CRISPR 101:
A Desktop Resource

Created and Compiled by Addgene
May 2017 (2nd Edition)
www.addgene.org

addgene
A better way to share plasmids

CRISPR 101: INTRODUCTION TO ADDGENE'S DESKTOP RESOURCE

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 in this handy new resource.

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


www.addgene.org


blog.addgene.org


www.facebook.com/addgene


www.twitter.com/addgene


www.linkedin.com/company/addgene


https://www.youtube.com/user/addgenemedia


http://blog.addgene.org/topic/podcast

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT'D)

TABLE OF CONTENTS (CONT'D)

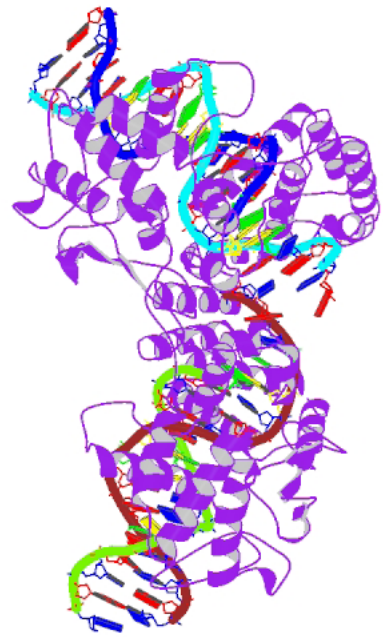
GENOME ENGINEERING OVERVIEW

INTRODUCTION TO GENOME ENGINEERING

By Mary Gearing | August 26, 2015

Genome engineering has made it possible to not only dissect complex gene interactions, but also to build new pathways through synthetic biology. The past few decades have seen tremendous advances in both the number and feasibility of genome engineering techniques, many of which are available from Addgene.

Genome engineering was born in the late 1970s, when multiple groups showed that exogenous DNA could be taken up by yeast or bacteria and randomly integrated into the genome. Subsequent work showed that this process could also occur in a targeted fashion. Addgene depositor Mario Capecchi realized that DNA microinjection into a cell's nucleus would stimulate cellular homologous recombination, permitting targeted genome modification. In 1989, he, Martin Evans and Oliver Smithies created the first knockout mouse, a watershed moment for genome engineering.



Cre Recombinase Complexed to DNA (1).

Cre-lox

The late 1980s also marked the introduction of Cre-lox recombination, a system derived from P1 bacteriophage now widely used to control gene expression. Today, Cre recombinase under the control of various promoters, or in its inducible form, provides sophisticated spatiotemporal control of gene expression, especially in mouse transgenics.

The Cre-lox system is a technology that can be used to induce site-specific recombination events. The system consists of two components derived from the P1 bacteriophage: the Cre recombinase and a loxP recognition site. The P1 bacteriophage uses these components as part of its natural viral lifecycle, and researchers have adapted the components for use in genome manipulation.

Cre recombinase, originally named because it "causes recombination" (although later referred to as the "cyclization recombinase"), is a 38 kDa protein responsible for intra- and inter-molecular recombination at loxP recognition sites. A key advantage of the system is that Cre acts independently of any other accessory proteins or co-factors, thus allowing for broad applications in a variety of experiments.

LoxP (locus of X(cross)-over in P1) sites are 34-base-pair long recognition sequences consisting of two 13-bp long palindromic repeats separated by an 8-bp long asymmetric core spacer sequence. The asymmetry in the core sequence gives the loxP site directionality, and the canonical loxP sequence is ATAACTTCGTATA-GCATA-CATTATACGAAGTTAT. The loxP sequence does not occur naturally in any known genome other than P1 phage, and is long enough that there is virtually no chance of it occurring randomly. Therefore, inserting loxP sites at deliberate locations in a DNA sequence allows for very specific manipulations as shown in the figure below.

