas13 (C2c2) was first identified in 2016 by the Zhang lab as an RNA guided RNase (Abudayyeh et al., 2016). Cas13 can be guided by a single CRISPR RNA (crRNA) to cleave ssRNA or mRNA. It also exhibits a "collateral effect" of non-specific ssRNA cleavage.

How does SHERLOCK work?

Cas13 can be programmed with crRNA to target an ssRNA of interest, for example, a sequence specific to a virus or pathogen. Once Cas13 recognizes and binds to the programmed sequence, it can promiscuously cleave surrounding ssRNA molecules. In SHERLOCK a quenched fluorescent ssRNA reporter is added to the reaction. Cleavage of the quenchable fluorescent RNA by the "activated" Cas13 produces a quantifiable signal that indicates the presence of your targeted nucleic acid. To increase the sensitivity of the assay, targeted DNA or RNA from a sample is first amplified using RPA (recombinase polymerase amplification) or reverse transcriptase (RT)-RPA, respectively. RPA is coupled with T7 transcription to convert amplified DNA to RNA for subsequent detection by Cas13. This amplification step in combination with the ssRNA reporter enables SHERLOCK to detect DNA or RNA with atomolar sensitivity and single base pair mismatch specificity.

SHERLOCKv2

SHERLOCK had a few limitations when it was first introduced in 2017 so the lab made further modifications to improve their detection system. They named this system SHERLOCKv2 for SHERLOCK version 2. Here are some of the improvements:

- SHERLOCKv2 uses far less primer in the pre-amplification step allowing for greater quantitation without compromising sensitivity.
- To detect multiple targets in one reaction, Cas enzymes, such as variations of Cas13 and Cas12a, are combined with different fluorescent reporters for each enzyme.
- In addition, Csm6, a CRISPR type-III effector nuclease, can be used in conjunction with Cas13 to amplify the signal of a single target. Csm6 can cleave ssRNA complementary to a crRNA of interest (Niewoehner and Jinek, 2016) and is conveniently activated by Cas13 ssRNA cleavage byproducts. Thus, binding of Cas13 to the programmed region of interest would lead to both the cleavage of a Cas13 specific reporter and the activation of Csm6. The Zhang lab showed that, by adding Csm6 and a Csm6 specific reporter (in the same fluorescent channel as the Cas13 reporter), they could significantly amplify the detection signal for a single target.
- SHERLOCKv2 was adapted so that a cleaved reporter could be detected on commercial lateral flow strips, similar to pregnancy tests. With lateral flow strips, the presence of your DNA or RNA of

Finding nucleic acids with SHERLOCK and DETECTR (CONT'D)

interest within a given sample is simply determined by the number of bands present on the strip. This type of readout allows for nucleic acid detection almost anywhere as lateral flow strips are easy to transport and work rapidly, providing reliable results in as little as an hour.

Find the SHERLOCK plasmids at Addgene!

Applications of the SHERLOCK detection system

The Zhang lab demonstrated that SHERLOCK could reliably distinguish between Zika and a closely-related virus, Dengue from multiple sample sources. SHERLOCK could also detect low-frequency cancer mutations from cell-free DNA fragments as well as health-related single nucleotide polymorphisms (SNPs) from human saliva.

SHERLOCKv2 provides a method to detect nucleic acids with high sensitivity and specificity without compromising speed, ease of use, and portability. This method can be used for an array of applications including clinical diagnostics (e.g. pathogen or virus detection), therapeutics and sensitive genotyping. The beauty of SHERLOCK systems is that it can be used easily and effectively in the lab and in the field (Myhrvold et al., 2018).

Now, the Zhang lab has shared a protocol for using SHERLOCK to detect SARS-CoV-2 RNA. The test is started using RNA purified from patient samples and can be read using a dipstick in under one hour. While the lab points out that the protocol is not approved for clinical use at this point, they hope that the protocol serves as a platform for advancing diagnostics.

DETECTR- DNA Endonuclease Targeted CRISPR Trans Reporter

Cas12a (Cpf1) is a CRISPR Cas variant that can also cleave dsDNA similarly to Cas9 (Zetsche et al., 2015). Cas12a however recognizes a different PAM site and generates 5' and 3' staggered ends after dsDNA breaks. The Doudna lab discovered Cas12a's ability to cleave non-specific (trans) ssDNA and harnessed this ability to create a DNA detection platform called DETECTR.

How does DETECTR work?

DETECTR works similarly to SHERLOCK. Cas12a is targeted to a specific DNA sequence, such as the Human papilloma virus (HPV) genome, via a crRNA. An ssDNA-fluorescently quenched (FQ) reporter, which will produce a signal when the ssDNA is degraded, is added to the reaction. To enhance sensitivity, the DNA is first amplified through isothermal amplification by RPA. When Cas12a-cRNA base pairs with the dsDNA of interest, the DNase activity of Cas12a is initiated. Surrounding trans-ssDNA, including the ssDNA-FQ reporter are subsequently degraded at a rate of ~1,250 cuts per second. A

Finding nucleic acids with SHERLOCK and DETECTR (CONT'D)

quantifiable fluorescent signal designates the presence of your DNA of interest, in this case HPV.

As proof of concept, the lab demonstrated that DETECTR could accurately distinguish between two similar types of HPV, HPV16 and HPV18, from human cells, at atomolar levels, within one hour. DETECTR thus has the ability to rapidly detect nucleic acids with high selectivity and sensitivity from patient samples.

Find the DETECTR plasmids here!

Applications of DETECTR in diagnostics

DETECTR allows for simple and efficient detection of nucleic acids in a mixed population for an array of molecular and clinical diagnostic applications. The company Mammoth Biosciences has been started based on DETECTR technology with the mission to "provide a CRISPR-based platform on which an infinite number of tests" for biosensing can be built upon.

Mammoth Biosciences and UCSF recently adapted the DETECTR platform to detect SARS-CoV-2 using a lateral flow strip format. They published this in Nature Biotechnology. Like the SHERLOCK detection system, DETECTR has not been approved for diagnostics yet, but is undergoing validation studies.

References

Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DBT, Shmakov S, Makarova KS, Semenova E, Minakhin L, Severinov K, Regev A, Lander ES, Koonin EV, Zhang F (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science 353:aaf5573 . https://doi.org/10.1126/science.aaf5573

Broughton JP, Deng X, Yu G, Fasching CL, Servellita V, Singh J, Miao X, Streithorst JA, Granados A, Sotomayor-Gonzalez A, Zorn K, Gopez A, Hsu E, Gu W, Miller S, Pan C-Y, Guevara H, Wadford DA, Chen JS, Chiu CY (2020) CRISPR–Cas12-based detection of SARS-CoV-2. Nature Biotechnology. https://doi.org/10.1038/s41587-020-0513-4

Chen JS, Ma E, Harrington LB, Da Costa M, Tian X, Palefsky JM, Doudna JA (2018) CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. Science 360:436–439 . https://doi.org/10.1126/science.aar6245

Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, Verdine V, Donghia N, Daringer NM, Freije CA, Myhrvold C, Bhattacharyya RP, Livny J, Regev A, Koonin EV, Hung DT, Sabeti PC, Collins JJ, Zhang F (2017) Nucleic acid detection with CRISPR-Cas13a/C2c2. Science 356:438–442

Finding nucleic acids with SHERLOCK and DETECTR (CONT'D)

. https://doi.org/10.1126/science.aam9321

Gootenberg JS, Abudayyeh OO, Kellner MJ, Joung J, Collins JJ, Zhang F (2018) Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. Science 360:439–444 . https://doi.org/10.1126/science.aaq0179

Myhrvold C, Freije CA, Gootenberg JS, Abudayyeh OO, Metsky HC, Durbin AF, Kellner MJ, Tan AL, Paul LM, Parham LA, Garcia KF, Barnes KG, Chak B, Mondini A, Nogueira ML, Isern S, Michael SF, Lorenzana I, Yozwiak NL, MacInnis BL, Bosch I, Gehrke L, Zhang F, Sabeti PC (2018) Field-deployable viral diagnostics using CRISPR-Cas13. Science 360:444–448 . https://doi.org/10.1126/science.aas8836

Niewoehner O, Jinek M (2016) Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. RNA 22:318–329 . https://doi.org/10.1261/rna.054098.115

Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, Semenova E, Minakhin L, Joung J, Konermann S, Severinov K, Zhang F, Koonin EV (2015) Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. Molecular Cell 60:385–397 . https://doi.org/10.1016/j.molcel.2015.10.008

Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F (2015) Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell 163:759–771 . https://doi.org/10.1016/j.cell.2015.09.038

Plasmids for Endogenous Gene Tagging in Human Cells

By the Allen Institute of Cell Science | Apr 6, 2017, Updated Oct 14, 2020 by Seth Kasowitz

A classic challenge in cell biology is making sure that what we observe through the microscope represents reality as accurately as possible. This is especially true in the case of protein tagging to elucidate cellular structures. Overexpression methods flood the cell with protein, which can both interfere with a cell's normal function and result in a ubiquitous background signal that makes it hard to visualize the precise location of the protein or structure of interest.

Endogenous gene tagging is an ideal solution because it allows for tagging and visualization of specific, individual proteins under endogenous regulatory control. But even with the advent of CRISPR/Cas9 technology, inserting large tags into a precise location in the genome is still inefficient, particularly in human cell lines. Furthermore, the quality control necessary to ensure the edited cells are behaving normally can be prohibitively expensive for many labs.

We've created and tested plasmids that use CRISPR/Cas9 to endogenously tag a wide variety of genes with GFP (Roberts et al., 2017). These plasmids are available through Addgene, and the stem cell lines we've made using them are available as part of the Allen Cell Collection at Coriell. More information and images generated from these cell lines can be found in the Cell Catalog on our data portal. The plasmids should be functional in many different human cell types, and while most of the genes we've tagged are widely recognized markers of cellular structures, we provide guidance on how to make your own gene-tagging constructs below. Read on to find out how these plasmids work and to discover more resources for your genomic tagging experiments.

Fluorescent tagging strategy

The first step in creating the plasmids was to identify the gene of interest and decide which terminus, N or C, to tag and which short amino acid linker to use. These decisions were made based on known functional aspects of the protein gleaned from the literature and engagement with researchers who study the protein or structure of interest. For a C-terminal tag, we inserted a linker and GFP tag at the end of the last exon of the gene to ensure it would be transcribed and translated. For an N-terminally tagged protein, we used the same strategy, inserting the tag and the linker preceding the first exon of the gene. In this manner, we have created an initial set of 10 plasmids representing 10 human genes and will add more to the collection.

Designing the donor plasmid

A key feature of our methodology is the use of a donor plasmid that contains long stretches of DNA on either side of the fluorescent protein sequence that is homologous to the sequence into which it is being inserted: one thousand base pairs on each side, in fact. Later, when we transfect the cells, these large regions of homology enable us to effectively utilize the host cell's inherent homology directed

Plasmids for Endogenous Gene Tagging in Human Cells (CONT'D)

repair (HDR) process following the double strand break CRISPR/Cas9 makes at the target site. The whole segment—a GFP tag and 1kb of homologous DNA flanking both sides (about 2.7kb total)—is inserted into a plasmid backbone for delivery to the cells. As a precautionary measure, it's also good to make a few single base pair mutations in crRNA target sites found within the homology areas of the plasmid to prevent the plasmid from being cut and destroyed by CRISPR/Cas9 during transfection.

Making the double strand break and introducing the tag through homology directed repair

To introduce the fluorescent tag to the cells, we used CRISPR/Cas9 to make a precise cut in the genome near either the beginning or end of the gene of interest —in our case in the genome of an hiPS cell. While the majority of the time the genome will attempt to repair itself without outside influence (non-homologous end joining, or NHEJ), in approximately one percent of cases, the stem cell will seek out a template for the repair process. By flooding the cell with a plasmid containing regions of

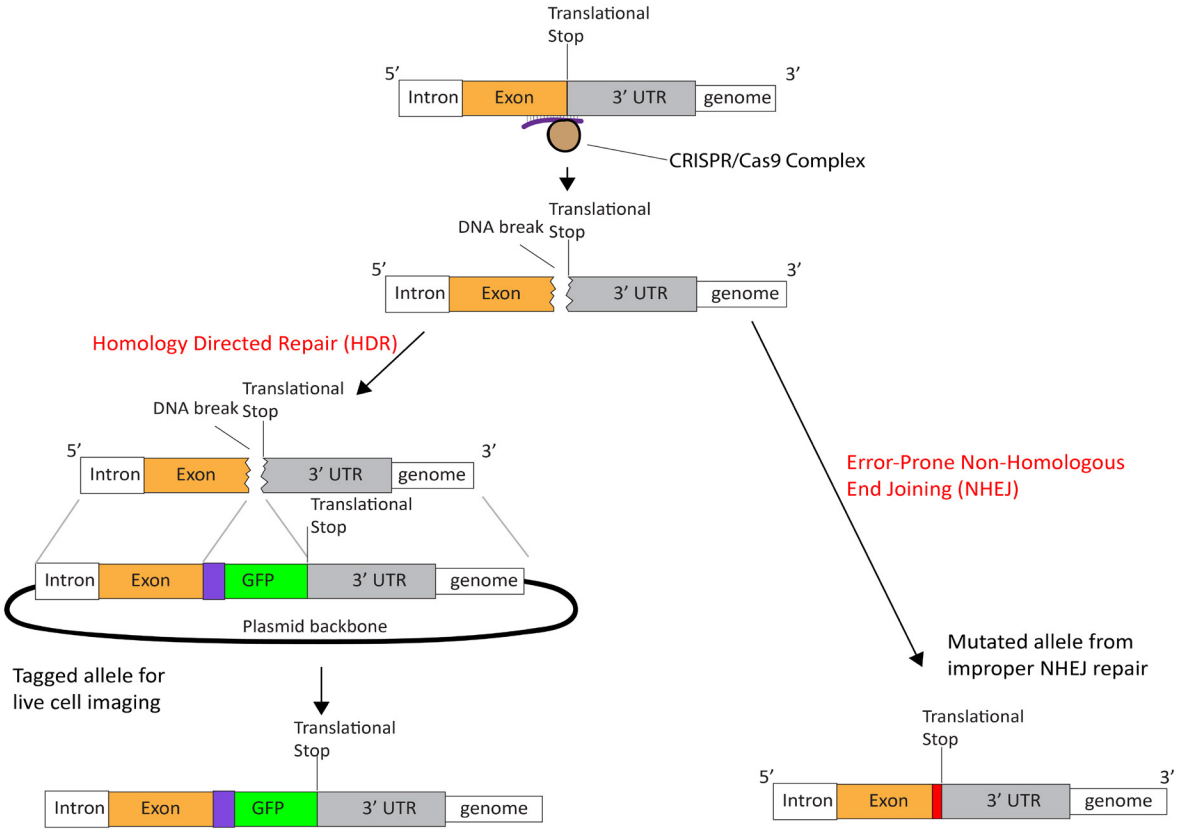


Figure 1: Steps to induce the double strand break and introduce the tag.

Plasmids for Endogenous Gene Tagging in Human Cells (CONT'D)

high homology to the gene with the cut, we trick the cell into using the introduced donor plasmid as a repair template. This results in the insertion of the GFP tag precisely where we need it, in this example, at the C-terminus of the gene of interest. To get the molecular components of CRISPR/Cas9 (Cas9 protein and the crRNA-tracrRNA complex) and the donor plasmid into the cells we use electroporation, a technique that briefly destabilizes the cell's outer membrane and allows the components to physically enter the cells so they can go to work.

Plasmids, plasmids everywhere

To find out more about gene edited hiPS cell lines generated using the Allen Institute for Cell Science Plasmid Collection, please visit the Cell Catalog on the Allen Cell Explorer, where you can also find the certificate of analysis provided with each cell line. The editing strategy and design described above can be used to create similar donor plasmids for introducing tags into your gene of interest.

Researchers have developed a variety of endogenous gene tagging systems. Addgene's CRISPR-based Protein Tagging collection includes systems for tagging mammalian cell lines, Drosophila cells, and *C. elegans*. In addition to generating fluorescent protein fusions, plasmids are available for tagging with epitope tags (e.g. FLAG and STREP) and auxin-inducible degrons (AID). Further, the Yamamoto PITCh system plasmids provides an alternative to HDR as the double-strand break repair pathway, relying instead on microhomology-mediated end-joining (MMEJ).

To find out more about gene edited hiPS cell lines generated using these plasmids, please visit the Cell Catalog on the Allen Cell Explorer, where you can also find the certificate of analysis provided with each cell line. The editing strategy and design described above can be used to create similar donor plasmids for introducing tags into your gene of interest.

Endogenous gene tagging with fluorescent proteins as described here enables live cell imaging to study cellular structures and processes and can lead to the better understanding of cell biology. We hope that these high-quality, highly specific plasmids, as well as the instructions on how to make your own, will make it easier for you to use genomic tagging in your experiments.

References

Roberts, Brock, et al. "Systematic gene tagging using CRISPR/Cas9 in human stem cells to illuminate cell organization." bioRxiv (2017): 123042.

Visualizing Genomic Loci with CRISPR-Sirius

By Jennifer Tsang | Apr 18, 2019

The catalytically dead Cas9 protein (dCas9) is well known for its ability to bind DNA targets without changing them. Thus, it has been widely adapted for CRISPR activation and inhibition experiments. But over the past few years, dCas9 has become a robust visualization tool to study the spatial and temporal arrangement of chromosomes and how these arrangements may affect nuclear processes. The latest of these visualization tools is CRISPR-Sirius, the brightest CRISPR-based tool to date for visualizing genomic loci in living cells.

Previous iterations of genomic loci visualization tools using CRISPR

In 2015, Thoru Pederson's lab at the University of Massachusetts Medical School developed a dCas9-based multicolor labeling system to visualize chromosomal loci in human cells (Ma et al., 2015). To do so, they fused a different fluorescent protein to various Cas9 variants, SpCas9, NmCas9, and St1Cas9. These fusions localized to specific loci based on the guide RNA sequence.

A year later, they developed CRISPRainbow (Ma et al., 2016), a set of gRNA scaffolds that contain two hairpin sequences (either MS2, PP7, or boxB) that each recruits a different fluorescent protein (BFP, GFP, or RFP) to the genomic loci of interest. Pairs of identical hairpins within one gRNA will result in a blue, green, or red readout. But gRNAs containing two different hairpins will produce a secondary color (yellow, cyan, or magenta) depending on the hairpin combination. A gRNA containing all three hairpins produces a white color. While CRISPRainbow allows for simultaneous visualization of sites on different chromosomes in living cells, the system requires high-copy chromosome-specific loci in the genome and the guides are relatively unstable.

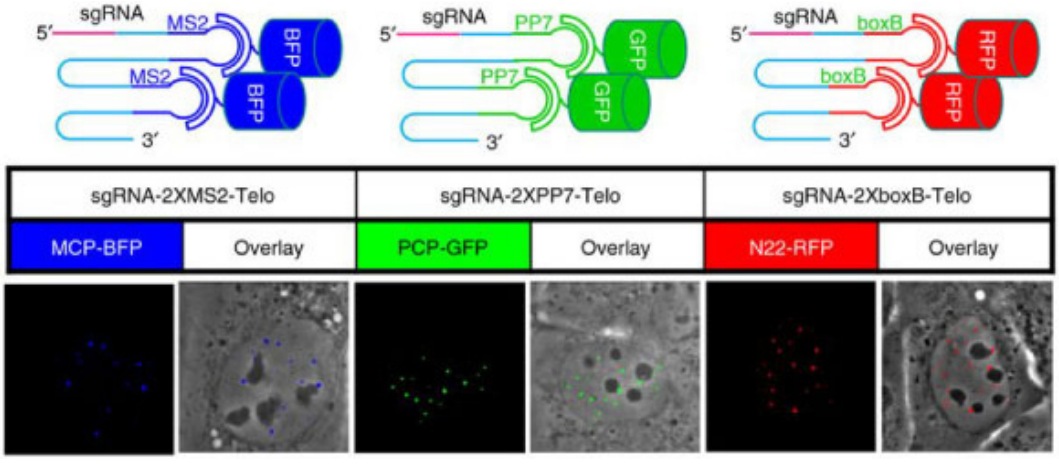


Figure 1: CRISPRainbow labeling. Two MS2, PP7, or boxB sequences inserted into the sgRNA generates blue, green, or red colors for labeling. A combination of MS2, PP7, or boxB sequences in the same sgRNA will generate additional colors. Image from Ma et al., 2016 with permission.

Visualizing Genomic Loci with CRISPR-Sirius (CONT'D)

Developing CRISPR-Sirius

Recently, Hanhui Ma and the Thoru Pederson lab developed CRISPR-Sirius (Ma et al., 2018). They improved on previous CRISPR-based loci visualization tools by increasing gRNA stability.