INTRODUCTION TO GENOME ENGINEERING (CONT'D)

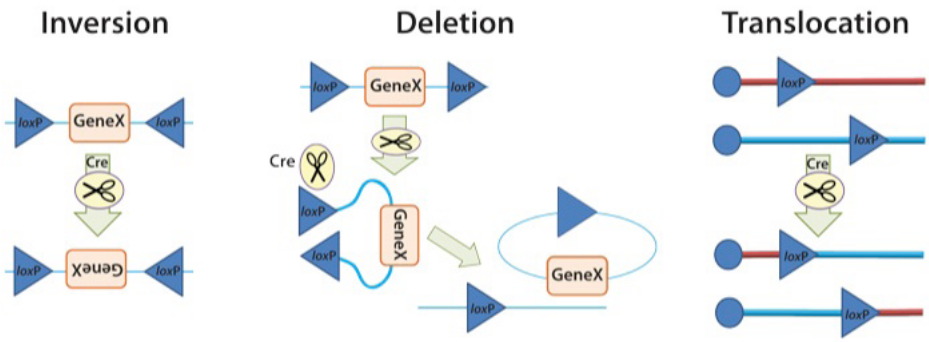


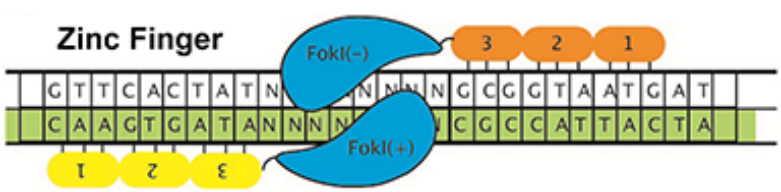
Image by Larissa Haliw

DNA Manipulations Possible with Cre-lox
Inversion: *If the loxP sites are on the same DNA strand and are in opposite orientations, recombination results in an inversion and the region of DNA between the loxP sites is reversed.* **Deletion:** *If the sites face in the same direction, the sequence between the loxP sites is excised as a circular piece of DNA (and is not maintained).* **Translocation:** *If the sites are on separate DNA molecules, a translocation event is generated at the loxP sites.*

Homologous Recombination and the Journey Towards CRISPR/Cas9

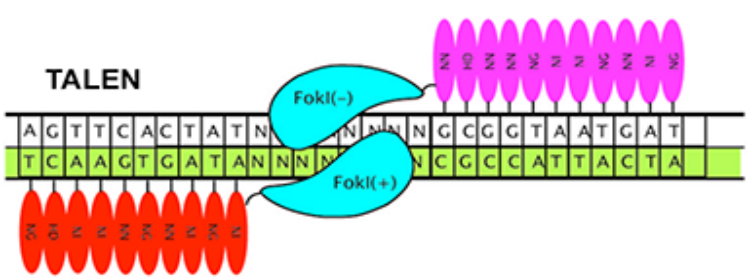
Homologous recombination is a cornerstone of genome engineering, but with the caveat that it occurs at low frequencies, limiting editing efficiency. To improve editing rates, researchers hijacked the function of endonucleases, which create difficult-to-repair DNA double-stranded breaks. Targeting these enzymes to a given locus results in DNA cleavage and forces the cell to undergo either non-homologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ occurs if no DNA repair template is provided, and its error-prone nature often leads to inactivating mutations. In the case of HDR, a repair template specifies desired genomic modifications, enabling precise editing. In addition to introducing point mutations or recombination sites, repair templates can also be used to introduce a gene of interest into a given locus.

Zinc Finger Nucleases



Zinc finger nucleases (ZFNs) represent the first step towards efficient, targeted nucleases. To create a ZFN, a series of zinc fingers is designed to bind to a specific genomic locus, and subsequently fused to FokI nuclease. Paired ZFNs recognizing two adjacent sites cleave DNA, initiating HDR. The utility of ZFNs is limited by their long synthesis time and nonmodular assembly process. Although computational tools helped improve targeting, it's not possible to design suitable ZFN pairs for every genomic locus.

TALENs

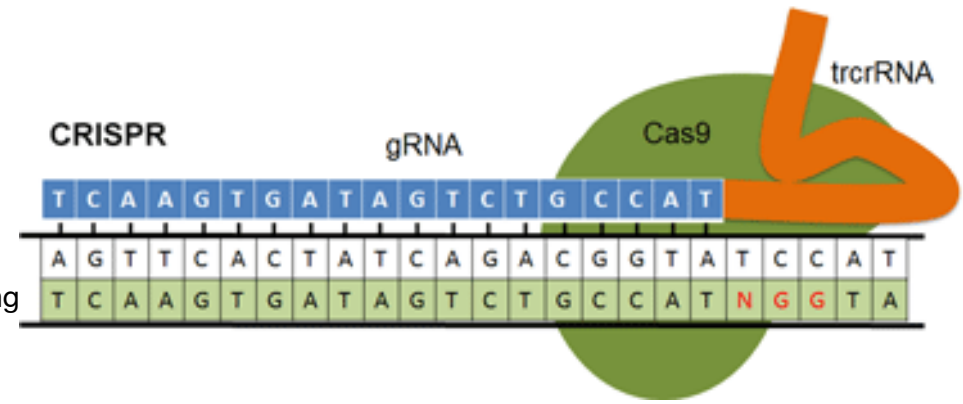


First reported in 2011, TALENs represented a huge step forward for genome engineering. This modular system is based on TAL effector DNA binding proteins, isolated from *Xanthomonas spp*, fused to FokI endouclease. Whereas it might take an experienced scientist six weeks to make a ZFN, a newbie could make a TALEN in just a few days! TALEN technology was rapidly adopted by the research community, with the Golden Gate TALEN kit becoming Addgene's most popular kit. The customizable DNA-binding properties of TALENs also enabled the design of custom transcription factors to modulate gene expression.

INTRODUCTION TO GENOME ENGINEERING (CONT'D)

CRISPR

Just when we thought genome engineering couldn't get any better, along came the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas9 system. CRISPR is an important component of the bacterial immune system that allows bacteria to remember and destroy phages. In genome engineering applications, Cas9 endonuclease is targeted by guide RNA (gRNA) sequence homology to a given locus,



where it induces a double stranded break. Like ZFNs and TALENs, CRISPR/Cas9 employs HDR, but the use of RNA to specify editing makes the system much less expensive and time-consuming, as well as more precise and scalable. For this reason, CRISPR/Cas9 has proven to be incredibly valuable for high-throughput genome engineering. CRISPR/Cas9 can also target multiple loci in one organism, and like TALENs, the system has also been adapted for other functions. CRISPRs are even more accessible to the research community than TALENs, with new papers using this technology published every week.

Other Important Genome Engineering Tools

With the success of CRISPR, it's easy to forget about other genome engineering methods...but you shouldn't! Another HDR-based method, recombineering, is commonly used in *E. coli* to make edits to the genome or a bacterial artificial chromosome (BAC). 50 bp homology arms flanking the repair template specify the site of recombination, catalyzed by phage recombinases. Since repair templates can be quickly generated using PCR, recombineering is easily scalable just like CRISPR. In addition to its applications in bacterial genome engineering, recombineering is also useful for creating BAC-based repair templates for other HDR genome engineering methods.

Outside of HDR, an exciting new gene transfer method uses the Sleeping Beauty transposon, reconstructed from fossils of ancient fish. Transposons are mobile DNA elements, and are thus ideal for gene transfer. This system is divided across two plasmids, one containing the gene of interest (GOI) flanked by inverted repeats, the other carrying the transposase. Following cotransfection, the transposase cleaves the GOI from the plasmid and facilitates a double strand break to allow the GOI to integrate into the genome. Sleeping Beauty was named 2009's Molecule of the Year, and it represents a robust alternative to viral vector-mediated gene transfer.