First, they began with the Broccoli aptamer system to visualize the gRNA in living cells. Broccoli aptamers are nucleic acid sequences that bind to the small molecule, DFHBI-1T to induce its fluorescence. They found that aptamer insertion in the gRNA tetraloop (a four base hairpin loop motif within the gRNA) resulted in a more stable gRNA than when they inserted the aptamer at the 3' end of the gRNA.

Because of the success of the MS2 and PP7 RNA hairpin sequences in CRISPRainbow, they knew they could be good candidates for further development. They inserted an octet of MS2 hairpins into the gRNA tetraloop and went through subsequent design iterations to increase stability and minimize misfolding and recombination during virus production needed to deliver CRISPR-Sirius into cells. They named this guide sgRNA-Sirius-8XMS2 which will recruit the HaloTag with red fluorescent ligands to the loci of interest.

Dual color detection

In order to label two loci simultaneously, Ma and Pederson created another sgRNA containing the PP7 hairpin sequence (sgRNA-Sirius-8XPP7) to recruit GFP to the loci of interest. To determine how the sgRNA-Sirius-8XPP7 and sgRNA-Sirius-8XMS2 scaffolds performed, they compared them head-to-head against CRISPRainbow. Both tools were targeted to the FBN3 intronic repeat which contains 22 copies of the gRNA target site. In addition to increased brightness, they also found that the signal to noise ratio can be increased from 2-fold (by CRISPRainbow) to 3-fold with CRISPR-Sirius.

Triple color detection

With a third sgRNA containing alternating MS2 and PP7 sequences (sgRNA-Sirius-4X(MS2-PP7)) that result in yellow fluorescence at the loci of interest, the team could visualize three loci and measure inter-locus distances. They tested this system out on chromosome 19 by targeting 46 repetitive loci that contained ≥20 copies of the gRNA target site. Of these, 26 of the 46 loci gave a detectable signal in human U2OS cells. Based upon this, they chose an intergenic DNA region (IDR3) as a reference loci to measure inter-locus distances of loci pairs from intergenic regions, intronic regions, and a pericentromeric region. All pairs were easily visualized and results were consistent with copy number variation (CNV) analysis of the U2OS cells.

Visualizing Genomic Loci with CRISPR-Sirius (CONT'D)

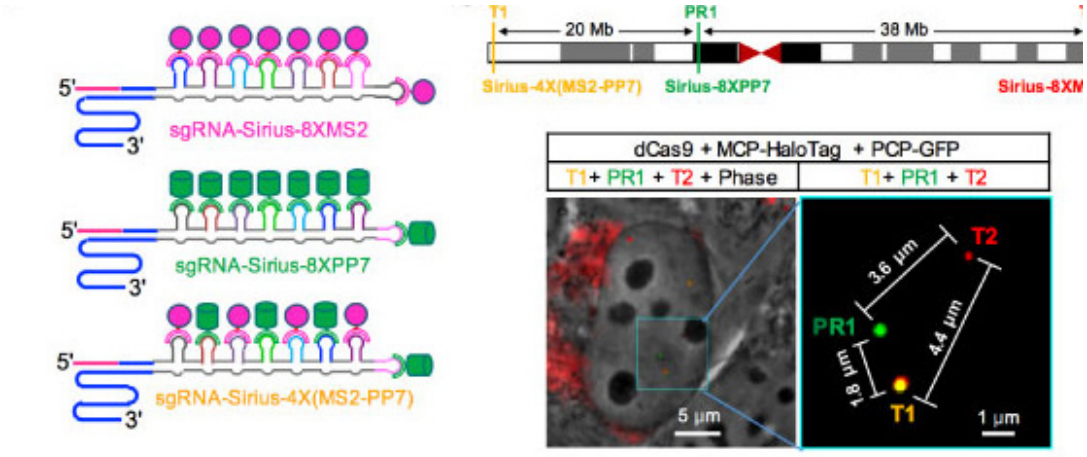


Figure 2: CRISPR-Sirius tricolor labeling. (Left Diagram of sgRNA-Sirius-8XMS2, sgRNA-Sirius-8XPP7, and sgRNA-Sirius-4X(MS2-PP7) showing recruitment of different fluorescent proteins. (Right) CRISPR-Sirius used to visualize relative locations of the subtelomeric region of the pericentromeric region (PR1), the p-arm (T1), and the q-arm (T2) of chromosome 19. Image from Ma et al., 2018 with permission.

Why try CRISPR-Sirius?

With all the tools out there to visualize genomic loci, what makes CRISPR-Sirius unique? Unlike fluorescent *in situ* hybridization, CRISPR-based visualization tools can be used in live cells. Compared to its predecessor CRISPRainbow, CRISPR-Sirius is brighter making it suitable for observing low-copy genomic loci or single copy genes. If you want to visualize genomic loci brightly, give CRISPR-Sirius a try (find the plasmids here).

References

Ma, Hanhui et al. "Multicolor CRISPR labeling of chromosomal loci in human cells." Proceedings of the National Academy of Sciences of the United States of America 112 10 (2015): 3002-7. PubMed PMID: 25713381 PubMed Central PMCID: PMC4364232.

Ma, Hanhui et al. "Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow." Nature Biotechnology 34 (2016): 528-530. PubMed PMID: 27088723 PubMed Central PMCID: 4864854.

Ma, Hanhui et al. "CRISPR-Sirius: RNA scaffolds for signal amplification in genome imaging." Nature Methods 15 (2018): 928-931. PubMed PMID: 30377374.

Multiplexed Capture of Promoter-enhancer 3D Chromatin Structures Using CRISPR

By Beth Kenkel | Dec 7, 2017, Updated Jul 2, 2020

Promoters may be the star of gene regulation, but enhancers and chromatin looping play important supporting roles. Enhancers are cis regulatory DNA sequences that, when bound by transcription factors, can increase a gene's transcription. Sometimes enhancers are located thousands of base pairs away from the gene they regulate, but are brought in proximity by the looping of chromatin, the complex of DNA and proteins that forms chromosomes.

These long range DNA interactions are often detected with chromosome conformation capture (3C) based methods, such as 5C or Hi-C (Han et al., 2018). Think of 3C methods as taking a snapshot of chromatin looping with the photo "developed" by analyzing DNA sequencing data. While "pictures" generated with current 3C methods provide useful information about chromatin interactions, they are grainy so it's hard to make out the details. To increase the resolution, the Xu lab created a dCas9-based CAPTURE (CRISPR Affinity Purification *in situ* of Regulatory Elements) method. The original CAPTURE method was published in 2017 and addressed many of the drawbacks of 3C methods, but could only detect chromatin interactions at one location in the genome at a time (Liu et al., 2017). Recently the Xu lab developed CAPTURE 2.0, an updated version of CAPTURE that detects chromatin interactions at hundreds of loci at once (Liu et al., 2020).

CRISPR Affinity Purification *in situ* of Regulatory Elements: The original CAPTURE method

The original CAPTURE method requires the creation of a stable cell line that co-expresses three components:

1. dCas9 with an added biotin acceptor site
2. BirA, a biotin ligase
3. a gRNA(s) targeting a single genomic location of interest

When these three components are co-expressed, the gRNA targets the dCas9 to the loci of interest and dCas9 is biotinylated by BirA. From here, the CAPTURE method is similar to other 3C methods: crosslinking captures chromatin interactions, a restriction enzyme digest cuts DNA into smaller pieces, and proximity ligation creates chimeric circles of DNA that are distant linearly but close together in space. The chromatin "picture" is then developed by detecting these chimeric DNA fragments by PCR or next-generation sequencing.

A drawback of the original CAPTURE method is that a new cell line has to be created for each genomic loci of interest. This is cumbersome and also makes it hard to compare results across multiple targets because different CAPTURE cell lines will have varying levels of dCas9 and gRNA expression. Additionally, CAPTURE 3C-based sequencing requires a large number of cells (~$5x10^7$). Together, these

Multiplexed Capture of Promoter-enhancer 3D Chromatin Structures Using CRISPR (CONT'D)

two requirements prevent the use of CAPTURE with primary cells or rare cell populations.

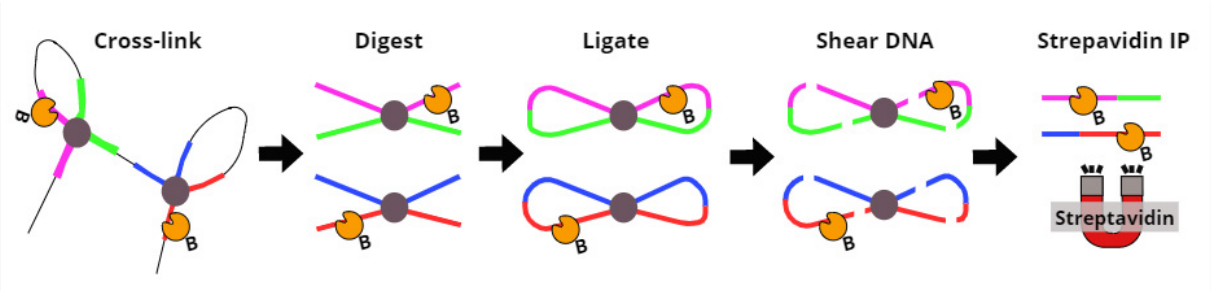


Figure 1: The CAPTURE workflow.

Key components of the CAPTURE 2.0 system

To improve upon the original CAPTURE method, the Xu Lab made the following changes to create CAPTURE 2.0:

1. **Replace the biotinalatable tag on dCas9 with a BioTAP-tag.** The BioTAP-tag is a 69 amino acid long biotinylation targeting sequence that is recognized by biotin ligases normally expressed in eukaryotic cells. Using this tag reduces the number of CAPTURE components that need to be delivered to cells from three to two.
2. **Lentiviral delivery of BioTAP-tagged dCas9 and gRNAs.** Using lentiviral delivery of the two remaining CAPTURE 2.0 components eliminates the need to create stable cell lines and allows CAPTURE 2.0 to be used with primary cells.

When compared to the original CAPTURE, CAPTURE 2.0 had a ~14-fold increase in its capture rate of chromatin interactions at a promoter in the well-characterized beta globin locus. The two methods had similar rates of specificity for the DNA sequence targeted in the experiment.

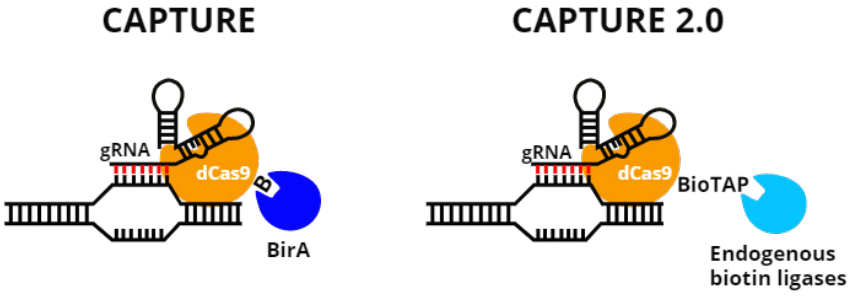


Figure 2: Comparing CAPTURE and CAPTURE 2.0.

Multiplexed Capture of Promoter-enhancer 3D Chromatin Structures Using CRISPR (CONT'D)

Table 1: Comparison of CAPTURE and CAPTURE 2.0.

	Original CAPTURE	CAPTURE 2.0
Biotinylation enzyme	BirA	Endogenous biotin ligase
Expression system	Stable cell line expressing 1) FLAG-biotin-tagged dCas9, 2) BirA, and 3) one or more gRNAs	Two separate lentiviruses delivering 1) a BioTAP-tagged dCas9 and an eGFP tag, and 2) gRNAs
Number of loci targeted	One	Many
Sample type(s)	Cell lines	Cell lines, primary cells

Applications of CAPTURE 2.0

CAPTURE 2.0's increased rate of detecting chromatin interactions opens the door for answering more questions about chromatin conformations. The Xu lab presented three ways to use CAPTURE 2.0:

1. Multiplex capture of chromatin interactions

The Xu lab performed a proof-of-concept experiment to show that CAPTURE 2.0 could detect multiple chromatin interactions at once. Chromatin interactions at five sites in the well-characterized beta-globin locus were analyzed with CAPTURE 2.0, with two gRNAs targeting each site. They found that DNA sequences targeted by the gRNAs were the top enriched sequences in the experiment and there was little enrichment of DNA targeted by a negative control gRNA or predicted gRNA off-target sites. Long-range chromatin interactions identified with the multiplexed CAPTURE 2.0 largely replicated interactions previously identified with the original CAPTURE method. Both CAPTURE methods outperformed other 3D chromatin methods such as ChIA-PET or Hi-C, by increasing both the number of unique chromatin interactions identified and the level of on-target enrichment.

2. Determining the spatial and hierarchical organization of chromatin interaction networks

A second application for CAPTURE 2.0 is the construction of spatial and hierarchical chromatin interaction networks. Think of these chromatin interactions networks as a social network. They can tell you things like which enhancers are interacting with which promoters, what part of a gene enhancers like to interact with, and how many promoters or genes each enhancer is interacting with. Super enhancers are particularly good models for building chromatin interaction networks since they are a large cluster of enhancers. CAPTURE 2.0's multiplexing and high resolution allowed the Xu Lab to simultaneously characterize the 'social networks' of 157 super enhancers.

Multiplexed Capture of Promoter-enhancer 3D Chromatin Structures Using CRISPR (CONT'D)

3. Measuring temporal changes in chromatin interactions

A third application for CAPTURE 2.0 is detection of temporal changes in chromatin conformation, like those that occur during development and cellular differentiation. Using a cell line model, the Xu Lab tracked enhancer-promoter interactions during erythroid differentiation.

CAPTURE 2.0 captures chromatin's dynamic role in gene regulation

The high resolution and multiplexed chromatin 'photos' generated with CAPTURE 2.0 allow us to uncover new details and patterns in chromatin interactions. The chromatin story told by CAPTURE 2.0 is important, but it's not the only plot line in gene regulation. When chromatin conformation data is compared with changes in gene expression, chromatin accessibility, and epigenetic modifications, we can start to build a timeline of events that surround changes in gene expression and determine who's a key driver of these changes. By generating a more vivid chromatin picture, CAPTURE 2.0 re-writes the script on gene regulation.

References

Han, J., Zhang, Z., & Wang, K. (2018). 3C and 3C-based techniques: the powerful tools for spatial genome organization deciphering. Molecular Cytogenetics, 11. https://doi.org/10.1186/s13039-018-0368-2

Liu X, Zhang Y, Chen Y, Li M, Zhou F, Li K, Cao H, Ni M, Liu Y, Gu Z, Dickerson KE, Xie S, Hon GC, Xuan Z, Zhang MQ, Shao Z, Xu J (2017) *In Situ* Capture of Chromatin Interactions by Biotinylated dCas9. Cell 170:1028–1043.e19 . https://doi.org/10.1016/j.cell.2017.08.003

Liu, X., Chen, Y., Zhang, Y., Liu, Y., Liu, N., Botten, G.A., Cao, H., Orkin, S.H., Zhang, M.Q., & Xu, J. (2020). Multiplexed capture of spatial configuration and temporal dynamics of locus-specific 3D chromatin by biotinylated dCas9. Genome Biology, 21. https://doi.org/10.1186/s13059-020-01973-w

Chapter 10

Expression and Delivery

Ribonucleoprotein (RNP) Delivery

By Andrew Hempstead | Sep 6, 2018

CRISPR has greatly enhanced the ability of scientists to make genomic alterations, bringing about a revolution in genome engineering, with new techniques rapidly being developed. Performing a CRISPR experiment requires delivery of, at minimum, two components: the Cas9 protein and a guide RNA (gRNA) targeting your genomic site of interest. This is commonly performed by transfecting cells with a plasmid, such as PX459, which encodes Cas9 and contains a site for inserting a custom gRNA. While this methodology has proven to be incredibly valuable to scientists, there are some potential complications that must be considered when using this method:

1. Cells must be amenable to transfection or viral transduction
2. Appropriate promoters must be chosen for both Cas9 and gRNA expression
3. Plasmid DNA may be incorporated into the genome
4. Off-target effects can occur due to prolonged Cas9 expression
5. The requirement for Cas9 transcription and translation delays editing

Cas9-gRNA ribonucleoproteins

One alternative approach, which avoids many of these complications, is to directly deliver a ribonucleoprotein (RNP), consisting of the Cas9 protein in complex with a targeting gRNA, to your cells of interest.

There are many advantages to using RNPs for CRISPR experiments. The RNP method can often be used in cells that are difficult to transfect, such as primary cells. Using RNPs can also alleviate difficulties with protein expression that occur in cells where common eukaryotic promoters (such as CMV or EF1A promoters found in many CRISPR plasmids) are not expressed. Lastly, because this method does not require the delivery of foreign DNA, and the Cas9-gRNA RNP is degraded over time, using RNPs may limit the potential for off-target effects.

As with plasmid based techniques, RNPs can be used to deliver Cas9 variants and to direct homology directed repair (HDR). In the former, modified versions of nuclease dead (dCas9) are used for a variety of experimental purposes, such as modifying gene expression, without altering the genome. In HDR experiments, Cas9-gRNA RNPs are used in combination with a template DNA to generate a specific mutation at a genomic site of interest. As new CRISPR techniques and tools are developed, it is likely that RNPs will play a major role in the application of these approaches.

Preparing the Cas9-gRNA ribonucleoprotein

The first step in this type of CRISPR experiment is the generation of the RNP complex. While one option is to purchase the Cas9 protein and a gRNA from a commercial vendor, this can often be

Ribonucleoprotein (RNP) Delivery (CONT'D)

expensive. An alternative approach is to express and purify His-tagged Cas9 from *E. coli* using a plasmid such as pET-28b-Cas9-His from Alex Schier's lab. gRNAs can be *in vitro* transcribed from ssDNA, which can be generated by commercial vendors such as IDT. These two components are then incubated together to form the RNP. Detailed protocols outlining these steps have been made publicly available by both the Jacob Corn and Alex Schier laboratories.

Delivering Cas9-gRNA ribonucleoproteins

An advantage of performing CRISPR experiments using RNPs is the variety of methods that can be used to deliver a Cas9-gRNA RNP. This advantage has allowed scientists to perform *in vitro* and *in vivo* CRISPR studies in experimental systems that may not be amenable to plasmid-based methods.

One of the most common techniques for delivery of RNPs is electroporation (A in the figure), which generates pores in the cell membrane, allowing for entry of the RNP into the cytoplasm. In addition to the use of this technique in cell culture, it has also been applied to genome editing of mouse zygotes, through a process known as CRISPR-EZ (CRISPR RNP Electroporation of Zygotes) (Chen et al. 2016). For CRISPR experiments that involve HDR, electroporation can be combined with cell-type specific reagents in a technique known as nucleofection, which forms pores in the nuclear membrane, allowing for entry of a DNA template.

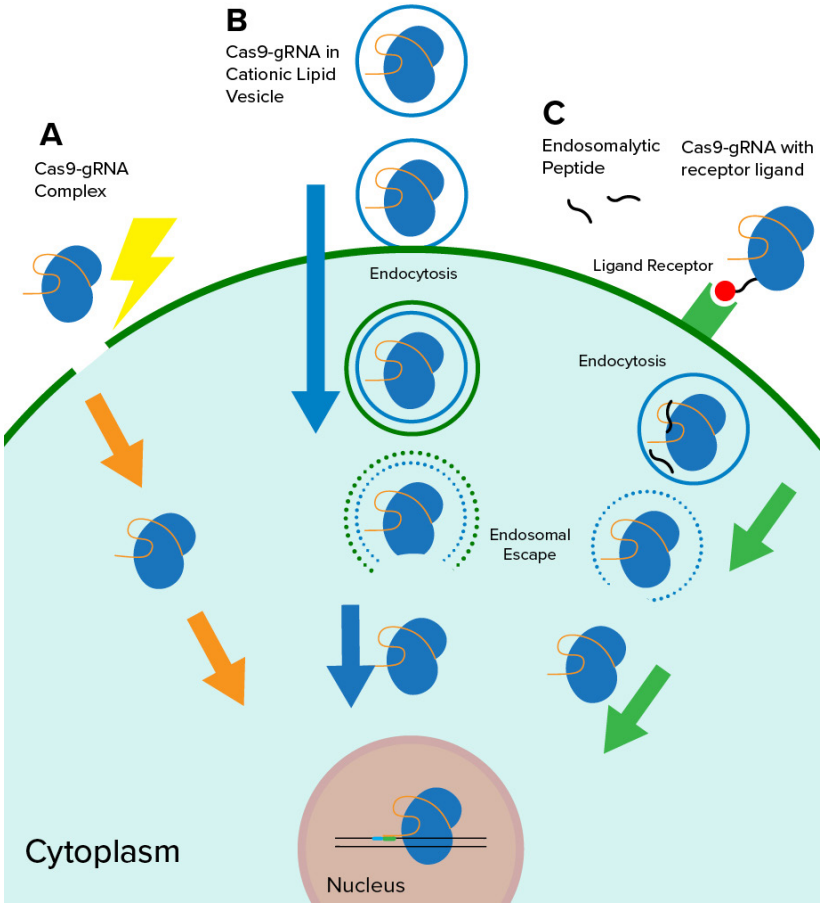


Figure 1: Three ways to delivery Cas9-gRNA ribonucleoproteins include (A) Cas9-gRNA complex by electroporation, (B) Cas9-gRNA in cationic lipid vesicles, (C) Cas9-gRNA binds a ligand to enter through a receptor.