Conclusion

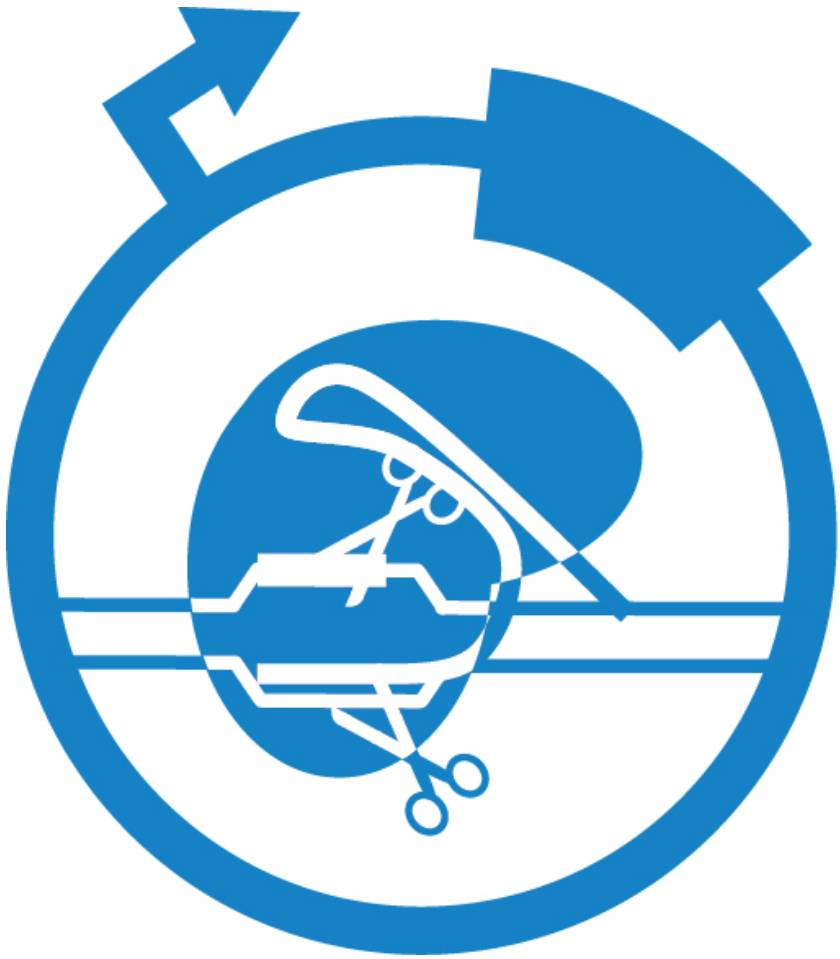
With multiple robust and efficient genome engineering methods at our fingertips, we have entered a Golden Age of genome engineering. Current work focuses on refining these techniques to ensure high specificity and activity, whatever the desired target locus (or loci) may be, with the hope that these methods will be useful clinically. What we at Addgene find most exciting is the democratization of genome engineering, which has and will continue to allow researchers all over the world to use these tools in their research. Continue reading to catch up on all things CRISPR.

INTRODUCTION TO GENOME ENGINEERING (CONT'D)

Further Reading

1. Guo, Feng, Deshmukh N. Gopaul, and Gregory D. Van Duyne. "Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse." Nature 389.6646 (1997): 40-46. PubMed PMID: 9288963.
2. Addgene's Genome Engineering Guide

WHAT IS CRISPR?



HISTORY OF CRISPR

By Mary Gearing | January, 2015

The CRISPR revolution has began, and it shows no signs of slowing down. This system, which is key to prokaryotic adaptive immunity, has proven to be especially amenable to genome engineering. CRISPR/Cas offers flexibility, as well as easy multiplexing and scaling, far beyond the capacities of previous systems. Prokaryotes have long utilized CRISPR/Cas as a powerful defensive strategy against viral invaders, and this system is proving to be just as useful (if not more so) for research applications.

CRISPR/Cas: An Adaptive Immune System

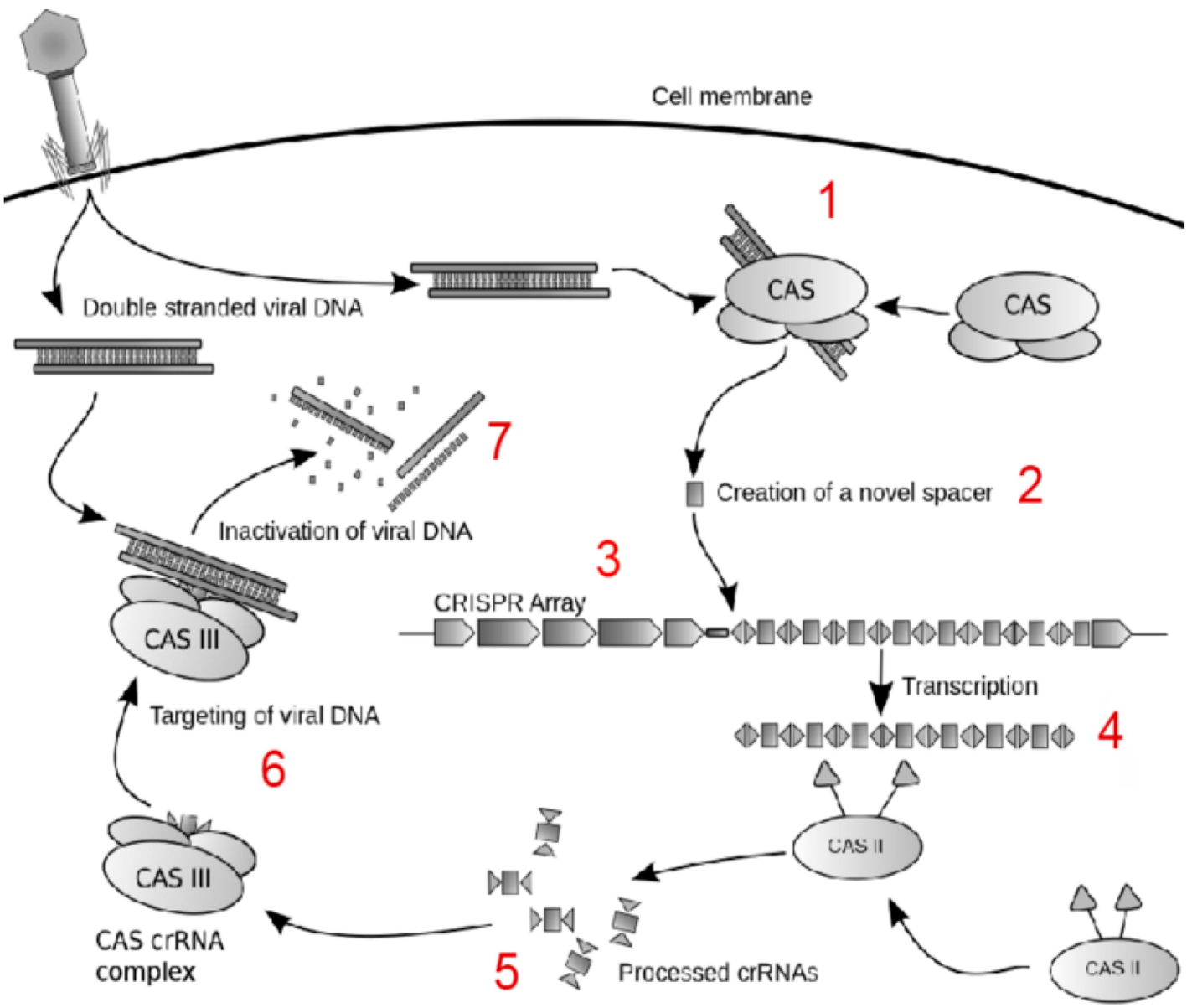
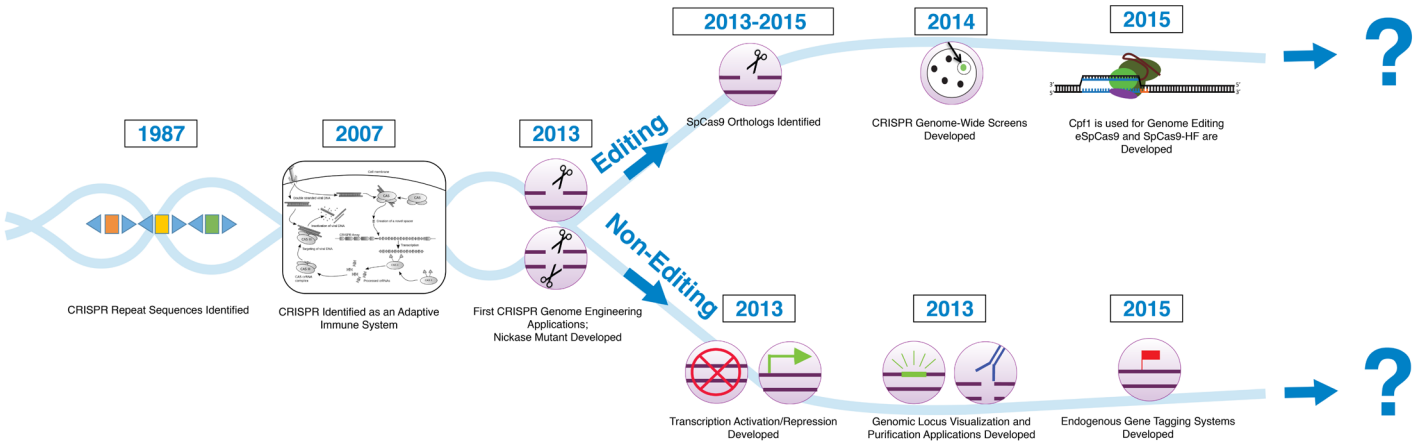