Multiple methods have been developed to enable Cas9-gRNA RNP-based genome editing *in vivo*. Though these techniques are in the early stages of development, they may serve as the basis for the use of CRISPR to treat genetic diseases. David Liu's Lab has demonstrated the use of

Ribonucleoprotein (RNP) Delivery (CONT'D)

cationic lipid-mediated methods (B in the figure) to deliver Cas9-gRNA RNPs to hair cells in the mouse inner ear (Zuris et al. 2015). A more targeted approach using RNPs has recently been described *in vitro*, with the hope of eventually using this method for *in vivo* applications. This technique utilizes Cas9 proteins harboring receptor ligands (C in the figure), which result in the internalization of Cas9-gRNA RNPs by specific cell types (Rouet et al. 2018). This was accomplished using a Cas9 mutant (Cas9M1C/C80S) that harbors two surface exposed cysteines, including the native C547, allowing for ligation to pyridyl disulfide-activated ligands.

Due to the concerns of off-target mutations and transgene integration in crops, delivery of Cas9-gRNA RNPs to plants, as an alternative to plasmid based techniques, has been an area of intense study. Two methods to accomplish this are Polyethylene Glycol (PEG) mediated transfection of protoplasts and biolistic bombardment of immature embryos (Liang et al. 2017). PEG-mediated methods, while technically less difficult to perform, often show low efficiency, while biolistic bombardment, which can have high efficiency, requires specialized equipment, such as a gene gun. Depending on the nature of the experiment, it is up to the researcher to determine which method is most appropriate for a given experiment.

Considerations for RNP-mediated CRISPR experiments

While using Cas9-gRNA RNPs for CRISPR experiments can reduce many potential complications, there are additional factors to take into account with these types of experiments. One important consideration is that the Cas9-gRNA RNP will be degraded by the cell over time. As a result, RNP-mediated CRISPR experiments may not be suitable for experiments where stable Cas9 expression is required, such as when a Cas9 fusion to a fluorescent protein is used to label nucleic acid targets. While Cas9-gRNA RNPs show promise for *in vivo* studies, little is known about the immunogenicity and stability of the Cas9 protein in the host. Indeed, research has shown that a Cas9 protein from Geobacillus stearothermophilus is more stable in human plasma than the standard spCas9 (Harrington et al. 2017). Lastly, as with any scientific research, choice of a technique often comes down to the researcher's familiarity with different systems. Using Cas9-gRNA RNPs in your experiments can require steps such as protein expression and purification, which will require additional laboratory reagents, equipment, and expertise.

Final thoughts

The Cas9-gRNA RNP method provides scientists with an additional means to deliver CRISPR components to their experimental systems of interest. Using this method can alleviate many issues that may arise when using plasmid-based systems, most notability the ability to perform CRISPR studies in cells that are difficult to transfect or transduce. While using Cas9-gRNA RNPs may require

Ribonucleoprotein (RNP) Delivery (CONT'D)

the production of the Cas9 protein and a gRNA in your laboratory, once these have been acquired, genomic editing can be performed rapidly and, in many cases, with reduced chance for off-target effects. As the CRISPR field continues to evolve, it is likely that the use of RNPs will play an important role in its continued advancement.

References

1. Chen, Sean, et al. "Highly efficient mouse genome editing by CRISPR ribonucleoprotein electroporation of zygotes." Journal of Biological Chemistry 291.28 (2016): jbc-M116. PubMed PMID: 27151215. PubMed Central PMCID: PMC4938170.

2. Zuris, John A., et al. "Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing *in vitro* and *in vivo*." Nature Biotechnology 33.1 (2015): 73. PubMed PMID: 25357182. PubMed Central PMCID: PMC4289409.

3. Rouet, Romain, et al. "Receptor-Mediated Delivery of CRISPR-Cas9 Endonuclease for Cell-Type-Specific Gene Editing." Journal of the American Chemical Society 140.21 (2018): 6596-6603. PubMed PMID: 29668265. PubMed Central PMCID: PMC6002863.

4. Liang, Zhen, et al. "Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes." Nature communications 8 (2017): 14261. PubMed PMID: 28098143. PubMed Central PMCID: PMC5253684.

5. Harrington, Lucas B., et al. "A thermostable Cas9 with increased lifetime in human plasma." Nature Communications 8.1 (2017): 1424. PubMed PMID: 29127284. PubMed Central PMCID: PMC5681539.

Mammalian Expression Systems and Delivery Methods

By Nicole Waxmonsky | Sep 24, 2016, Updated Sep 30, 2020 by Marcy Patrick

CRISPR technology has been widely adopted for genome editing purposes because it's cheaper, faster, and easier than prior editing techniques. More and more CRISPR tools are being published each month, making CRISPR a great choice for your next experiment!

Here, we'll provide an overview of some CRISPR mammalian expression systems, the typical applications for each, and potential delivery methods.

General considerations for planning your CRISPR experiment

As with any experiment, there are many factors that need to be considered during the planning process. For CRISPR experiments, the following framework can help get you started:

1. Determine what type of outcome you are trying to achieve: Do you want to permanently knock out or knock-in a gene? Do you want to activate or repress gene expression? Are you trying to create a single point mutation? Do you want to create a fusion with a reporter protein such as GFP? All of these outcomes can be most effectively achieved with different CRISPR components.
2. Select the appropriate CRISPR tools for your application: Wildtype Cas9 or Cas9 nickase are appropriate for generating knockouts or knock-ins, or introducing mutations and tags, while a catalytically dead dCas9 can be used in conjunction with activator or repressor domains to control gene expression. Base editors can help you make precise point mutations near the gRNA target site.
3. Choose an appropriate expression system and delivery method: Do you need stable integration or is transient expression sufficient? Which cell types will you be editing? Do you want to deliver the components as DNA, or would mRNA or protein delivery be more suitable?
4. Determine how you will evaluate the outcome: Will you be detecting insertions/deletions using a mismatch repair assay? Or is PCR followed by gel electrophoresis and/or next-generation sequencing more appropriate?

If you already have an end product in mind, steps 1 and 2 will generally be straightforward. Likewise for step 4, as this ties directly back to your specific goal. When thinking about step 3, however, you may be surprised at the number of options available - how do you choose between them?

One of the first steps is to identify what CRISPR components you will need to deliver. Minimally, one or more sgRNAs and Cas9 are required for any application. If you want to include a homology directed repair (HDR) template to create knock-ins, point mutations, or to add a tag, you will also need to deliver a donor plasmid or single-stranded DNA oligonucleotide, so you will need to make sure your expression system and delivery methods are compatible with all your components.

Next, consider the best form for those CRISPR components based on your model system. CRISPR

Mammalian Expression Systems and Delivery Methods (CONT'D)

reagents can be delivered via transfection, nucleofection, viral transduction or injection as either protein, RNA, or DNA. Finally, once you have identified the best expression system, you can then choose the best method for introducing those CRISPR components into your target cells.

Mammalian CRISPR expression systems

Each model system has its own best practices for efficient delivery of CRISPR components. If you are new to CRISPR or your model system, a good first step would be to consult the literature to see if anyone has published a protocol that would work for your system. Addgene has depositor submitted protocols and links to a CRISPR forum where you may be able to find information regarding your system of choice. The table below summarizes the various components included with each expression system as well as suitable applications.

Table 1: Various expression systems, their components and applications

Expression system	Component of system	Application
Mammalian expression vector	Promoter driving Cas9 expression can be constitutive or inducible. U6 promoter is typically used for gRNA. May contain reporter gene (e.g. GFP) to identify and enrich positive cells or a selection marker to generate stable cell lines.	Transient or stable expression of Cas9 and/or gRNA in a mammalian cell line that can be transfected at high efficiency.
Lentiviral transduction	Cas9 and gRNA can be present in a single lentiviral transfer vector or separate transfer vectors. May contain reporter gene (e.g. GFP) to identify and enrich positive cells. Likely contains a selection marker to generate stable cell lines. Packaging and Envelope plasmids provide the necessary components to make lentiviral particles.	Stable, tunable expression of Cas9 and/or gRNA in a wide variety of mammalian cell lines. Useful for difficult to transfect cell types and can be used *in vivo*. A common choice for conducting genome-wide screens using CRISPR.

Mammalian Expression Systems and Delivery Methods (CONT'D)

Expression system	Component of system	Application
AAV transduction	Only compatible with SaCas9 (packaging limit ~4.5kb). CRISPR elements are inserted into an AAV transfer vector and used to generate AAV particles.	Transient or stable expression of SaCas9 and/or gRNA. Infects dividing and non-dividing cells. AAV is least toxic method for *in vivo* viral delivery.
Cas9 mRNA and gRNA	Plasmids containing gRNA and Cas9 are used in *in vitro* transcription reactions to generate mature Cas9 mRNA and gRNA. RNA is delivered to target cells using microinjection or electroporation.	Transient expression of CRISPR components, expression decreases as RNA is degraded within the cell. Can be used for generating transgenic embryos.
Cas9-gRNA ribonucleoprotein complexes	Purified Cas9 protein and *in vitro* transcribed gRNA are combined to form a Cas9-gRNA complex and delivered to cells using cationic lipids.	Transient expression of CRISPR components, expression decreases as gRNA and Cas9 protein are degraded within the cell.

Delivery methods for mammalian cell lines

As mentioned above, the expression system you choose in many ways dictates the best method for introducing those CRISPR components into your target cells. Delivery of these components into mammalian cell lines is quite broad and includes several different methods, so we've summarized some key features of common delivery methods below:

Table 2: Chemical transfection

	Lipid mediated	Cationic polymers	Calcium phosphate
Cells	Transformed cell lines	Transformed cell lines	Transformed cell lines
Cargo	Nucleic acid or protein	Nucleic acids	Nucleic acids
Throughput	High	High	High
Efficiency	Medium	Medium	Medium
Difficulty	Easy	Easy	Easy

Mammalian Expression Systems and Delivery Methods (CONT'D)

Table 3: Physical transfection

	Electroporation/nucleofection	Microinjection
Cells	Many types	Single cells
Cargo	Nucleic acids or proteins	Nucleic acids or proteins
Throughput	High	Very low
Efficiency	High	Very high
Difficulty	Very easy	Very hard

Table 4: Viral delivery

	Lentivirus	Retrovirus	AAV
Cells	Dividing or nondividing	Dividing	Dividing or nondividing
Cargo	Nucleic acids	Nucleic acids	Nucleic acids
Throughput	Low	Low	Low
Efficiency	Variable	Variable	Variable
Difficulty	Hard	Hard	Hard

If your expression system is not well characterized in terms of CRISPR use, you will want to invest some time in optimizing and testing the efficiency of CRISPR delivery. There are a few plasmids at Addgene that have been published as CRISPR testing tools:

- Traffic Light Reporter System: Can evaluate lentiviral component delivery as well as genome repair by non-homologous end joining (NHEJ) or HDR.
- EGFP validation of sgRNAs: Can evaluate component delivery and sgRNA efficacy by cloning in genome target sequence into EGFP reporter.
- Target DNA reporter system: Can evaluate component delivery with validated tools.

Adenoviral Delivery of CRISPR/Cas9 Expands Genome Editing to Primary Cells

By Kendall Morgan | Sep 30, 2014, Updated Dec 10, 2020 by Benoit Giquel

Adenoviral vectors (AdVs) have been used for many years for *in vivo* delivery and gene therapy and represent one of the most studied viruses. Adenoviruses are non-enveloped double-stranded DNA viruses with a genome of around 36 kb in size, which does not integrate into the host genome, removing the risk of oncogenicity or genotoxicity seen with other vectors (Lukashev et al., 2016). The protein capsid encapsulating the genome is majorly composed of three proteins that can be easily modified in order to facilitate its entry into the target cells (Beatty et al., 2012). These features make AdV an attractive candidate for *in vivo* delivery of advanced genome editing machinery such as CRISPR-Cas.

Delivering gRNA and Cas9 using adenoviral vectors

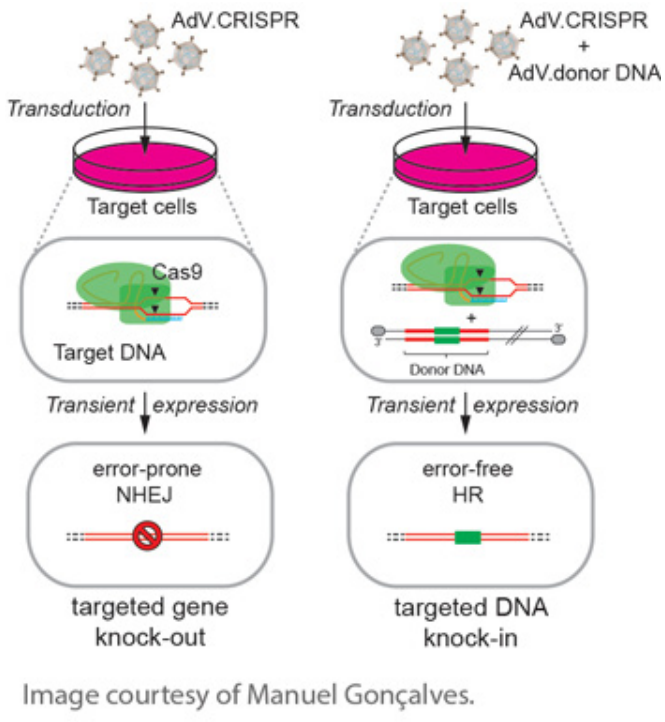


Figure 1: Adenoviral vectors can be used to deliver CRISPR components and a repair template (for HDR) to generate knock-outs (NHEJ) or knock-ins (HDR).

As compared to other viral vectors, AdVs can provide a large transgene packaging capacity. For instance, with a packaging capacity of about 6 kb, the 1st generation of AdVs already have a larger packaging capacity than AAV vectors. This capacity is large enough to carry the Cas9 gene and a gRNA expression cassette in one single viral particle. The earliest explorations of the use of AdV for *in vivo* CRISPR-Cas delivery utilized the Non Homologous End Joining (NHEJ) DNA repair mechanism to knock out genes as a result of an insertion, a deletion or a frameshift (Ding et al., 2014, Wang et al., 2015). In these studies, the strategy employed was to deliver a standard AdV vector to the targeted organs or tissues with Cas9 and a gRNA designed to target a mutated gene. In a Scientific Reports paper introducing the delivery method in 2014, Manuel Goncalves' lab reported that AdV-mediated transduction of gRNA:Cas9 ribonucleoprotein complexes into transformed and non-transformed cells yielded rates of targeted mutagenesis similar to those achieved by isogenic AdVs encoding TALENs targeting the same chromosomal region. The CRISPR/Cas9-derived RNA-guided nuclease-induced gene disruption frequencies in the various cell types ranged from 18% to 65% (Maggio et al., 2014).

An alternative approach to NHEJ is to utilize the Homology Directed Repair (HDR) to knock in

Adenoviral Delivery of CRISPR/Cas9 Expands Genome Editing to Primary Cells (CONT'D)

corrective genes. For this you would need to provide a DNA template to repair the double strand break via homology directed repair and you would probably need to use a two viral vector systems: one vector for Cas9/sgRNA and one vector for your specific DNA template. Using this two vectors system, Manuel Gonçalves' lab found in a second paper published in Nature Methods the same year that delivering RNA-guided nucleases or TALENs together with AdV donor DNA leads to a vast majority of AdV-modified human cells being subjected to scarless homology-directed genome editing (Maggio et al., 2014). Gonçalves said they attribute this phenomenon to the presence of terminal proteins capping the ends of linear double-stranded AdV genomes. Such protein-DNA structures presumably reduce the likelihood that donor DNA will interact with sporadic double-stranded chromosomal DNA breaks "that always happen naturally."

These adenoviral CRISPR/Cas9 genome editing tools developed and demonstrated by Manuel Gonçalves and his colleagues at Leiden University Medical Center are now available at Addgene along with a description of their experimental protocol. The three plasmids which have been deposited to Addgene are: pAdSh.PGK.Cas9, pAdSh.U6.gRNAS1, and pAdSh.U6.gRNAGFP.

Gonçalves says that advantages of AdVs include their episomal nature and very efficient introduction of DNA into therapeutically relevant, non-transformed mammalian cells. These viral vector systems also work equally well in dividing and quiescent, post-mitotic mammalian cells.

Applications of CRISPR delivery by adenoviral vectors

Since then, other groups have successfully developed and used AdV vectors to edit genes in several biological systems. It has been used for instance in inherited disorders such as B-thalassemia, Sickle Cells Disease, Duchenne Muscular Dystrophy, Alpha1-antitrypsin deficiency (most widely used so far) or for cancer gene therapy such as nonalcoholic steatohepatitis cancer. In deficiencies of plasma proteins where AdV vector delivery is often used, the CRISPR approach based on HDR has the advantage of being a potentially universal treatment (Stephens et al., 2019). If sufficient gene transfer and knock-in is achieved, producer cells could produce corrective amounts of protein for any such disease. This technology therefore represents a potential platform that could be realized via simple switching of the relevant gRNAs and donor DNA.

Adenoviral CRISPR/Cas9 genome editing tools can also be used to deliver Cas9 and DNA templates in somatic cells of adult animals in order to create animal models for particular cancer. That's exactly what the lab of Andrea Ventura at the Memorial Sloan Kettering Cancer Center did to create a model of Eml4–Alk-driven lung cancer (Maddalo et al., 2014). These tools are available at Addgene and could be useful to labs wishing to create this kind of animal models.

Adenoviral Delivery of CRISPR/Cas9 Expands Genome Editing to Primary Cells (CONT'D)

Start using adenoviral vectors with your CRISPR/Cas9 research!

To find more information about the adenoviral delivery of CRISPRS/Cas9 using the Gonçalves and Ventura labs' plasmids, including protocols, check out the plasmids at Addgene: pAdSh.PGK.Cas9 (expresses *S. pyogenes* Cas9 from the PGK promoter) and U6 promoter-driven guide RNA constructs, pAdSh.U6.gRNAS1 and pAdSh.U6.gRNAGFP, and also Adeno Cas9 and Adeno EA. Or if you're looking for a broader range of CRISPRs plasmid tools, find more plasmids, CRISPR technology guides, FAQs, and CRISPR resources on Addgene's CRISPR page.

References

Beatty MS, Curiel DT (2012) Adenovirus Strategies for Tissue-Specific Targeting. In: Applications of viruses for cancer therapy. Elsevier, pp 39–67

Ding Q, Strong A, Patel KM, Ng S-L, Gosis BS, Regan SN, Cowan CA, Rader DJ, Musunuru K (2014) Permanent Alteration of PCSK9 With *In Vivo* CRISPR-Cas9 Genome Editing. Circ Res 115:488–492 . https://doi.org/10.1161/circresaha.115.304351

Holkers M, Maggio I, Henriques SFD, Janssen JM, Cathomen T, Gonçalves MAFV (2014) Adenoviral vector DNA for accurate genome editing with engineered nucleases. Nat Methods 11:1051–1057 . https://doi.org/10.1038/nmeth.3075

Lukashev AN, Zamyatnin AA Jr (2016) Viral vectors for gene therapy: Current state and clinical perspectives. Biochemistry Moscow 81:700–708 . https://doi.org/10.1134/s0006297916070063

Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han Y-C, Ogrodowski P, Crippa A, Rekhtman N, de Stanchina E, Lowe SW, Ventura A (2014) *In vivo* engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. Nature 516:423–427 . https://doi.org/10.1038/nature13902

Maggio I, Holkers M, Liu J, Janssen JM, Chen X, Gonçalves MAFV (2014) Adenoviral vector delivery of RNA-guided CRISPR/Cas9 nuclease complexes induces targeted mutagenesis in a diverse array of human cells. Sci Rep 4: . https://doi.org/10.1038/srep05105

Stephens CJ, Lauron EJ, Kashentseva E, Lu ZH, Yokoyama WM, Curiel DT (2019) Long-term correction of hemophilia B using adenoviral delivery of CRISPR/Cas9. Journal of Controlled Release 298:128–141 . https://doi.org/10.1016/j.jconrel.2019.02.009

Wang D, Mou H, Li S, Li Y, Hough S, Tran K, Li J, Yin H, Anderson DG, Sontheimer EJ, Weng Z, Gao G,

Adenoviral Delivery of CRISPR/Cas9 Expands Genome Editing to Primary Cells (CONT'D)

Xue W (2015) Adenovirus-Mediated Somatic Genome Editing of Pten by CRISPR/Cas9 in Mouse Liver in Spite of Cas9-Specific Immune Responses. Human Gene Therapy 26:432–442 . https://doi.org/10.1089/hum.2015.087

Overcoming the AAV Size Limitation for CRISPR Delivery

By Mary Gearing | Jul 13, 2015, Updated Sep 16, 2020 by Beth Kenkel

CRISPR genome editing has quickly become a popular system for *in vitro* and germline genome editing, but *in vivo* gene editing approaches have been limited by problems with Cas9 delivery. Adeno-associated viral vectors (AAV) are commonly used for *in vivo* gene delivery due to their low immunogenicity and range of serotypes allowing preferential infection of certain tissues. However, packaging *Streptococcus pyogenes* (SpCas9) and a gRNA together (~4.2 kb) into an AAV vector is challenging due to its packaging capacity of AAV (~4.7 kb). While this approach has been proven feasible, it leaves little room for additional regulatory elements. Feng Zhang's group previously packaged Cas9 and multiple gRNAs into separate AAV vectors, increasing overall packaging capacity but necessitating purification and co-infection of two AAVs.