Figure 1. An overview of CRISPR/Cas as a bacterial adaptive immune system. When foreign viral or plasmid DNA enters the cell, a Cas complex recognizes it and cleaves it into small fragments (1), adding a new spacer (2) to the end of the CRISPR array. This array contains small pieces of DNA from past invaders (3), but does not contain PAM sites, so it is recognized as "self." The CRISPR array is transcribed into a long RNA (4) that is subsequently cleaved into mature crRNAs (5). These crRNAs direct the Cas complex to the foreign DNA based on sequence specificity (6), allowing the DNA to be cleaved and destroyed (7). Wikipedia, accessed 25 November 2013. Author: James Atmos (3).

HISTORY OF CRISPR (CONT'D)

CRISPR Timeline 1987-2015



CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat) sequences were initially discovered in the *E. coli* genome in 1987 (1), but their function as a safeguard against bacteriophages was not elucidated until 2007 (2). Initial experiments exposed *S. thermophilus*, a bacterium important for yogurt and cheese production, to predatory phages to test if exogenous phage DNA could be incorporated into the bacterial genome as part of the CRISPR repeats. Cas (CRISPR associated) genes, which code for polymerases, nucleases, and helicases, were also disrupted to determine their various roles in this process. Scientists hypothesized that prokaryotes had developed an adaptive immune system - utilizing various Cas genes to not only store a record of invading phages but also to destroy the phage upon re-exposure (2,3) (Figure 1). More specifically, specialized Cas proteins snip foreign DNA into small fragments approximately 30 bp in length and paste them into the CRISPR sequence. Separate Cas proteins then express and process the CRISPR loci to generate CRISPR RNAs (crRNAs). Through sequence homology, these crRNAs guide a Cas nuclease to the specified exogenous genetic material, located next to the species-specific protospacer adjacent motif (PAM). The CRISPR/Cas complex binds to the foreign DNA and cleaves it to destroy the invader.