Cas9 orthologs: shorter, but just as potent and specific?

The two AAV strategies described above showed successful target modification, indicating that AAV is a good delivery vehicle for Cas9. However, to allow for more room for regulatory sequences while still fitting Cas9 and its gRNAs into one AAV, Cas9 must be made smaller. Previous attempts to "shrink" Cas9 include the use of St1Cas9 (~3.3 kb) from *Streptococcus thermophilus* and a rationally-designed truncated Cas9. Unfortunately, certain drawbacks limit the utility of these systems: St1Cas9 requires a very specific PAM sequence that limits the number of targetable loci, and truncated Cas9 has much lower efficiency than its wild-type counterpart.

In search of a smaller, but equally potent Cas9 protein, the Zhang lab took a different approach. They analyzed over 600 Cas9 orthologs and found that they could be divided into two groups: longer orthologs approximately 1,350 amino acids in size, which includes SpCas9, and shortern orthologs approximately 1000 amino acids in size. From the pool of shorter orthologs, only *Staphylococcus aureus* Cas9 (SaCas9, 1,053 amino acids) displayed cleavage activity in mammalian cells. SaCas9 produced indels at a similar efficiency to SpCas9, leading the group to focus their efforts on SaCas9 characterization for *in vivo* studies.

One of the pitfalls of CRISPR/Cas9 genome editing is the potential for off-target effects. To compare the off-target effects of SpCas9 and SaCas9, Zhang's group used an approach called BLESS (direct *in situ* breaks labeling, enrichment on streptavidin and next-generation sequencing). Using this sensitive method, they found that SaCas9 did not display higher levels of off-target activity than SpCas9, confirming its suitability for *in vivo* studies.

Testing AAV-SaCas9 *in vivo*

To test the efficiency of AAV-SaCas9 *in vivo*, the Zhang lab created an all-in-one SaCas9 and sgRNA construct using the liver-specific serotype AAV8. Since the efficiency of CRISPR/Cas9 genome editing

Overcoming the AAV Size Limitation for CRISPR Delivery (CONT'D)

varies across targets, they tested two genes in mice. For both genes, they saw indel formation and phenotypic changes as early as 1 week post-injection. Livers from these mice were histologically normal and liver injury markers were not increased compared to a control AAV-GFP. Not only did the AAV-SaCas9-sgRNA constructs mediate genome modification, but they did so without a substantial immune response or toxicity.

More AAV-based CRISPR systems

Smaller Cas proteins

SaCas9 isn't the only CRISPR enzyme that's small enough to package into AAV.

- At 984 amino acids in length, Cas9 from *Campylobacter jejuni* (CjCas9) is the smallest Cas9 ortholog characterized to date. Kim et al. successfully used CjCas9 with AAV to target genes in mouse muscle and eye tissue.
- *Neisseria meningitidis* Cas9 (NmeCas9) at 1,082 amino acids, can also be packaged into AAV. NmeCas9 has the added advantage that it can be turned off by anti-CRISPR proteins. The Sontheimer lab used NmeCas9 with AAV to edit two different disease-causing loci in mouse liver.
- Two other smaller Cas proteins include Cas12b and CasX. While there are no papers yet published that deliver these Cas proteins with AAV, at just 1,108 and 986 amino acids, respectively, the size of both is within AAV's packaging limit.
- For CRISPR-based RNA editing, the REPAIR (RNA Editing for Programmable A to I Replacement), system is also small enough to deliver with AAV. This system fuses catalytically dead dCas13b to the catalytic domain of RNA deaminase ADAR2. Constructs containing the ADAR2 truncation ADAR2DD(delta984-1090) are approximately 4.1 kb in length, allowing them to be packaged in AAV.

Split AAV approaches

While finding smaller Cas9 orthologs works for some applications, it doesn't eliminate the need to deliver larger cargo such as base editors and prime editors. While using AAV to deliver large transgenes might seem daunting, it's actually a challenge that the field has overcome with split AAVs. In general split AAVs break a large transgene in two pieces and package each piece into an individual AAV. When a cell is transduced by both AAVs, the full length gene and/or protein is reconstituted.

One way to reconstitute a split protein like Cas9 is to use split inteins. Split inteins are a pair of naturally occuring polypeptides that, when at the ends of two proteins, mediate protein trans-splicing, similar to an intron in pre-mRNA splicing. In 2016, Fine et al developed a proof-of-concept split intein SpCas9 which had modest editing rates in HEK-293T cells when compared to the full-length SpCas9. In 2016, Chew et al. developed a split intein spCas9-AAV toolbox that retains the gene-targeting capabilities of full-length SpCas9. This set of plasmids includes AAV-Cas9C-VPR for targeted gene activation. Split inteins have also been used to express base editors.

Overcoming the AAV Size Limitation for CRISPR Delivery (CONT'D)

References

Chew WL, Tabebordbar M, Cheng JKW, Mali P, Wu EY, Ng AHM, Zhu K, Wagers AJ, Church GM (2016) A multifunctional AAV–CRISPR–Cas9 and its host response. Nat Methods 13:868–874 . https://doi.org/10.1038/nmeth.3993

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339:819–823 .https://doi.org/10.1126/science.1231143

Cox DBT, Gootenberg JS, Abudayyeh OO, Franklin B, Kellner MJ, Joung J, Zhang F (2017) RNA editing with CRISPR-Cas13. Science 358:1019–1027 . https://doi.org/10.1126/science.aaq0180

Fine EJ, Appleton CM, White DE, Brown MT, Deshmukh H, Kemp ML, Bao G (2015) Trans-spliced Cas9 allows cleavage of HBB and CCR5 genes in human cells using compact expression cassettes. Sci Rep 5: . https://doi.org/10.1038/srep10777

Garneau JE, Dupuis M-È, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, Moineau S (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468:67–71 . https://doi.org/10.1038/nature09523

Ibraheim R, Song C-Q, Mir A, Amrani N, Xue W, Sontheimer EJ (2018) All-in-one adeno-associated virus delivery and genome editing by *Neisseria meningitidis* Cas9 *in vivo*. Genome Biol 19: . https://doi.org/10.1186/s13059-018-1515-0

Kim E, Koo T, Park SW, Kim D, Kim K, Cho H-Y, Song DW, Lee KJ, Jung MH, Kim S, Kim JH, Kim JH, Kim J-S (2017) *In vivo* genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. Nat Commun 8: . https://doi.org/10.1038/ncomms14500

Levy JM, Yeh W-H, Pendse N, Davis JR, Hennessey E, Butcher R, Koblan LW, Comander J, Liu Q, Liu DR (2020) Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. Nat Biomed Eng 4:97–110 . https://doi.org/10.1038/s41551-019-0501-5

Liu J-J, Orlova N, Oakes BL, Ma E, Spinner HB, Baney KLM, Chuck J, Tan D, Knott GJ, Harrington LB, Al-Shayeb B, Wagner A, Brötzmann J, Staahl BT, Taylor KL, Desmarais J, Nogales E, Doudna JA (2019) CasX enzymes comprise a distinct family of RNA-guided genome editors. Nature 566:218–223 . https://doi.org/10.1038/s41586-019-0908-x

Overcoming the AAV Size Limitation for CRISPR Delivery (CONT'D)

Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O (2014) Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. Cell 156:935–949. https://doi.org/10.1016/j.cell.2014.02.001

Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F (2015) *In vivo* genome editing using *Staphylococcus aureus* Cas9. Nature 520:186–191 . https://doi.org/10.1038/nature14299

Senís E, Fatouros C, Große S, Wiedtke E, Niopek D, Mueller A-K, Börner K, Grimm D (2014) CRISPR/Cas9-mediated genome engineering: An adeno-associated viral (AAV) vector toolbox. Biotechnology Journal 9:1402–1412 . https://doi.org/10.1002/biot.201400046

Strecker J, Jones S, Koopal B, Schmid-Burgk J, Zetsche B, Gao L, Makarova KS, Koonin EV, Zhang F (2019) Engineering of CRISPR-Cas12b for human genome editing. Nat Commun 10: .https://doi.org/10.1038/s41467-018-08224-4

Swiech L, Heidenreich M, Banerjee A, Habib N, Li Y, Trombetta J, Sur M, Zhang F (2014) *In vivo* interrogation of gene function in the mammalian brain using CRISPR-Cas9. Nat Biotechnol 33:102–106 . https://doi.org/10.1038/nbt.3055

Wang D, Zhang F, Gao G (2020) CRISPR-Based Therapeutic Genome Editing: Strategies and *In Vivo* Delivery by AAV Vectors. Cell 181:136–150 . https://doi.org/10.1016/j.cell.2020.03.023

Nanoblades: Tiny CRISPR Ninjas for Genome Editing Difficult Cells

By Beth Kenkel | Sep 26, 2019

CRISPR is a simple and versatile tool for genome engineering, but its utility is dependent on its ability to infiltrate cells. Options for CRISPR delivery include plasmid transfection, RNP electroporation, and viral transduction; but these methods aren't stealthy enough to gain access to some cells and tissues, such as human induced pluripotent stem cells (hiPSCs). Nanoblades, a new CRISPR delivery method developed by the Ricci Lab and the T. Ohlmann Lab, adds a covert tool to the CRISPR tool box.

Nanoblades are engineered murine leukemia virus (MLV)-like particles loaded with Cas9-gRNA ribonucleoproteins (RNPs). According to the authors, they named their tool Nanoblades because they think of these particles as tiny DNA cutting ninjas that deliver CRISPR to multiple types of targets: primary cells, embryos, and animals. Nanoblades can also carry different types of CRISPR cargo including RNPs for indel formation via NHEJ, precise modification with Homology Directed Repair (HDR), CRISPR activation or repression, transgenic mouse line creation, and *in vivo* gene editing of mice, with the potential to be used for other types of CRISPR applications. Let's take a look at the key components of Nanoblades, how their editing efficiency and off-target effects compare to existing technologies, and their use for HDR editing and mouse-model generation.

Key components of Nanoblades

Nanoblades are a Cas9 RNP delivery vehicle that's based on mouse leukemia virus (MLV) virus-like particles (VLPs). VLPs lack viral genomes which means they are non-infectious and non-replicative. Nanoblades are produced by transfecting the following four components into 293T cells.

MLV GAG-POL polyprotein

Nanoblades are formed when viral structural proteins, such as the MLV protein Gag, multimerize and spontaneously assemble into particles at the cell membrane. Mature Nanoblades then bud into the cell culture supernatant and can be purified by centrifugation.

Gag-Cas9 fusion protein

The MLV Gag protein fused to a Flag-tagged SpCas9 is incorporated into Nanoblades, just like the unmodified MLV Gag protein. After Gag-Cas9 is incorporated, the MLV protease, supplied as part of the MLV Gag-Pol polyprotein, cleaves the proteolytic site connecting the two proteins, thus releasing Cas9. Nanoblades can also deliver Cas9 variants, such as a transcriptional activator dCas9-VPR, if these variants are fused to the MLV Gag in place of SpCas9.

gRNA

A gRNA targeting a DNA sequence of interest is expressed in 293T cells and loaded into Nanoblades

Nanoblades: Tiny CRISPR Ninjas for Genome Editing Difficult Cells (CONT'D)

by association with the the Cas9 portion of the Gag-Cas9 fusion protein. A single Nanoblade particle can be loaded with 1-4 different gRNAs.

Viral envelope proteins

To alter the cell tropism of Nanoblades, different viral envelope protein(s) are expressed to pseudotype Nanoblades. The team found that Nanoblades pseudotyped with a mix of VSV-G and the baboon endogenous retrovirus Rless glycoprotein (BaEVRless) had high rates of transduction for most cell types they tested.

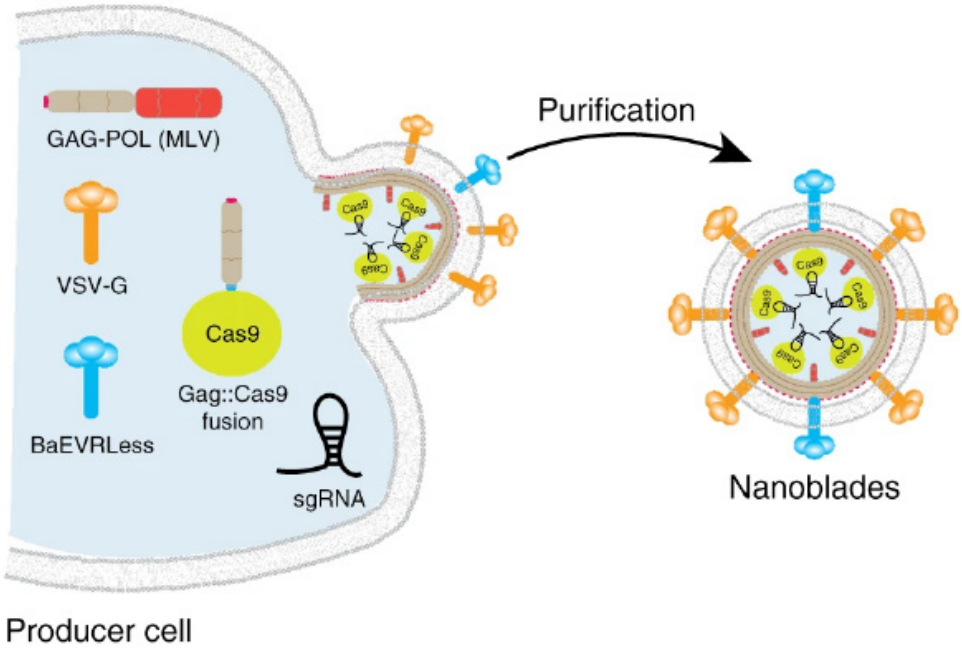


Figure 1: The key components of Nanoblades. Image from Mangeot et al., 2019.

Genome editing primary cells with Nanoblades

Primary cells are often refractory to gene delivery methods, but the researchers were able to edit hiPSCs, mouse bone marrow (BM) cells, primary human hepatocytes, and hematopoietic stem cells (HSCs) using Nanoblades. Overall, editing efficiencies were between ~50-70% for these four cell types (Table 1). Additionally, hiPSCs and mouse BM cells still behaved like iPSCs and BM cells after Nanoblade treatment as measured by their maintenance of pluripotency marker expression and response to LPS stimulation, respectively.

Nanoblades: Tiny CRISPR Ninjas for Genome Editing Difficult Cells (CONT'D)

Table 1: Editing efficiency using Nanoblades in various cell types.

Primary cell type	Gene targeted by guide	Editing efficiency
hiPSCs	EMX1	67%
Mouse BM cells from GFP transgenic mouse	GFP	~60-65%
Human hepatocytes	Myd88	~50%
Human HSCs	Myd88	~50%

Comparing Nanoblades to existing CRISPR delivery methods

When compared head-to-head with existing CRISPR delivery technologies, Nanoblades had higher rates for editing efficiency and lower rates of off-target effects. Nanoblades generated up to 76% editing efficiency in mouse BM cells when using a gRNA targeting the Fto gene, while editing was not detected in cells electroporated with RNPs using the same guide. Nanoblades also had lower off-target effects when compared to DNA transfection of gRNA and Cas9 expression plasmids. These results are in line with previous research showing that the method of Cas9 and gRNA delivery can influence the level of off-target effects, with electroporation of RNPs generally having lower off-target edits compared to DNA transfection.

Homology directed repair (HDR) with Nanoblades

CRISPR-based HDR allows for the precise insertion of a DNA sequence ranging from a single nucleotide change to large insertions like the addition of a fluorophore or tag. To utilize HDR, a donor template containing the desired sequence flanked by homology arms must be delivered to cells along with the gRNA and Cas9. An "all-in-one" Nanoblade which carries a gRNA, Cas9, and a HDR donor template DNA can be generated by incubating Nanoblade particles with polybrene, a cationic polymer. Polybrene has long been used to increase retroviral transduction efficiency by neutralizing the charge repulsion between virions and the cell surface, and also helps form DNA and retroviral-particle or VLPs complexes. The team found that "all-in-one" Nanoblades facilitated HDR with either a ssDNA or dsDNA donor template.

In vivo editing of mouse zygotes and mice with Nanoblades

CRISPR has also been extensively used to create transgenic mice by microinjecting zygotes with CRISPR components. However, this sometimes requires intracellular injections into the cytoplasm or pronucleus of the zygote which is both technically challenging and can negatively impact zygote viability. Nanoblades, however, avoid this pitfall since they are injected into the space surrounding the

Nanoblades: Tiny CRISPR Ninjas for Genome Editing Difficult Cells (CONT'D)

zygote, where they can then fuse with the target cell membrane, and transfer their CRISPR cargo into the zygote. As a proof-of-principle, the team used Nanoblades to disrupt the tyrosinase (Tyr) gene in mouse zygotes, which results in albinism in mice. Of the eight mice born from Tyr Nanoblade treated zygotes, five had detectable Tyr editing at the phenotypic and genotypic levels. The editing efficiencies of the chimeric and fully albino mice ranging from 11% to 100%, a level similar to other approaches.

Nanoblades can also be used to non-invasively gene edit the liver of adult mice. When mice were IV injected with Nanoblades targeting the hydroxyphenylpyruvate dioxygenase (HPD) gene, all nine treated mice had ~10% editing in their livers, while non-injected and control gRNA injected mice had no detectable editing. Additionally, off-target spleen editing was not detected and Nanoblade injection did not induce morbidity. Together these results demonstrate that Nanoblades are a potential alternative method for generating transgenic mice strains and for non-invasive *in vivo* gene editing.

The bottom line on Nanoblades

Nanoblades are tiny DNA-cutting ninjas which can stealthily deliver many types of CRISPR editing cargo to difficult to target cells. When compared to existing Cas9 delivery methods, Nanoblades have just as good if not better on-target editing efficiencies and have lower off-target edits. Another great thing about Nanoblades? They're easily produced in a BSL2 lab, making them a versatile and accessible genome editing tool. Get Nanoblades for your research!

References

Mangeot, Philippe E., et al. "Genome editing in primary cells and *in vivo* using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins." Nature communications 10.1 (2019): 45. PubMed PMID: 30604748. PubMed Central PMCID: PMC 6318322.

Chapter 11

Using CRISPR in Model Organisms

Generating Mouse Models Using CRISPR/Cas9

By Wenning Qin and Haoyi Wang | Jul 12, 2016, Updated Oct 1, 2020 by Aliyah Weinstein

CRISPR/Cas9 is revolutionizing the mouse gene-targeting field. Mice have long been extremely useful in the lab – they are relatively small and easy to work with, making them the go-to choice for studying mammalian biology. Similar to any model, mice are not without their problems, but much genetic and physiological data have been accumulated over the years using them. Indeed, the future of mouse work is bright as it is now easier than ever to manipulate the mouse genome using CRISPR/Cas9.

Similar to the human genome, the mouse genome is made up of 3×10^9 nucleotides (nt), and encodes 23,000 or so genes. It would be great if we could just go in and quickly manipulate individual mouse genes and study their function in health and disease, but, until recently, it just wasn't that easy. In the 1980s, gene targeting technology was invented to introduce specific changes into the mouse genome. With these early technologies, a researcher would first introduce a mutation into a mouse embryonic stem cell (ESC) line, enrich and select for cells that had successfully incorporated the desired mutation, and then derive mice from these engineered ESCs. To do so, the engineered ESCs would be injected into developing mouse embryos, the embryos allowed to develop into chimeric mice (with a fraction of the cells in the adult mice derived from the engineered ESCs), and chimeric adults mated to produce completely transgenic offspring. Although powerful, this technology is cumbersome to use, not that efficient, often takes more than a year to generate a mouse model, and its success is not always guaranteed.