CRISPR systems, found in 95% of archaeal and 48% of bacterial genomes, are highly diverse, with variation in PAM sequences and the number and type of Cas proteins (4). Makarova et al.'s classification defines 5 types and 16 subtypes based on shared characteristics and evolutionary similarity. These are grouped into two large classes based on the structure of the effector complex that cleaves genomic DNA (5). The Type II CRISPR/Cas system was the first harnessed for genome engineering, with Type V following in 2015.

Fighting Back: Evolution of Anti-CRISPR Genes in Phage

The CRISPR/Cas adaptive immune system seems like a winning strategy to ensure prokaryotic survival, but it is not invincible. In the original *S. thermophilus* experiments, virulent phage 858 was able to generate single point mutations in the S1 spacer region, preventing the Cas nucleases from re-identifying it, and thus circumventing the CRISPR/Cas defense mechanism (3). Surprisingly, only a few genes that neutralize CRISPR/Cas have been identified. A 2013 study found five distinct "anti-CRISPR" genes in bacteriophages infecting *Pseudomonas aeruginosa*. Addition of one of the five anti-CRISPR genes allowed a phage to evade *P. aeruginosa*'s Type I-F CRISPR/Cas system, while mutating a phage's anti-CRISPR gene had the opposite effect (6). These genes were only able to deactivate the Type I-F system and are not translatable to other CRISPR subtypes. Subsequent work found anti-CRISPRs inhibiting Type I-E CRISPR/Cas in *P. aeruginosa*, but these genes do not

HISTORY OF CRISPR (CONT'D)

inhibit type I-E activity in other bacteria such as *E. coli* (7).

Later work has shown that these anti-CRISPR genes employ varied mechanisms to block CRISPR/Cas. Of the three proteins studied, two prevented the CRISPR-Cas complex from binding to DNA, but did so by binding different complex subunits. Another anti-CRISPR binds to the helicase-nuclease Cas3, thus preventing its recruitment to the CRISPR-Cas complex (8).

The high diversity and near ubiquity of CRISPR systems suggests that there may be a variety of undiscovered anti-CRISPR genes, but these genes will likely provide limited protection, as seen with the previously characterized anti-CRISPRs. It will be important to monitor this molecular arms race to see what new strategies both bacteria and phage develop to oppose one another. Like the current CRISPR systems, these battles may provide previously unimagined tools for genome engineering.

Genome Engineering Applications

In 2012, Jinek et al. first demonstrated that CRISPR/Cas could be programmed for targeted DNA cleavage *in vitro* (9). In 2013, Cong et al. and Mali et al. described CRISPR/Cas-based genome editing in cell culture (10, 11). Less than four years later, PubMed lists more than 2600 CRISPR-related publications, many of which detail work to improve the tool's specificity, orthogonality, and multiplexibility in various species, as well as the development of new applications. The first CRISPR papers described two main categories of genome edits. First, the double-stranded break induced by Cas9 could be repaired through non-homologous end joining (NHEJ), an error-prone process that produces insertions/deletions disrupting a target gene. On the other hand, if a repair template with homology arms was supplied, the break could be repaired according to this template, allowing for precise gene editing. These edits can include mutations seen in human patients, protein tags, or loxP/FRT sites. Homology-directed repair (HDR) allows users to create a variety of modifications in endogenous loci with unprecedented speed and specificity.

CRISPR/Cas technology development has occurred at a truly astonishing pace, with much work directed towards increasing specificity. A "nickase" mutant, which cuts only one strand of DNA, is now commonly used with paired gRNAs to lower off-target cleavage frequency (10). gRNA length has also been optimized; truncated gRNAs with <20 base homology display less off-target activity (12). Two exciting new Cas9 variants display low off-target activity due to rationally designed mutations. Mutations in enhanced SpCas9 (eSpCas9) reduce the enzyme's helicase activity, lowering cleavage at off-target sites without compromising on-target editing efficiency (13). Similarly, SpCas9-HF contains engineered mutations in the SpCas9 residues that normally form hydrogen bonds with DNA. Like eSpCas9, SpCas9-HF lowers off-target effects to nearly undetectable levels (14).