To learn more about this process check out our Mouse Modeling blog posts: Part 1: Genetically Engineered Mice and Part 2: Breeding and Crossing Mice.

My first CRISPR mouse experiment

Everything changed with the advent of CRISPR in 2013 (Cong et al., 2013)! Haoyi got his first exposure to CRISPR when he used this new technology to engineer mouse models (Wang et al., 2013, Yang et al., 2013). Like everyone else in the world, I (Wenning) was fascinated by CRISPR and rushed to test it. I still remember the day, March 16, 2014, that I got my first preliminary results from a CRISPR genome editing experiment. I was in my office and opened a sequence file. Then I saw it, the chromatogram showing the telltale signs of a mutation introduced by CRISPR. The sequence started clean and then, about 100 bps into the run, it got "dirty", with multiple sequence traces overlapping each other! When the count was done, 75% of the mice in this experiment showed a mutation! I looked out my window and I saw the Dorr Mountain with the rest of the skyline of Acadia National Park. I was confused. It was supposed to be hard. I was mentally prepared to go through a learning curve that would result in success only after many attempts. Was I really successful in my first attempt at using CRISPR?

Later on, I found out that my experience was shared by many others around the world. Yes, finally, we have a technology, CRISPR, that is simple in concept, straightforward to use, and robust in performance. In its natural setting, CRISPR-Cas9 is an acquired immune system in bacteria and

Generating Mouse Models Using CRISPR/Cas9 (CONT'D)

archaea. As you know well if you've been following the Addgene blog, it has been repurposed for genome editing in eukaryotes, with the most widely used CRISPR genome editing system derived from *Streptococcus pyogenes*. An alternative Cas9, from Streptococcus aureus, is as effective at editing mouse embryos as Cas9 from *S. pyogenes* and has the advantage of being smaller (Zhang et al., 2016). For editing the genome, the CRISPR/Cas system makes use of 3 components, a guide RNA (gRNA) of about 125 nt that specifies the target, the Cas9 endonuclease that creates the DNA double-strand break (DSB) at the target site, and a donor oligonucleotide or plasmid as the repair material if needed (for knock in models).

CRISPR mouse model basics

To create a mouse model, the gRNA, Cas9, and donor oligonucleotide or plasmid components are brought together and microinjected into either the pronucleus or the cytoplasm of fertilized mouse eggs. Or, to avoid handling embryos *ex vivo*, the components can be electroporated into the oviduct of pregnant females, a technique called Genome editing via Oviductal Nucleic Acids Delivery, or GONAD (Gurumurty et al., 2016). AAV can also be used as a vector to deliver CRISPR/Cas9 either through injection into either an oocyte or into the oviduct of pregnant females (Yoon et al., 2018, Mizuno et al., 2018). Alternatively, the gRNA, Cas9, and donor oligonucleotide can be electroporated into the mouse zygote (Qin et al., 2015).

When inside the zygotes, the gRNA will seek out its target among the 3 X 10^9 nt of genetic content in the mouse genome and the Cas9 enzyme will make a cut at the target site. This is when it gets exciting – the cell sends out an "SOS" signal and cellular repair mechanisms rush in to repair the damage. If all they can do is stitch the two broken ends

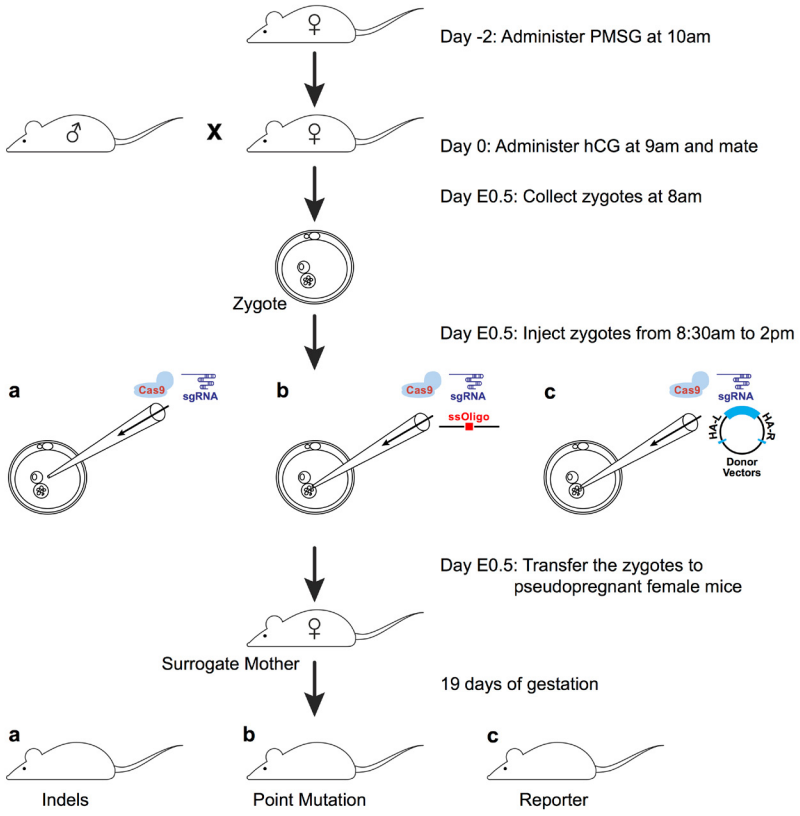


Figure 1: Process for generating genome-edited mouse lines using CRISPR/Cas9. Purified CRISPR/Cas9 components are directly injected into mouse zygotes and zygotes implanted into a surrogate mother. Offspring are then screened for the desired mutations. Figure adapted from Qin et al., 2016.

Generating Mouse Models Using CRISPR/Cas9 (CONT'D)

together through non-homologous end joining (NHEJ), this will leave behind a "scar", with nucleotides missing or added at the broken ends and, as such, cripple the gene. However, if repair material is provided (in the form of an oligonucleotide or plasmid), precise changes can be made in the genome via the homology directed repair pathway (HDR), be it a single nucleotide change, insertion of a reporter gene, or replacement of the murine sequence with a human gene.

Advantages of using CRISPR for mouse genome engineering

From the very start, generating mouse models using CRISPR is easier than more conventional methods. Time and money savings come from the fact that CRISPR is so efficient that you can inject the reagents directly into fertilized mouse eggs, circumventing the need for enrichment and selection offered by mouse ESCs. For example, we often screen only 15-25 mice when generating knockout models with CRISPR and find that many, if not all, the mice carry a frameshift mutation. For knock in with a donor oligonucleotide, we aim to generate 50 to 100 mice and are usually successful in deriving mice carrying the intended mutation. With a donor plasmid, the outcome is less predictable. In our best case, we saw that 2 out of 3 mice carried a 5 kb insertion in the ROSA locus as assessed by Southern blot.

From the efficiency discussed above extends 4 major advantages of CRISPR mouse editing when compared to more conventional methods. First, one can work with almost any strain of mice, as compared with conventional gene targeting, which is limited to a few strains, including 129 and C57BL/6, for which we have germline competent ESC lines. Second, the process is much quicker. It takes 3 months to generate founder mice using CRISPR, as compared with 8 to 10 months going through conventional gene targeting. It also costs less. With CRISPR, it may cost about $5,000 to generate founder mice, while with conventional gene targeting, the cost may be $50,000. Finally, compared to conventional methods of genome engineering, CRISPR is much more efficient at generating the desired mutant offspring, with 85% success compared to 50% when using ESC (Cohen, 2016).

Disadvantages of using CRISPR for generating mouse models

Although CRISPR is incredibly useful for generating mutations by NHEJ and generating small mutations with HDR, when it comes to larger scale genome editing, such as replacement of a mouse gene with its human ortholog (greater than 5 kb), it remains to be seen whether CRISPR is as robust as conventional gene targeting. CRISPR tools were recently used to introduce a 25 kb human gene into the mouse genome. However, with only 3.4% of founders establishing germline transmission of the transgene, this technique is still in need of improvement. Also when working with CRISPR, one must beware that not all gRNAs are created equal. Some work better than others. There are a variety of resources

Generating Mouse Models Using CRISPR/Cas9 (CONT'D)

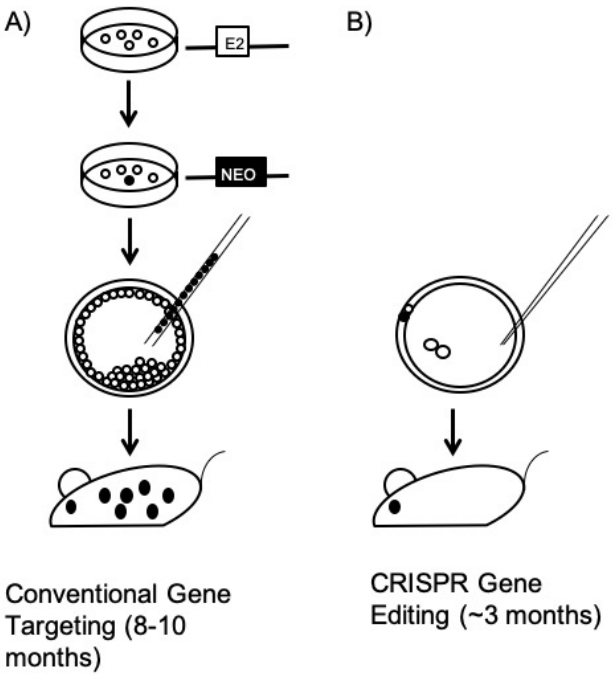


Figure 2: In conventional gene editing (A), you must first generate your desired mutation in mouse ES cells, select for the mutation and inject the ES cells into a mouse embryo. This laborious process takes longer (~8-10 months) than CRISPR gene editing in mice (B). With CRISPR, you can inject CRISPR components (see Figure 1) directly into a zygote and get your desired mouse with the appropriate gene edits in much less time (~3 months). Image courtesy of Wenning Qin.

online to guide you in gRNA design. We have been using Benchling, but there are many other gRNA design tools available. Finally, always remember that you are working with RNAs, which are prone to degradation by the omnipresent RNase. You may want to declare your bench "RNase free" and refrain from talking to friends and colleagues while working with these reagents. Other than that, be nice to your microinjectionist, who has the strategically important job of delivering the payload!

Last but not least, when using CRISPR, remember to appreciate the fact that it was first discovered in an obscure bacterium and that we still have much to learn from biology in all its forms!

References

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339:819–823 . https://doi.org/10.1126/science.1231143

Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE (2014) Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. Nat Biotechnol 32:1262–1267 . https://doi.org/10.1038/nbt.3026

Gurumurthy CB, Takahashi G, Wada K, Miura H, Sato M, Ohtsuka M (2016) GONAD: A Novel CRISPR/Cas9 Genome Editing Method that Does Not Require *Ex Vivo* Handling of Embryos. Current Protocols in Human Genetics 88: 15.8.1-15.8.12. https://doi.org/10.1002/0471142905.hg1508s88

Leidy-Davis T, Cheng K, Goodwin LO, Morgan JL, Juan WC, Roca X, Ong ST, Bergstrom DE (2018) Viable Mice with Extensive Gene Humanization (25-kbp) Created Using Embryonic Stem Cell/Blastocyst and CRISPR/Zygote Injection Approaches. Sci Rep 8: 15028. https://doi.org/10.1038/s41598-018-33408-9

Generating Mouse Models Using CRISPR/Cas9 (CONT'D)

Mizuno N, Mizutani E, Sato H, Kasai M, Ogawa A, Suchy F, Yamaguchi T, Nakauchi H (2018) Intra-embryo Gene Cassette Knockin by CRISPR/Cas9-Mediated Genome Editing with Adeno-Associated Viral Vector. iScience 9:286–297 . https://doi.org/10.1016/j.isci.2018.10.030

Qin W, Kutny PM, Maser RS, Dion SL, Lamont JD, Zhang Y, Perry GA, Wang H (2016) Generating Mouse Models Using CRISPR/Cas9 Mediated Genome Editing. Current Protocols in Mouse Biology 6:39–66 . https://doi.org/10.1002/9780470942390.mo150178

Qin W, Dion SL, Kutny PM, Zhang Y, Cheng AW, Jillette NL, Malhotra A, Geurts AM, Chen Y-G, Wang H (2015) Efficient CRISPR/Cas9-Mediated Genome Editing in Mice by Zygote Electroporation of Nuclease. Genetics 200:423–430 . https://doi.org/10.1534/genetics.115.176594

Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R (2013) One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering. Cell 154:1370–1379 . https://doi.org/10.1016/j.cell.2013.08.022

Yoon Y, Wang D, Tai PWL, Riley J, Gao G, Rivera-Pérez JA (2018) Streamlined *ex vivo* and *in vivo* genome editing in mouse embryos using recombinant adeno-associated viruses. Nat Commun 9: 412. https://doi.org/10.1038/s41467-017-02706-7

Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R (2013) One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Cell 153:910–918 . https://doi.org/10.1016/j.cell.2013.04.025

Zhang X, Liang P, Ding C, Zhang Z, Zhou J, Xie X, Huang R, Sun Y, Sun H, Zhang J, Xu Y, Songyang Z, Huang J (2016) Efficient Production of Gene-Modified Mice using *Staphylococcus aureus* Cas9. Sci Rep 6: 32565. https://doi.org/10.1038/srep32565

CRISPR Methods for Bacteria

By Mary Gearing | Mar 3, 2016, Updated Sep 28, 2020 by Will Arnold

Although CRISPR systems were first discovered in bacteria, most CRISPR-based genome engineering has taken place in other organisms. In many bacteria, unlike other organisms, CRISPR-induced double stranded breaks are lethal because the non-homologous end-joining (NHEJ) repair pathway is not very robust. In many cases, homology-directed repair does not function effectively either, but scientists have devised means of co-opting phage genetic systems to facilitate homologous recombination in bacteria. These quirks change the way CRISPR-mediated genome engineering functions in bacteria, but have no fear - plasmids from Addgene depositors are making it easier than ever to use CRISPR in *E. coli* and other bacterial species. Read on to learn about the tools available for bacteria and some of the applications for which they've been used.

The beginnings of bacterial CRISPR engineering

One common way of accomplishing bacterial genome engineering is done with recombineering, a technique that utilizes phage recombination machinery to promote homologous recombination of linear DNA fragments. Since recombineering does not contain a selection step for successful modifications, efficiency can be low, especially for larger modifications.

What's the solution to this inefficiency? Use CRISPR to make it a selectable process! As NHEJ is ineffective in bacteria, CRISPR-induced double stranded breaks (DSB) are lethal. Addgene depositor Luciano Marraffini's lab took advantage of this lethality to design the first synthetic bacterial CRISPR system in *E. coli*. The system available from Addgene consists of two plasmids:

- pCas9: carries Cas9 and chloramphenicol resistance
- pCRISPR: carries a spacer targeting the gene of interest and kanamycin resistance

E. coli carrying the phage recombineering machinery are first electroporated with pCas9. Then, pCRISPR is introduced along with an oligonucleotide repair template.

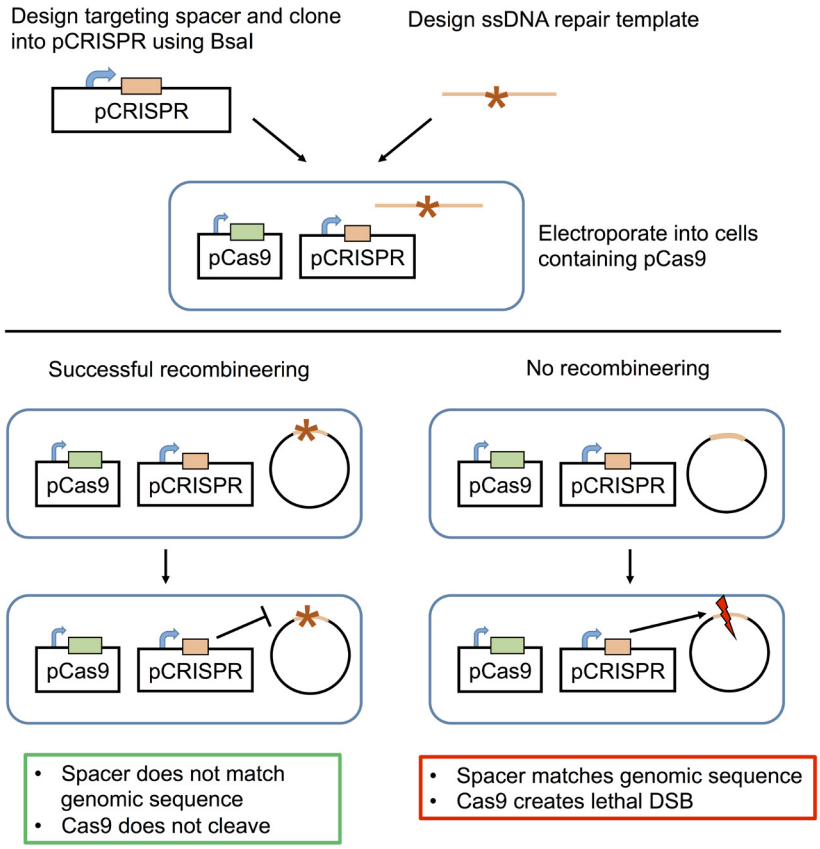


Figure 1: Schematic of pCRISPR and pCas recombineering.

CRISPR Methods for Bacteria (CONT'D)

Through recombineering, the locus of interest is modified to match the repair template, and the locus cannot be recognized by the spacer-derived crRNA. However, if recombineering is unsuccessful and the wild-type sequence persists, Cas9 will cleave the gene of interest, inducing a lethal DSB.

This system is distinct from those used in eukaryotes in that CRISPR isn't the primary editing force. In contrast, in *E. coli*, CRISPR is primarily a means of selection that targets cells in which homologous recombination has not occurred. This powerful negative selection system ensures high editing efficiency; the only non-edited cells to survive have inactivating mutations in the Cas9 or spacer sequence, and these rare events are easily detectable using PCR. The system also functions in *S. pneumoniae* and can be used to generate multiple mutations simultaneously (Jiang et al. 2013).

Genome editing tools for bacteria

CRMAGE system combines CRISPR and recombineering-based MAGE

The Nielsen lab's CRMAGE system is a fast, multiplexable method that combines CRISPR and recombineering-based MAGE (Multiplex Automated Genome Engineering) technology (Ronda et al., 2016). pMA7CR_2.0 expresses lambda Red and Cas9, which are separately inducible by L-arabinose and anhydrotetracycline (aTet), respectively. pMAZ-SK contains an aTet-inducible gRNA and a backbone-targeting gRNA cassette for plasmid curing through "self-destruction" after induction with L-rhamnose and aTet. CRMAGE is much more efficient than traditional recombineering, with 96-99% efficiency for point mutations and 66% efficiency for small insertions. Multiplexing of two targets simultaneously is possible with efficiency >70%. CRMAGE is an incredibly fast protocol, with only 5 hours incubation time needed for a single round of editing, and a subsequent curing protocol that requires only 2-3 hours incubation.

E. coli and *T. citrea* scarless editing plasmids

Sheng Yang's lab describes a two-plasmid system that combines recombineering with CRISPR to create a system for scarless, iterative genome engineering (Jiang et al., 2015). pCas contains Cas9 and the phage recombination gene lambda Red. pTargetF contains the specific gRNA(s), and the repair template is supplied as a dsDNA fragment. Gene deletion efficiency is as high as ~69%, but insertion efficiency varies with the length of homology supplied with the template (40 bp - 6% vs 400 bp - 28%.) Each round of editing takes two days, and the pTargetF and pCas plasmids can be cured from the bacteria via non-selection and growth at 37 °C, respectively. Although developed in *E. coli*, the system was used successfully in *Tatumella citrea*, another species of Enterobacteriaceae, without the need for modification. This finding suggests that the system may be functional in most Enterobacteriaceae.

CRISPR Methods for Bacteria (CONT'D)

pCRISPomyces for editing in *Streptomyces*

Streptomyces bacteria produce a wide variety of bioactive natural products. To easily explore and engineer pathways within this genus, Huimin Zhao's lab created two "pCRISPomyces" systems for use in *Streptomyces* (Cobb et al., 2015). pCRISPomyces-1 includes Cas9, a tracrRNA, and a CRISPR array, while pCRISPomyces-2 contains Cas9 and a gRNA cassette. The simpler system of pCRISPomyces-2 displays a higher editing efficiency, perhaps due to its condensed design. For both systems, custom spacers/gRNAs are easily inserted using BbsI and Golden Gate Assembly. Either plasmid can be also linearized with XbaI to insert extra elements, like a repair template, using Gibson Assembly or restriction enzyme cloning. *Streptomyces* bacteria are more recombinogenic than *E. coli*, so this system functions more similarly to CRISPR/Cas9 systems adapted for eukaryotes in that Cas9-mediated cleavage induces HDR to directly modify a given gene.

CRISPR-tranposons

By combining CRISPR editing and transposons, the Sternberg and Zhang labs developed CRISPR-transposons. The Sternberg lab's system called INTEGRATE (Insertion of Transposable Elements by Guide RNA-Assisted Targeting) and the Zhang lab's system called CAST (CRISPR-associated Transposase) both enable gRNA-directed transposition. INTEGRATE consists of four major components including (1) a CRISPR RNA, (2) four proteins forming the QCascade DNA-targeting module with the crRNA, (3) three transposase proteins, and (4) the donor DNA. By using a multi-spacer CRISPR array, CRISPR-transposons can be multiplexed.

Transcriptional repression (CRISPRi) in bacteria

As RNA interference does not function in bacteria, most efforts to regulate gene expression were limited to inducible promoters or direct gene knockout. In contrast, CRISPR offers a much more user-friendly way to modulate gene expression. The Marraffini lab and Qi lab developed early systems for *E. coli*; while the Marraffini lab used a native minimal CRISPR array (Bikard et al., 2013), the Qi lab employed a gRNA-based design more familiar to those using CRISPR in eukaryotes (Qi et al., 2013). As in other systems, catalytically dead (dCas9) targeted to a promoter or gene body can repress transcription by physically blocking the elongation complex from binding the DNA or extending the transcript. Different Cas orthologues and subtypes function somewhat differently and can be more effective targeting either the promoter or the coding sequence and the template or non-template strand, among other parameters.

Stanley Qi lab's CRISPRi plasmids

Stanley Qi's lab demonstrated that a minimal, regulatable CRISPR system effectively silenced

CRISPR Methods for Bacteria (CONT'D)

transcription of one or more genes in *E. coli*. and in a mammalian cell line (Qi et al., 2013). Two important features of this early system are its regulatable nature and its ability to multiplex repression. Using a Tet responsive promoter to drive expression of a catalytically dead Cas9 enzyme allows reversible repression based on the presence of the inducer, in this case anhydro-tetracycline (aTc). By cloning in two tandem copies of the gRNA cassette, one can reliably knockdown expression of multiple target transcripts simultaneously.

Marraffini lab's CRISPRi/a plasmids

In addition to the early work of the Marraffini lab pioneering genome editing using CRISPR in bacteria, they also led early efforts creating tools for CRISPRi and CRISPRa (Bikard et al., 2013). By targeting either a catalytically dead Cas9 (dCas9) alone or a dCas9 fused to the omega subunit of RNA polymerase, the lab successfully repressed or activated transcription of select genes in *E. coli*. Furthermore, to demonstrate wider applicability they demonstrated success of Cas9 mediated repression in the Gram-positive bacterium, *Streptococcus pneumoniae*.

Robert Husson lab's *M. tuberculosis* CRISPRi plasmids

A group from the Husson lab used CRISPRi to study essential genes in *M. tuberculosis*. dCas9 and a gRNA are expressed on pRH2502 and pRH2521, respectively. Both dCas9 and the gRNA are Tet-inducible, and a gRNA can be cloned easily into pRH2521 using BbsI. This system results in 80-90% RNA knockdown across multiple gRNAs per gene (Singh et al. 2016).

Sarah Fortune's lab *M. tuberculosis* CRISPRi plasmids

Noting that the more common spCas9 system has generally low efficiencies or can be proteotoxic when used in Mycobacterial species, Sarah Fortune's lab screened 11 other Cas9 orthologs to identify a more efficient, well tolerated, Cas enzyme. They identified the Cas9 orthologue from *Streptococcus thermophilus*, Cas9sth1 as a robust and highly active enzyme for transcriptional silencing *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. In addition to their identification of a new CRISPRi tool, they also optimized the system by modulating the PAM to further tune efficiency of knockdown and developed new, more tightly controlled, Tet responsive promoters to avoid leaky activation of the system.

Mobile CRISPRi for Gram-negative and Gram-positive bacteria

Jason Peters, Oren Rosenberg, and Carol Gross used two different genetic systems, targeting either Gram-negative bacteria or Gram-positive firmicutes, to bring CRISPRi to a much wider swath of bacterial species (Peters et al., 2019). The systems differ on how exactly the CRISPRi locus is

CRISPR Methods for Bacteria (CONT'D)

introduced (Tn7 in Gram-positives and ICEbs1 in Gram-negatives) but share the common CRISPR-dCas9 system for targeted repression of gene expression. They successfully knock down expression of essential and non-essential genes in a wide variety of bacteria including in pooled, genome-wide, formats. Interested in knowing more? Check out our Mobile CRISPRi blog post.

CRISPR activation (CRISPRa)

While the activation of transcription is a more challenging problem than repression, scientists have deposited several plasmid tools that can be used in this way. As described in the above CRISPRi section, the Marraffini Lab published a robust tool for CRISPRi and CRISPRa using RNAP-Omega-dCas9 fusions. Here's some more:

SoxS based CRISPRa

In the search for a CRISPR activation method in *E. coli*, Jesse Zalatan's lab used an approach to identify known proteins involved in the activation of transcription including transcription factors, phage proteins, and RNAP components (Dong et al., 2018). They found a transcription factor called SoxS for CRISPRa. Adding complexity, they found that they could also include a CRISPRi system for multiplexed modulation.

CRISPRa from sigma-54 promoters

Noting that previous work in bacterial CRISPRa had focused on sigma-70 dependent promoters, Baojun Wang's lab developed a system that can activate sigma-54 dependent promoters (Liu et al., 2019). While sigma-70 promoters are ubiquitous, they are not universal, particularly when it comes to non-model bacteria. The authors were able to engineer a dCas9 variant fused to bacterial enhancer binding protein required for the activation of sigma-54 type promoters. The system was functional in activating promoters in *Pseudomonas syringae* and *Klebsiella oxytoca*.

Microbiome engineering

Scientists have begun to realize the potential CRISPR technology can have on the human microbiome. You can consider the microbiome as a collection of organisms, their genes, and metabolic processes, occupying the gut. Many human diseases have connections to either specific organisms or organism/gene interactions with the host.

This scenario allows for several potential CRISPR applications. First, CRISPR can be used to target locations in the genomes of pathogenic or undesirable bacteria. This system can be delivered a

CRISPR Methods for Bacteria (CONT'D)

number of ways but success has been observed with Phagemids (Selle et al., 2020, Citorik et al., 2014, Bikard et al., 2014). When the CRISPR system targets these genomic locations, and introduces double strand breaks, the bacteria are unable to resolve them, and the cells die.

This could be applied to diseases where the presence of just a single organism, such as *Clostridium difficile*, could be targeted and preferentially depleted from gut. While still far from the clinic, CRISPR-Cas3 systems have been used successfully in mice for just such a purpose (Selle et al., 2020).

Targeting a subset of bacteria in a population

Work from the Lu lab found early success using CRISPR-Cas to target subsets of bacteria in population based on the presence of specific target sequences (Citorik et al., 2014). These could be genes, specific polymorphisms, including antibiotic resistance genes. This approach successfully targeted specific organisms when using bacteriophages or mobile plasmids in a bacterial strain introduced to the population as methods of introduction.

Engineering the bee gut microbiome

Beyond the human gut, the Barrick lab has deposited plasmids that can be used to engineer the bee gut microbiome. They created a modular system using a broad host range plasmid that can be maintained in a number of bacteria native to the honey bee and bumble bee gut microbiomes titling it the Bee microbiome toolkit (Leonard et al., 2018).

Base Editing

Among the newest approaches that leverage the remarkable properties of CRISPR-Cas systems is base editing. This process is generally accomplished by fusing an enzyme capable of converting one base to another to dCas9. The typical gRNA targeting approach is then used to direct the complex to a specific locus. This approach provides a more granular level of control where individual bases can be modified.

Cytidine deaminase PmCDA1

Akihiko Kondo's lab identified a cytidine deaminase, PmCDA1, that could be fused to dCas9 and efficiently introduce cytosine-to-thymine substitutions in target sequences (Banno et al., 2018). These substitutions were limited to approximately a five-base pair window of the target sequence which could be adjusted by changing the sgRNA length. The system could be further refined by the addition of a uracil DNA glycosylase inhibitor and degradation tag on the fusion protein to limit the activity of the system as a whole.

CRISPR Methods for Bacteria (CONT'D)

Cytidine deaminase for *Pseudomonas* species

Expanding this tool to other organisms, the Ji lab brought genome and base editing to *Pseudomonas* species (Chen et al., 2018). Using the cytidine deaminase, APOBEC1 and a nickase version of Cas9 they were effectively able to introduce point mutations in *P. aeruginosa*, *Pseudomonas putida*, *Pseudomonas fluorescens*, and *Pseudomonas syringae*. These vectors contain BsaI sites for easy cloning of a gRNA of interest.

Cytidine and adenine base editing in *Streptomyces*

Some of the most industrially and clinically relevant bacterial species reside in the genus *Streptomyces*, yet genome engineering remains challenging for most organisms in the genus. Work from the Sun lab demonstrates the utility of cytidine and adenine base editors that enable C-to-T and A-to-G base substitutions respectively. Substitutions seem most efficient when using a nickase version of Cas9, but were also successful using a dCas9 variant as well. Furthermore, they provide a high-fidelity (HF) variant of the BE3 that further improves editing rates.

References

Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, Fischetti VA, Marraffini LA (2014) Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nat Biotechnol 32:1146–1150 . https://doi.org/10.1038/nbt.3043

Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Research 41:7429–7437 . https://doi.org/10.1093/nar/gkt520

Citorik RJ, Mimee M, Lu TK (2014) Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nat Biotechnol 32:1141–1145 . https://doi.org/10.1038/nbt.3011

Cobb RE, Wang Y, Zhao H (2014) High-Efficiency Multiplex Genome Editing of *Streptomyces* Species Using an Engineered CRISPR/Cas System. ACS Synth Biol 4:723–728 . https://doi.org/10.1021/sb500351f

Dong C, Fontana J, Patel A, Carothers JM, Zalatan JG (2018) Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. Nat Commun 9: . https://doi.org/10.1038/s41467-018-04901-6

Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013) RNA-guided editing of bacterial genomes using

CRISPR Methods for Bacteria (CONT'D)

CRISPR-Cas systems. Nat Biotechnol 31:233–239 . https://doi.org/10.1038/nbt.2508

Leonard SP, Perutka J, Powell JE, Geng P, Richhart DD, Byrom M, Kar S, Davies BW, Ellington AD, Moran NA, Barrick JE (2018) Genetic Engineering of Bee Gut Microbiome Bacteria with a Toolkit for Modular Assembly of Broad-Host-Range Plasmids. ACS Synth Biol 7:1279–1290 . https://doi.org/10.1021/acssynbio.7b00399

Liu Y, Wan X, Wang B (2019) Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. Nat Commun 10: . https://doi.org/10.1038/s41467-019-11479-0

Peters JM, Koo B-M, Patino R, Heussler GE, Hearne CC, Qu J, Inclan YF, Hawkins JS, Lu CHS, Silvis MR, Harden MM, Osadnik H, Peters JE, Engel JN, Dutton RJ, Grossman AD, Gross CA, Rosenberg OS (2019) Enabling genetic analysis of diverse bacteria with Mobile-CRISPRi. Nat Microbiol 4:244–250 . https://doi.org/10.1038/s41564-018-0327-z

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell 152:1173–1183 . https://doi.org/10.1016/j.cell.2013.02.022

Ronda C, Pedersen LE, Sommer MOA, Nielsen AT (2016) CRMAGE: CRISPR Optimized MAGE Recombineering. Sci Rep 6: . https://doi.org/10.1038/srep19452

Selle K, Fletcher JR, Tuson H, Schmitt DS, McMillan L, Vridhambal GS, Rivera AJ, Montgomery SA, Fortier L-C, Barrangou R, Theriot CM, Ousterout DG (2020) *In Vivo* Targeting of Clostridioides difficile Using Phage-Delivered CRISPR-Cas3 Antimicrobials. mBio 11: . https://doi.org/10.1128/mbio.00019-20

Singh AK, Carette X, Potluri L-P, Sharp JD, Xu R, Prisic S, Husson RN (2016) Investigating essential gene function in*Mycobacterium tuberculosis*using an efficient CRISPR interference system. Nucleic Acids Res 44:e143–e143 . https://doi.org/10.1093/nar/gkw625

An "elegans" Approach to Better CRISPR/Cas9 Editing Efficiency

By Jordan Ward | Jan 27, 2015, Updated Dec 17, 2020

Emerging CRISPR/Cas9 editing technologies have transformed the palette of experiments possible in a wide range of organisms and cell lines. In *C. elegans*, one of the model organisms which I use to study gene regulation during developmental processes, CRISPR/Cas9 allows us to delete sequences and introduce mutations and epitopes with unprecedented ease. These new *C. elegans* approaches rapidly enrich for editing events without the need for any selective marker to remain in the edited animal.

Development of CRISPR/Cas9 editing strategies in *C. elegans*

As brilliantly and irreverently illustrated in the parable of the Geneticist vs the Biochemist, we geneticists are a lazy bunch who love to rely on the awesome power of genetic selection. Knock-in events tend to be rare, and the challenge with any experiment is in efficient recovery of these edits. Initial work in *C. elegans* relied on selective markers such as drug resistance (Chen et al., 2013), fluorescence (GFP) (Tzur et al., 2013), or rescue of mutant phenotypes (unc-119) (Dickinson et al., 2013). These approaches allowed effective recovery of knock-ins, but did result in 1-2 kilobases of additional sequence being introduced. Cre-mediated excision of the selective cassette minimizes the sequence added, leaving a 34 bp "scar", but currently requires additional experimental manipulation. In some cases, injection of Cre is required, however the development of the self-excising cassette (Dickinson et al., 2015) made this step as simple as heat-shocking to induce the expression of a Cre recombinase contained within the knock-in. Including Cre-mediated marker excision, it takes approximately 3-4 weeks to obtain a homozygous, outcrossed knock-in ready for experimentation.

In 2014, several manuscripts published in the journal Genetics detailed approaches in which selection for an editing event that produces a visible phenotype enriches for knock-outs and knock-ins at other genomic loci. The first approach – co-CRISPR from Craig Mello's lab – used inactivation of the unc-22 gene as their selection marker (Kim et al., 2014). Meanwhile Andy Fire's lab used oligo-mediated knock-in of a dominant mutation, known as co-conversion (Arribere et al., 2014). In both cases, the selected mutation must be removed, which can be done by isolating animals with particular visible phenotypes. These approaches may become even more powerful following a report that one can use linear repair templates (ie. PCR-derived dsDNA) with 30-60 bp homology arms to knock in large epitopes, such as GFP (Paix et al., 2014). A further development of this approach reported in 2017 uses the direct injection of a CRISPR/Cas9 ribonucleoprotein complex with linear DNA as repair templates which reduces time to generate edits to only four days (Paix et al., 2017). The co-CRISPR and co-conversion approaches have the advantage of being used in any genetic background, but require variable amounts of experimental manipulation and screening.

An "elegans" Approach to Better CRISPR/Cas9 Editing Efficiency (CONT'D)

A co-selection strategy for CRISPR editing in *C. elegans* improves efficiency

I was fortunate – or unfortunate enough, depending on perspective – to work with inefficient sgRNAs in my initial direct screening efforts, which were similar to the approach detailed by Paix et al. Although I was able to knock-in a 2xFLAG epitope into my gene of interest, I encountered low efficiency (0.13%) and laborious handling (screening 768 F1 animals). Recovering rare editing events in a sea of unedited animals struck me as a "needle in the haystack" type of problem and led to me to explore alternate approaches. Being a "lazy" geneticist, I developed a co-selection approach relying on repair of a conditional-lethal mutation to identify edits with minimal screening effort. Selecting for repair of a lethal mutation should remove much of the unedited "haystack", facilitating recovery of edited animals.

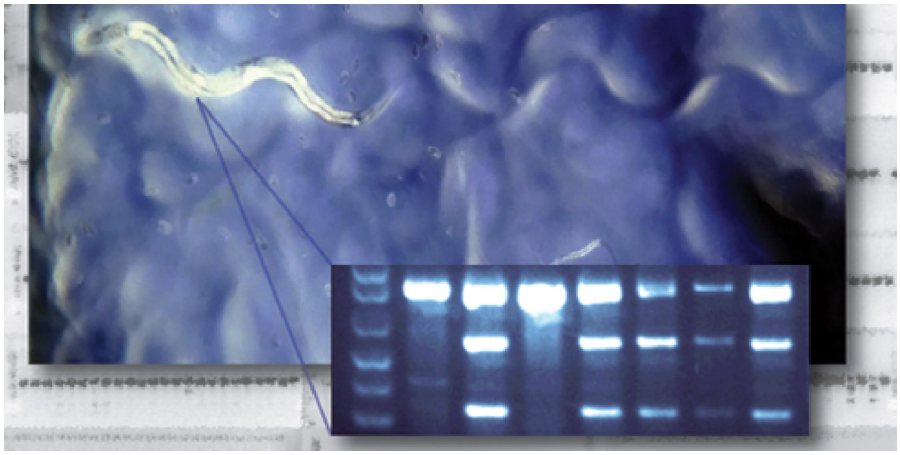


Figure 1: Only animals rescued for the conditional lethal mutation survive the selection. Image from Jordan Ward, UCSF.

In 2015 (Ward, 2015), I demonstrated that selection for repair of a temperature-sensitive pha-1 mutation significantly enriches for knock-in of 2x and 3xFLAG epitopes into other, non-linked loci; pha-1(e2123) mutant worms are perfectly viable at 15°C, yet exhibit complete embryonic lethality at 25°C. This method resulted in efficiencies ranging from 11-100% of F1 animals carrying precise knock-ins, and homozygous knock-in animals can be obtained in eight days. The only animals on a plate, other than the parental animal, are rescued progeny, which makes screening extremely rapid. This stringent selection allowed me to optimize a variety of editing parameters: oligo repair templates with homology arms of 35-80 bp, and DNA double-strand breaks up to 54 bp from the desired insertion site result in efficient editing. Repair oligos do not need to be PAGE purified, although doing so increases knock-in efficiency. Finally, as shown in Drosophila S2 cells (Böttcher et al., 2014), inactivation of NHEJ results in a further increase in knock-in efficiency, presumably by channeling DNA breaks into the homologous recombination repair pathway. Reagents required to perform pha-1 co-conversion are available through Addgene. In 2019, Farboud et al. described another similar co-conversion marker (repair of a zen-4(cle10ts) allele), increasing experimental options (Farboud et al., 2019).

An "elegans" Approach to Better CRISPR/Cas9 Editing Efficiency (CONT'D)

Expanding and improving the CRISPR/Cas editing toolbox for *C. elegans*

Recent work has further improved the efficiency of genome editing by CRISPR whether it be by improved gRNA or repair template design, delivery, screening, or other methods. Here are some highlights from the past few years:

Repair template design

The Meyer lab meticulously characterized several aspects of editing. First, they devised guidelines for designing single-stranded repair templates (Farboud et al., 2019).

The Mello lab also reported that modifying the 5' modification of oligos used to generate linear dsDNA repair templates by PCR boosted efficiency (Ghanta et al., 2018). They further built on this work, discovering that melting and rapidly cooling dsDNA repair templates dramatically boosted editing efficiency (Ghanta et al., 2020). They also found that the majority of knock-ins were occurring 24 hours post-injection, making screening for knock-ins much more efficient. Pure PCR repair templates are critical for efficient knock-ins (Ghanta et al., 2020). Together, these improvements from the Mello lab resulted in strikingly high editing efficiencies, obtaining as many as 100 independent GFP knock-ins from a single injected animal.

PAM sequence considerations

For changes 5' to the PAM, sequences corresponding to the same strand as the PAM produced a higher editing efficiency, while edits 3' to the PAM sequences corresponding to the opposite strand were better (Farboud et al., 2019).

The Meyer lab also determined optimal PAM orientation for editing using two Cas9 targets. An orientation where the 5' ends of the PAMs are facing outwards, ensuring the PAMs remain on the chromosome after DSB formation, was most efficient (Farboud et al., 2019). Using this approach, they were able to insert 10 kb of non-homologous sequence or a series of single-nucleotide substitutions along a 1.5 kb tract, all without selectable markers.

CRISPR in *C. elegans*

The Meyer and Mello labs identified several experimental set ups that could improve CRISPR efficiency in *C. elegans*. The Meyer lab also found that mating improved editing efficiency suggesting that sperm are more permissive to Cas9 editing (Farboud et al., 2019). Work from the Mello lab has dramatically improved RNP editing with linear repair templates. They found that Cas9 RNPs could be toxic at high concentrations, and editing efficiency improved by using more dilute mixtures (Doksin et al., 2018).

An "elegans" Approach to Better CRISPR/Cas9 Editing Efficiency (CONT'D)

The Mello lab reported efficient editing using a co-injection marker that is mainly lost in F2 animals, removing the need to segregate away co-conversion markers.

Expanding beyond Cas9

Since the development of CRISPR/Cas9 editing methods, other Cas proteins have been adapted for use in *C. elegans*. In 2017, the Korswagen lab developed Cas12a (Cpf1) for editing in *C. elegans* (Ebbing et al., 2017). Because *C. elegans* has an AT-rich genome, it can be difficult to find an appropriate NGG PAM site to target. The (T)TTN PAM motif required by the *Acidaminococcus* sp. BV3L6 Cas12a (AsCpf1) eases this constraint and allows editing from more sites of the *C. elegans* genome.

Toolkits for editing in *C. elegans*

The development of toolkits such as the SapTrap CRISPR/Cas Toolkit and SapTrap-SEC kit for *C. elegans* genome engineering have also facilitated CRISPR editing in *C. elegans*. These modular toolkits build single plasmid targeting vectors that encode the gRNA transcript and the repair template of your choosing. Tools such as SapTrap Builder greatly simplifies design of these constructs, allowing novice researchers to quickly and easily start genome editing (Schwartz et al., 2018).

To see more CRISPR tools for *C. elegans*, visit Addgene's *C. elegans* CRISPR plasmids page.

Future perspectives

It is interesting to note that recipients for the 2015 Breakthrough Prize included Jennifer Doudna and Emmanuelle Charpentier for their pioneering CRISPR/Cas9 work (now recognized by the 2020 Nobel Prize in Chemistry), and Victor Ambros and Gary Ruvkun for their seminal work on micro RNAs in *C. elegans*. The *C. elegans* work on small RNAs informed and drove work in countless other systems, whereas import of CRISPR technology into *C. elegans* has been a transformative innovation for our community. Developments in mammalian and yeast cells using modified Cas9 proteins to regulate gene expression (CRISPRi/CRISPRa), visualize specific genomic loci (CRISPR-imaging), or to find the proteins associated with a given genomic locus (CRISPR-ChAP-MS) could further transform the range of experiments imaginable in *C. elegans*. Equally, some of the "elegans" methods outlined in this post – particularly the use of co-selection methods – could greatly streamline editing in other organisms and systems, allowing rapid progress in a wide range of fields.

References

Arribere JA, Bell RT, Fu BXH, Artiles KL, Hartman PS, Fire AZ (2014) Efficient Marker-Free Recovery of

An "elegans" Approach to Better CRISPR/Cas9 Editing Efficiency (CONT'D)

Custom Genetic Modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. Genetics 198:837–846 . https://doi.org/10.1534/genetics.114.169730

Bottcher R, Hollmann M, Merk K, Nitschko V, Obermaier C, Philippou-Massier J, Wieland I, Gaul U, Forstemann K (2014) Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells. Nucleic Acids Research 42:e89–e89 . https://doi.org/10.1093/nar/gku289

Chen C, Fenk LA, de Bono M (2013) Efficient genome editing in *Caenorhabditis elegans* by CRISPR-targeted homologous recombination. Nucleic Acids Research 41:e193–e193 . https://doi.org/10.1093/nar/gkt805

Dickinson DJ, Ward JD, Reiner DJ, Goldstein B (2013) Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. Nat Methods 10:1028–1034 . https://doi.org/10.1038/nmeth.2641

Dickinson DJ, Pani AM, Heppert JK, Higgins CD, Goldstein B (2015) Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette. Genetics 200:1035–1049 . https://doi.org/10.1534/genetics.115.178335

Dokshin GA, Ghanta KS, Piscopo KM, Mello CC (2018) Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in *Caenorhabditis elegans*. Genetics 210:781–787 . https://doi.org/10.1534/genetics.118.301532

Ebbing A, Shang P, Geijsen N, Korswagen HC (2017) Extending the CRISPR toolbox for *C. elegans*: Cpf1 as an alternative gene editing system for AT-rich sequences. microPublication Biology. https://doi.org/10.17912/W2237D

Farboud B, Severson AF, Meyer BJ (2018) Strategies for Efficient Genome Editing Using CRISPR-Cas9. Genetics 211:431–457 . https://doi.org/10.1534/genetics.118.301775

Ghanta KS, Doksin GA, Mir A, Krishnamurthy PM, Gneid H, Edraki A, Watts JK, Sontheimer EJ, Mello CC (2018) 5′ Modifications Improve Potency and Efficacy of DNA Donors for Precision Genome Editing. BioRxiv. https://doi.org/10.1101/354480

Ghanta KS, Mello CC (2020) Melting dsDNA Donor Molecules Greatly Improves Precision Genome Editing in *Caenorhabditis elegans*. Genetics 216:643–650 . https://doi.org/10.1534/genetics.120.303564

Kim H, Ishidate T, Ghanta KS, Seth M, Conte D Jr, Shirayama M, Mello CC (2014) A Co-CRISPR Strategy

An "elegans" Approach to Better CRISPR/Cas9 Editing Efficiency (CONT'D)

for Efficient Genome Editing in *Caenorhabditis elegans*. Genetics 197:1069–1080 . https://doi.org/10.1534/genetics.114.166389

Paix A, Folkmann A, Seydoux G (2017) Precision genome editing using CRISPR-Cas9 and linear repair templates in *C. elegans*. Methods 121–122:86–93 . https://doi.org/10.1016/j.ymeth.2017.03.023

Paix A, Wang Y, Smith HE, Lee C-YS, Calidas D, Lu T, Smith J, Schmidt H, Krause MW, Seydoux G (2014) Scalable and Versatile Genome Editing Using Linear DNAs with Microhomology to Cas9 Sites in *Caenorhabditis elegans*. Genetics 198:1347–1356 . https://doi.org/10.1534/genetics.114.170423

Schwartz ML, Jorgensen EM (2018) SapTrap vectors for introducing point mutations with unc-119+ selection. microPublication Biology. https://doi.org/10.17912/DDVH-BG64

Tzur YB, Friedland AE, Nadarajan S, Church GM, Calarco JA, Colaiácovo MP (2013) Heritable Custom Genomic Modifications in *Caenorhabditis elegans* via a CRISPR–Cas9 System. Genetics 195:1181–1185 . https://doi.org/10.1534/genetics.113.156075

Ward JD (2014) Rapid and Precise Engineering of the *Caenorhabditis elegans* Genome with Lethal Mutation Co-Conversion and Inactivation of NHEJ Repair. Genetics 199:363–377 . https://doi.org/10.1534/genetics.114.172361

Engineering the Plant Genome Using CRISPR

By Joel McDade | Oct 11, 2016, Updated Dec 22, 2020 by Benoit Giquel

CRISPR has taken the genome engineering world by storm owing to its ease of use and utility in a wide variety of organisms. While much of current CRISPR research focuses on its potential applications for human medicine (Waltz, 2016), the potential of CRISPR for genome engineering in plants is also being realized. There are a variety of reasons to consider using genome editing to change the genetic code of plants, including the development of crops with longer shelf life and the development of disease-resistant crops to increase agricultural yield (Wang et al., 2016; Wang et al., 2014). While it is certainly possible to select for desirable traits using traditional plant breeding approaches, these techniques are cumbersome, often requiring several rounds of selection to isolate plants with the phenotype of interest. Genome engineering, on the other hand, allows for targeted modification of known or suspected genes that regulate a desired phenotype. In fact, CRISPR has already been used to engineer the genome of many plant species, including commonly used model organisms like Arabidopsis and Medicago truncatula and several crop species including potato, corn, tomato, wheat, mushroom, and rice (Khatodia et al., 2016). Despite the almost universal functionality of the CRISPR system in most organisms, some plant-specific changes to CRISPR components are necessary to enable genome editing in plant cells.

CRISPR components for plant genome engineering

CRISPR can be used to knockout, activate or repress target genes in plants using the same general experimental design principles developed in other model organisms (see our CRISPR guide for common CRISPR principles). However, plant-specific modifications to commonly used CRISPR plasmids are necessary to use the CRISPR system in plant cells. Like other model systems, expression of *S. pyogenes* Cas9 or Cas9 variants (hereafter referred to as Cas9) and a single stranded guide RNA (gRNA) is sufficient to modify the genome of plant cells. The structure of the gRNA (composed of a ~20 nucleotide targeting sequence and ~75 nucleotide scaffold sequence) is consistent between plants and other organisms, but the promoter used to drive gRNA expression is dependent upon the cell type in question. In plant cells, gRNA expression is achieved by placing the gRNA downstream of a plant-specific RNA pol III promoter, such as AtU6, TaU6, OsU6 or OsU3, which are commonly used to drive expression of small RNAs in their respective species. Addgene carries >30 "empty gRNA" backbones which contain a plant pol III promoter and gRNA scaffold sequence and allow researchers to insert targeting oligos with minimal cloning required. As with other model systems, multiple gRNAs can be expressed to modify several genomic loci at once (get more information on multiplexing gRNAs here).

Cas9 is commonly tagged with a nuclear localization sequence to enhance targeting to the nucleus, and several codon optimized Cas9 variants have been created in an effort to increase translation in a particular plant species or cell type (Belhaj et al., 2013). Nuclease dead Cas9 (dCas9) based activators (such as dCas9-VP64) or repressors (dCas9-KRAB or dCas9-SRDX) can also be used to activate or repress target genes in plant cells, respectively. Cas9 expression is commonly driven by plant-derived RNA pol II promoters which regulate expression of longer RNAs (such as mRNAs for gene expression).

Engineering the Plant Genome Using CRISPR (CONT'D)

Examples of commonly used RNA pol II promoters for Cas9 expression include the ubiquitously expressing cauliflower mosaic virus 35S promoter (CaMV 35S) or ubiquitin promoters (Belhaj et al., 2013). Addgene carries Cas9-containing plasmids for knockout, activation and repression of target genes in plants and many of the aforementioned empty gRNA backbones also contain Cas9, which enables expression of both Cas9 and the gRNA off of the same plasmid.

Delivering CRISPR components to plant cells

Once you have selected the correct CRISPR components for your application, it is time to deliver these components to your target cells. Remember, efficient delivery of CRISPR components is essential for any CRISPR experiment, and failure to express either the gRNA or Cas9 in your cell line will result in a failed experiment. CRISPR components can be expressed stably or transiently depending on the delivery method and cell type in question. CRISPR components can be delivered and expressed transiently using a standard detergent, Polyethylene Glycol (PEG), although the application of this approach is limited to protoplast cells (plant cells whose cell wall has been removed). Another common delivery method is agrobacterium-mediated delivery, which uses the soil derived bacterium Agrobacterium tumefaciens as a vehicle to deliver your gene of interest into a target cell line or organism (Presented in figure 1). More information on Agrobacterium-mediated transformation can be found in this blog post. The pDGE Dicot Genome Editing Kit from the Stuttmann lab contains a variety of Agrobacterium-compatible, Cas9 containing vectors ready for Golden Gate mediated cloning of your gRNA of interest.

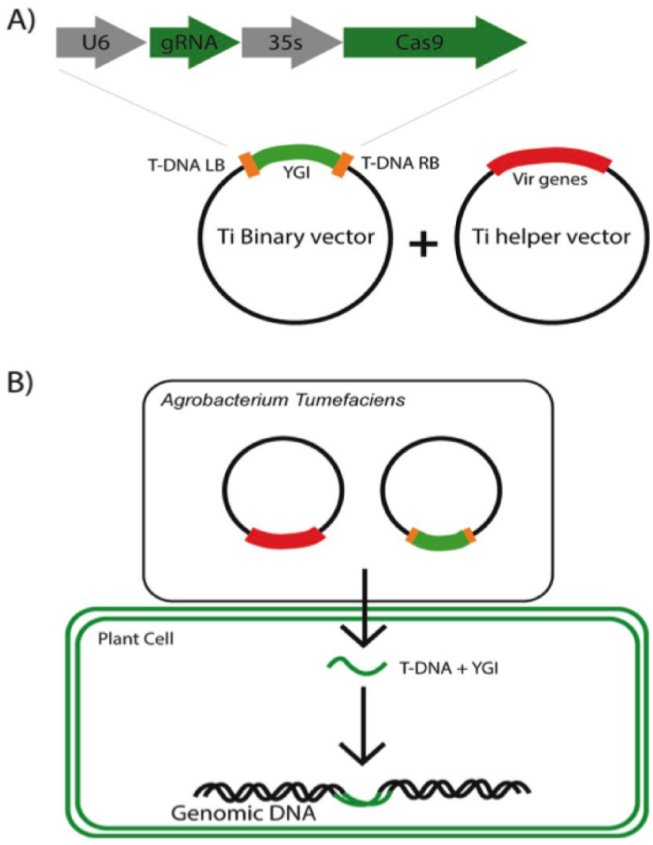


Figure 1) Agrobacterium tumefaciens can be used as a vehicle to deliver your gene of interest (YGI) into plant cells. The system typically consists of a Ti plasmid into which YGI is inserted and a Ti "helper" plasmid, which contains VIR genes necessary for T-DNA processing and insertion into the plant genome.

In recent years, deaminase-mediated base editing (cytosine base editor or adenine base editor) and reverse transcriptase-mediated prime editing technologies have been shown to be excellent alternative genome editing technologies especially in human cells. In contrast to homology directed repair, these methods do not involve double strand break (DSB) formation and do not require donor DNA. These

Engineering the Plant Genome Using CRISPR (CONT'D)

precise edits tend to be more efficient than HDR (reviewed in Zhu et al., 2019). Both cytosine base editors and adenine base editors have been developed for plants and prime editing for plants has been developed for rice by Yiping Qi's lab and for rice and wheat by Caixia Gao's lab.

More recently developed methods can also be used to efficiently deliver CRISPR-Cas9 components into plants. Nanoparticles such as carbon nanotubes (Kwak et al., 2019), DNA nanostructures (Zhang et al., 2019) or cell-penetrating peptides (Santana et al., 2020) and plants viruses (barley yellow striate mosaic virus (Gao et al., 2019) or sonchus yellow net rhabdovirus (Ma et al., 2020)) have already been shown to be efficient ways to deliver CRISPR-Cas9 to plant cells and should be considered alternative to PEG or agrobacterium-mediated delivery.

Summary

While different in many of the specifics - promoters used, precise protein sequences or domains, and methods of delivery - the underlying technique of CRISPR mediated genome engineering in plants isn't all that different from how it's used in other systems. Luckily, you don't have to look far for plasmids that have the plant-specific modifications required for targeting your favorite plant gene; you can find many plasmids for a wide variety of CRISPR applications in plants available through Addgene. In addition to the plasmids described above, Addgene carries several useful CRISPR toolkits for creating plant expression plasmids, including plant CRISPR plasmids from the Yiping Qi lab and the MoClo Plant Parts Kit from the Patron Lab. As with all plasmids in the repository, we highly recommend reading the associated publications or protocols to get the most out of the plasmid you chose for your experiment, but, if you're working with plants, don't be afraid to try your hand at CRISPR.

References

Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods 9:39 . https://doi.org/10.1186/1746-4811-9-39

Gao Q, Xu W, Yan T, Fang X, Cao Q, Zhang Z, Ding Z, Wang Y, Wang X (2019) Rescue of a plant cytorhabdovirus as versatile expression platforms for planthopper and cereal genomic studies. New Phytol 223:2120–2133 . https://doi.org/10.1111/nph.15889

Gelvin SB (2003) Agrobacterium-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool. MMBR 67:16–37 . https://doi.org/10.1128/mmbr.67.1.16-37.2003

Khatodia S, Bhatotia K, Passricha N, Khurana SMP, Tuteja N (2016) The CRISPR/Cas Genome-

Engineering the Plant Genome Using CRISPR (CONT'D)

Editing Tool: Application in Improvement of Crops. Front Plant Sci 7: 506. https://doi.org/10.3389/fpls.2016.00506

Kwak S-Y, Lew TTS, Sweeney CJ, Koman VB, Wong MH, Bohmert-Tatarev K, Snell KD, Seo JS, Chua N-H, Strano MS (2019) Chloroplast-selective gene delivery and expression in planta using chitosan-complexed single-walled carbon nanotube carriers. Nat Nanotechnol 14:447–455 . https://doi.org/10.1038/s41565-019-0375-4

Lowder LG, Zhang D, Baltes NJ, Paul JW III, Tang X, Zheng X, Voytas DF, Hsieh T-F, Zhang Y, Qi Y (2015) A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. Plant Physiol 169:971–985 . https://doi.org/10.1104/pp.15.00636

Ma X, Zhang X, Liu H, Li Z (2020) Highly efficient DNA-free plant genome editing using virally delivered CRISPR–Cas9. Nat Plants 6:773–779 . https://doi.org/10.1038/s41477-020-0704-5

Santana I, Wu H, Hu P, Giraldo JP (2020) Targeted delivery of nanomaterials with chemical cargoes in plants enabled by a biorecognition motif. Nat Commun 11: . https://doi.org/10.1038/s41467-020-15731-w

Wang F, Wang C, Liu P, Lei C, Hao W, Gao Y, Liu Y-G, Zhao K (2016) Enhanced Rice Blast Resistance by CRISPR/Cas9-Targeted Mutagenesis of the ERF Transcription Factor Gene OsERF922. PLoS ONE 11:e0154027 . https://doi.org/10.1371/journal.pone.0154027

Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, Qiu J-L (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat Biotechnol 32:947–951 . https://doi.org/10.1038/nbt.2969

Waltz E (2016) Gene-edited CRISPR mushroom escapes US regulation. Nature 532:293–293 . https://doi.org/10.1038/nature.2016.19754

Zhang H, Demirer GS, Zhang H, Ye T, Goh NS, Aditham AJ, Cunningham FJ, Fan C, Landry MP (2019) DNA nanostructures coordinate gene silencing in mature plants. Proc Natl Acad Sci USA 116:7543–7548 . https://doi.org/10.1073/pnas.1818290116

Zhu H, Li C, Gao C (2020) Applications of CRISPR–Cas in agriculture and plant biotechnology. Nat Rev Mol Cell Biol 21:661–677 . https://doi.org/10.1038/s41580-020-00288-9

Chapter 12

Biosafety

General CRISPR Safety Considerations

By Mary Gearing | Jan, 2016

Addgene loves supplying CRISPR plasmids, but we want to make sure you're aware of potential safety issues. These include biosafety concerns, as well as biocontainment concerns for researchers working with highly mobile model organisms like Drosophila. This section provides basic background information to help you use CRISPR safely, but please be sure to contact your institution's Biosafety Committee before beginning work.

Viral vectors

If you're using CRISPR with a viral vector system, please refer to our Viral Vector Biosafety Guide.

Lentiviral CRISPR systems are very common for both single gene and pooled library knockout experiments. The lentiviral systems available at Addgene are derived from HIV, but their organization across multiple plasmids and the deletion of many HIV proteins lowers the probability of generating replication-competent virus. Lentiviral vectors are handled at BSL-2/2+ safety levels.

Addgene also supplies some CRISPR plasmids for retroviral infection. Retroviral vectors are classified based on the cell types they infect; vectors that can infect human cells are handled at BSL-2/2+, while other vectors may be handled at BSL-1 depending on the target gene(s).

One popular CRISPR application uses SaCas9 with an adeno-associated viral vector (AAV) for *in vivo* gene editing. Since AAV is replication-incompetent and is not known to cause disease in humans, it is usually handled at BSL-1, provided that the gRNA(s) used do not have potential oncogenic, apoptotic or toxic effects. Please note, if you use a helper virus instead of a helper plasmid system to deliver your AAV cargo, your work should be done in BSL-2 conditions.

Preventing unintended editing

Since CRISPR is such a robust editing system, scientists need to be extra cautious when designing experiments to avoid the possibility of "accidental researcher self-editing." When working with model organisms, the easiest way to reduce this possibility is to design gRNAs that target sequences not conserved in humans. Once you've used your favorite CRISPR software to design gRNAs, BLAST them against the human genome to check for potential off-targets.

The method used to deliver Cas9 and gRNA(s) can also affect biosafety risk. Cell culture treatment or animal injection are contained, relatively low risk methods. Inhalation-based delivery presents a higher risk since it is more difficult to contain the viral particles. To minimize the risks associated with CRISPR inhalation, Addgene depositor Andrea Ventura has used an aerosol-based system with a replication-incompetent virus targeting mouse-specific loci.

General CRISPR Safety Considerations (CONT'D)

As alluded to above, the question of "what" you're trying to edit is often just as important as "how." Work targeting tumor suppressor or oncogenic genes, like P53 or KRAS, warrants a high level of prudence. Introducing human disease alleles into model organisms also comes with risk, especially if the gRNA target sequence is conserved in humans. In general, any editing that promotes oncogenesis or apoptosis, or could be potentially toxic, should be carefully designed to maximize biosafety and minimize researcher risk.

Gene drive containment

If you've kept up with CRISPR, chances are you've heard of gene drives. If not, check out our blog post on gene drives from guest writer Kevin Esvelt. In short, gene drives allow a genetic modification to spread rapidly through a sexually reproducing population. In standard inheritance, a heterozygous parent has a 50% chance of passing a modified gene to its offspring. In gene drive inheritance, nearly 100% of offspring inherit the modification. CRISPR/Cas9-based gene drives consist of Cas9 and a gRNA positioned adjacent to each other in the locus targeted by the gRNA. If this cassette is present on one chromosome, Cas9 activity can result in the gene drive being copied to the other chromosome, increasing its inheritance rate and allowing it to spread rapidly throughout the population.

Gene drives have many potential uses - in fact, a recently published gene drive renders mosquitoes resistant to the malaria-causing parasite Plasmodium falciparum. If released into the wild, this gene drive could end malaria as we know it. However, the key advantage of a gene drive is also its main disadvantage - if accidentally released, a gene drive can spread rapidly, with uncertain consequences. Thus, working with gene drives, especially in organisms likely to escape confinement, like flies and mosquitoes, requires experimental confinement strategies.

Avoiding accidental creation of a gene drive is simple: don't use a DNA vector that contains both Cas9 and a gRNA. This separation prevents Cas9 and a gRNA from integrating together into the genome, and is analogous to the common strategy of dividing viral vector components across multiple plasmids. For example, Cas9 can be maintained episomally or integrated into the AAVS1 safe harbor locus.

If you do want to work with gene drives, see the infographic below for some different types of confinement strategies. Ideally, multiple types of confinement strategies should be used together to prevent accidental release.

General CRISPR Safety Considerations (CONT'D)

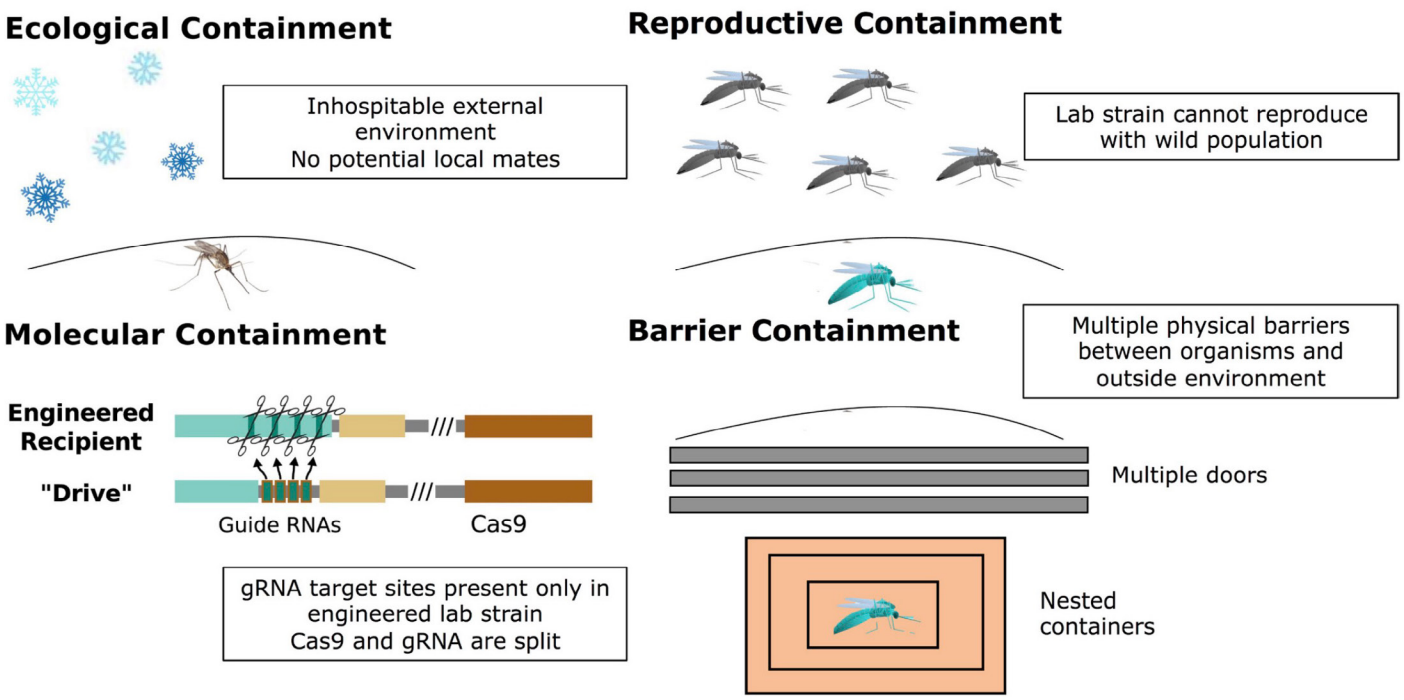


Figure 1: Visual depiction of common containment strategies. For optimal biocontainment, multiple strategies should be used together. Figure adapted from Esvelt et al.

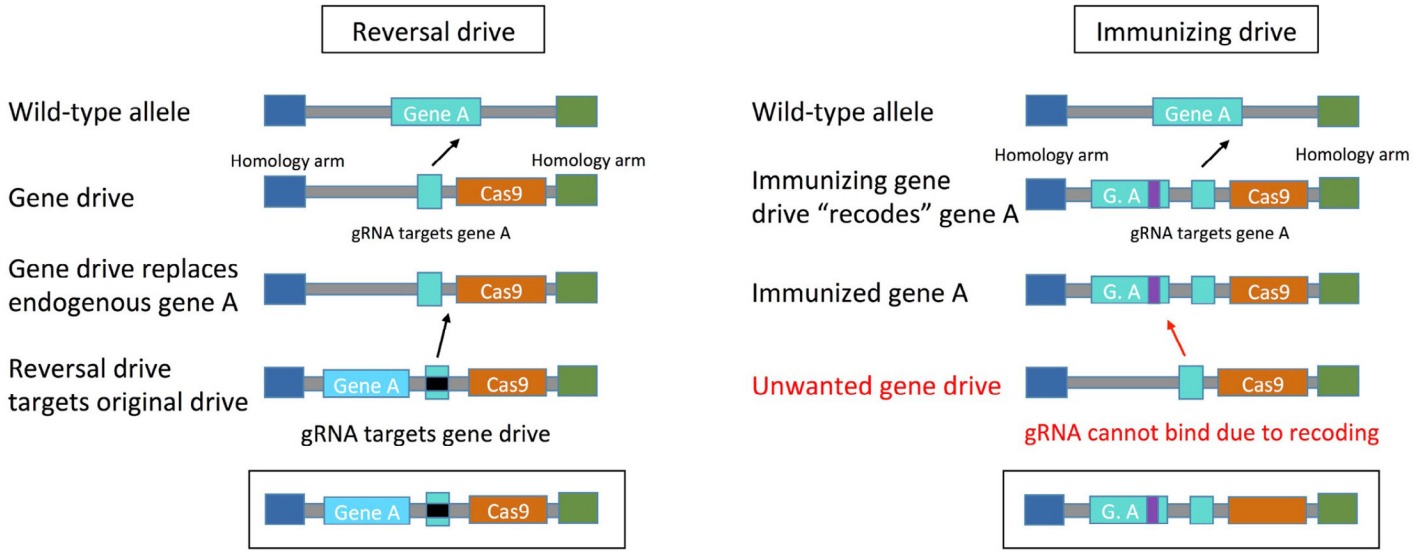
Gene drive reversal and immunization

Stringent confinement is only a part of the best practices for working with gene drives. Ideally, researchers should also develop tools to rapidly reverse engineered gene drives. The first reversal gene drive was recently developed in yeast. This reversal drive cuts the previously inserted gene drive to insert a functional version of the originally targeted gene (see figure below). Although Cas9 and a gRNA remain in the genome, wild-type gene function is restored.

Gene drive immunization can also prevent unwanted modifications due to accidental release. In this situation, the immunizing drive recodes a portion of the gene such that the gRNA in the gene drive cannot modify it, thus preventing a gene drive from affecting this "immunized" population (see figure 2).

Specific questions about how biosafety relates to your research should be directed to your institution's Biosafety Committee.

General CRISPR Safety Considerations (CONT'D)



- Modified/recoded version of previously deleted gene A is inserted to restore wild-type function
- Cas9 and gRNAs remain in genome

- Unwanted gene drive cannot target immunized gene A
- Cas9 and gRNAs remain in genome

Figure 2: Reversal and immunizing drives. Reversal drives "reverse" previous gene drives by reintroducing previously edited/deleted genes. Immunizing drives modify a gene of interest to prevent other gene drives from targeting it. Figure adapted from Esvelt et al.

References

Akbari OS, Bellen HJ, Bier E, Bullock SL, Burt A, Church GM, Cook KR, Duchek P, Edwards OR, Esvelt KM, Gantz VM, Golic KG, Gratz SJ, Harrison MM, Hayes KR, James AA, Kaufman TC, Knoblich J, Malik HS, Matthews KA, O'Connor-Giles KM, Parks AL, Perrimon N, Port F, Russell S, Ueda R, Wildonger J (2015) Safeguarding gene drive experiments in the laboratory. Sci Screen Shot 2020-12-23 at 8.40.18 AM ence 349:927–929 . https://doi.org/10.1126/science.aac7932

DiCarlo JE, Chavez A, Dietz SL, Esvelt KM, Church GM (2015) Safeguarding CRISPR-Cas9 gene drives in yeast. Nat Biotechnol 33:1250–1255 . https://doi.org/10.1038/nbt.3412

Esvelt KM, Smidler AL, Catteruccia F, Church GM (2014) Concerning RNA-guided gene drives for the alteration of wild populations. eLife 3: . https://doi.org/10.7554/elife.03401

Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, Bier E, James AA (2015) Highly

General CRISPR Safety Considerations (CONT'D)

efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. Proc Natl Acad Sci USA 112:E6736–E6743 . https://doi.org/10.1073/pnas.1521077112

Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han Y-C, Ogrodowski P, Crippa A, Rekhtman N, de Stanchina E, Lowe SW, Ventura A (2014) *In vivo* engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. Nature 516:423–427 . https://doi.org/10.1038/nature13902

Acknowledgements And Final Words

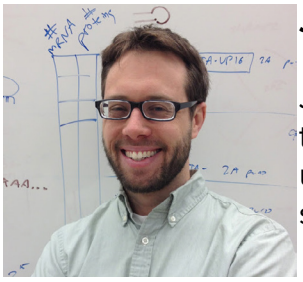
Special Thanks to Our Guest Contributors!

Allen Institute for Cell Science

The Allen Institute for Cell Science is a division within the Allen Institute, a nonprofit research organization founded by the late Paul G. Allen. The Allen Institute for Cell Science creates image-based, multi-scale models of cell organization, dynamics, and activities to predict and understand cells in normal and pathological contexts.

Alejandro (Alex) Chavez

Alejandro (Alex) Chavez is a member of the Addgene Scientific Advisory Board. His laboratory at Columbia University is focused advancing scientific discovery through the creation of innovative technologies. To date his group has develop several popular CRISPR-based tools for altering DNA sequence or regulating gene expression.

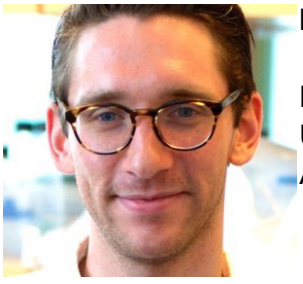


John Doench

John Doench is the Director of R&D in the Genetic Perturbation Platform at the Broad Institute and has worked with many Addgenies to help improve the understanding, curation, and explanation of our CRISPR resources. He really likes small RNAs.

Søren Hough

Søren Hough is a Biochemistry PhD candidate in the laboratory of Prof Steve Jackson at the Gurdon Institute at the University of Cambridge. Outside of the lab, he works as a freelance journalist and science writer with a focus on the ways science can be used to create a more just and equitable society. Read more at https://magazine.scienceforthepeople.org/.



Dylan Kwart

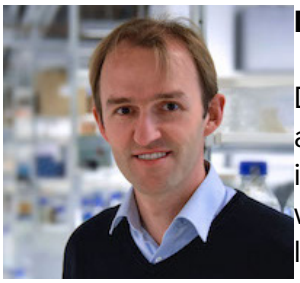
Dylan Kwart is a graduate student Marc Tessier-Lavigne's lab at the Rockefeller University. He currently works to understand pathological mechanisms of Alzheimer's disease by developing novel iPSC-based model systems.

Acknowledgements And Final Words

Special Thanks to Our Guest Contributors!

Alissa Lance-Byrne

Alissa Lance-Byrne is a graduate student at the University of California, Santa Cruz, where her research is focused on engineering nanoparticles as drug delivery systems. She is also a NASA Fellow studying biosignature preservation in Mars analogue environments.



Dominik Paquet

Dominik Paquet is currently a Professor for Neurobiology at the Institute for Stroke and Dementia Research, Ludwig-Maximilians-University Munich. His lab develops iPSC-based models of neurodegenerative and neurovascular diseases. He worked on the described project while he was a Postdoc in Marc Tessier-Lavigne's lab at the Rockefeller University in NYC.

Wenning Qin

Wenning Qin is a molecular biologist with Biogen and is particularly interested in applying CRISPR for mouse genome editing.

Dale Ramsden

Dale Ramsden is a member of the curriculum in Genetics and Molecular Biology, the Dept. of Biochemistry and Biophysics and the Lineberger Comprehensive Cancer Center at UNC Chapel Hill.

Marcelle Tuttle

Marcelle Tuttle is a resident in the Tufts Medical Center Internal Medicine Residency and former researcher at the Wyss Institute for Biologically Inspired Engineering.

Acknowledgements And Final Words

Special Thanks to Our Guest Contributors!

Haoyi Wang

Haoyi Wang is a Principle Investigator of the State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences. He has developed genome engineering technologies, including transposon-based method for determining transcription factors binding specificity, TALEN- and CRISPR-mediated genome editing in human pluripotent stem cells and mice.

Jordan Ward

Jordan Ward is an assistant professor at the University of California, Santa Cruz where his group focuses on gene regulation in nematode molting and spermatogenesis. His group also builds tools to improve genome editing and conditionally manipulate proteins.

David Wyatt

David Wyatt was formerly a graduate student in Dale Ramsden's lab interested in determining how the structure of broken DNA ends impacts how they are repaired.

Acknowledgements And Final Words

Thanks Addgenies, past and present!

Will Arnold

Will Arnold is a Senior Scientist II at Addgene. He earned his PhD from the University of Kentucky studying gene regulatory mechanisms in pathogenic bacteria. Will is interested in bioinformatics, his dogs, and all things Kentucky basketball.

Alyssa Cecchetelli

Alyssa Cecchetelli is a Quality Control Scientist at Addgene. She received her PhD from Northeastern University studying cell signaling and communication in *C. elegans*. She is interested in molecular biology, science communication, all NY sports teams, and traveling. Follow her on Twitter at @ACecchetelli

Gabrielle Clouse

Gabrielle Clouse is a Lab Technician for the Viral Vectors team at Addgene and has previously worked for Cardea Bio, the first group to integrate CRISPR/Cas systems with electronic sensors. She is interested in increasing accessibility in science, CRISPR technology, and synthetic biology.

Chari Cortez

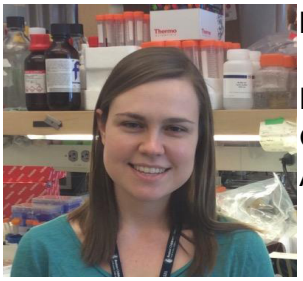
Chari is a former Scientist at Addgene. Her primary research experiences were in Immunology, infectious disease, and host cell receptor signaling in the context of initiating inflammation and host-pathogen interactions. She loves science, T cells, signaling pathways, and writing about science!

Melina Fan

Melina is co-founder and Chief Scientific Officer of Addgene. She received her PhD from Harvard University studying diabetes and obesity. She is dedicated to improving the infrastructure for research and discovery and is passionate about science, education, and entrepreneurship.

Acknowledgements And Final Words

Thanks Addgenies, past and present!



Mary Gearing

Mary Gearing is a former Scientist at Addgene. She got her start as a Science Communications Intern writing for the Addgene blog and website. As a full-time Addgenie, she enjoyed blogging about CRISPR and other cool plasmids!

Benoit Giquel

Benoit Giquel is the Outreach and Business Development Manager at Addgene. Based in the UK, he leads Addgene's outreach activities and helps furthering new international partnerships. Expert in Immunology and Cell Biology, he is also interested in genome editing and fluorescent-based techniques. Outside his duties at Addgene, he likes running, cooking and playing music.

Andrew Hempstead

Andrew Hempstead is a Senior Scientist at Addgene. Andrew's interests include genome engineering, synthetic biology, and microbiology.

Seth Kasowitz

Seth Kasowitz is a Senior Scientist at Addgene. Seth is interested in data analysis and visualization, RNA biogenesis, and meiosis. He loves helping scientists share their tools with the global research community.

Beth Kenkel

Beth Kenkel is currently a research scientist at a cell therapy company. She is particularly interested in science communication and viral vectors. Follow Beth on Twitter @ElizabethKenkel.

Acknowledgements And Final Words

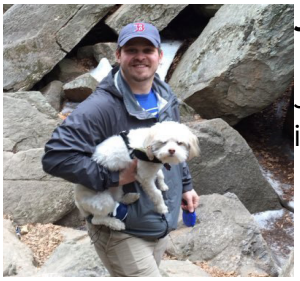
Thanks Addgenies, past and present!

Caroline LaManna

As Director of Marketing & Communications, Caroline leads Addgene's brand, marketing, communications, and educational content strategies. She oversees the customer and scientific support teams. She received her PhD in Biomedical Engineering from Boston University. Her goal is to make it easier for scientists to share reagents, communicate their research, and teach science in the community.

Michael Lemieux

Michael G. Lemieux is the Scientific Support Manager at Addgene with interests in professional development, graduate education reform, and molecular biology.



Joel McDade

Joel R. McDade is a former Business Development Manager at Addgene whose interests include Genome Editing and Science Education.

Kendall Morgan

Kendall is a science writer based in North Carolina. She has a PhD in Biology from the University of Oregon and a certificate in science writing from the University of California, Santa Cruz. She writes about science, medicine and science culture and is dedicated to making it easier for scientists to share what they make and do with each other and the world.

Christina Mork

Christina Mork is a Scientific Curator at Addgene and loves helping turn the gears of science. She received her PhD in microbiology from Boston University and has worked as a staff scientist at the Parkinson's Institute in California.

Acknowledgements And Final Words

Thanks Addgenies, past and present!

Nyla Naim

Nyla is Senior Scientist on the Scientific Support Team at Addgene. She received her PhD at the University of Pittsburgh and continued her postdoctoral research at the University of Vermont studying cellular signaling, biosensors, and optogenetics. Nyla supports biomedical research by connecting researchers with resources and promoting reproducible science.

Marcy Patrick

Marcy is the Director of Production at Addgene and oversees the plasmid intake, storage, and QC processes. She loves that her job allows her to support the scientific community day in and day out, while still keeping up to date on all the cool science out there.

Leah Schwiesow

Leah Schwiesow is a Scientific Curator at Addgene. She received her PhD from the University of California, Santa Cruz and studied virulence gene regulation in pathogenic bacteria.

Jennifer Tsang

Jennifer Tsang is the Science Communications and Marketing Coordinator at Addgene and manages the blog and social media. She is particularly interested in science communication and microbiology. Follow her on Twitter @jw_tsang.

Cary Valley

Cary Valley is a Senior QC Scientist at Addgene and is interested in data analysis.

Acknowledgements And Final Words

Thanks Addgenies, past and present!

Nicole Waxmonsky

Nicole is an UI/UX Lead & Strategist at Addgene. She received her PhD in Cell and Molecular Biology from the University of Minnesota. She enjoys using the web to increase scientific sharing and to improve access to basic science information.

Aliyah Weinstein

Aliyah Weinstein is Addgene's Marketing and Communications Manager. She has a research background in cancer immunology and is engaged in science writing and advocating for diversity and equity in STEM. Outside of science, Aliyah enjoys yoga, board games, and traveling. You can find Aliyah on Twitter @desabsurdites.

If you have any questions, comments or suggestions about how Addgene can improve its educational content, please contact us at blog@addgene.org.

Join our community!

@addgene

blog.addgene.org

www.addgene.org