Multiple Cas9 orthologs from other bacterial species, such as *Staphylococcus aureus*, have also been characterized (15). These orthologs have distinct PAM sites that do not cross react, allowing multiple Cas9s to be used together. SpCas9 mutants with alternative PAM sites have also been identified. One mutant, D1135E, retains the same 3'-NGG PAM, but is more selective for this PAM sequence than wt SpCas9, which may reduce the frequency of off-target effects (16). Cpf1, another CRISPR nuclease, recognizes a 5'-TTN PAM, improving targeting in AT-rich genomes (17). This expansion of possible PAM sequences moves us closer to being able to target every locus in every genome. PAM diversity will also improve HDR applications, as editing efficiency increases when the PAM is in close proximity to the desired edit site.

CRISPR/Cas has truly revolutionized genetics in a variety of model organisms. *C. elegans* was one of the

HISTORY OF CRISPR (CONT'D)

first organisms to which CRISPR was applied - today, you can generate a protein fusions or loss of function mutations in *C. elegans* with just a few hours of hands-on time (18)! Zebrafish and Drosophila are also easily modified using CRISPR (19,20). For those interested in mammalian systems, CRISPR has made it easier than ever to generate knockout or mutant animals; injecting RNA encoding Cas9 and the appropriate gRNA into mouse embryos results in efficient germline editing (21). Combining SaCas9 with adeno-associated viral vectors has also made postnatal genome editing a reality in mice (15).

Genome editing with CRISPR/Cas is not only simple, but also scalable. Multiple gRNAs targeting multiple loci can be easily expressed in the same cell or organism (10, 11). The use of RNA to target a locus makes CRISPR especially amenable to genetic screens. Pooled gRNA libraries can be used to identify genes that are important to a given phenotype (22, 23). Current libraries are available for gene knockouts, as well as transcriptional activation or repression (24). Combined with the power of next-generation sequencing, CRISPR/Cas9 is the most robust system yet for genome-wide screening.

Using CRISPR for Non-editing Applications

Scientists have also used the targeting capability of CRISPR Cas9 to perturb and study biology without altering DNA sequence. For instance, a catalytically dead Cas9 mutant (dCas9) has been especially useful, with fusion proteins available to help researchers activate or deactivate transcription of given genes (25). Fluorescently labeled dCas9 helps visualize genomic loci of interest prompting new work into nuclear organization (26). Epitope-tagged dCas9 can also be used to purify a genomic locus and its associated proteins or RNAs, a process called enCHIP (27). Multiple methods have been developed to tag endogenous proteins, enabling specific protein isolation without the need for a custom antibody (28, 29).

Future Directions

It's amazing to see how much progress has been made in just a few years, with the skyrocketing popularity of CRISPR echoing that of RNAi and iPS cells and quickly surpassing that of its predecessors ZFNs and TALEs. Genome engineers continue to work to develop a highly specific, programmable platform well-suited for various biological and translational technologies. Biotechnology companies are exploring therapeutic applications of CRISPR to treat genetic disease, with the caveat that off-target editing risk must be made very low or nonexistent. Three studies published concurrently in *Science* demonstrate *in vivo* treatment of mice suffering from Duchenne muscular dystrophy; these studies represent the first evidence of successful postnatal genome editing in a disease model (30, 31, 32). Scientists continue to wrestle with the possibility of postnatal and germline editing, with the International Summit of Human Gene Editing convened in 2015 to discuss the ethics associated with CRISPR editing in agricultural and therapeutic applications.

Despite the ethical controversies surrounding non-research applications, it's clear that CRISPR is a truly disruptive technology for basic research. The beauty of CRISPR/Cas is that it's adaptable to almost any model system or biological topic, and you don't need to be an expert to see results! Addgene has empowered researchers to harness previous experimental successes and further develop the CRISPR/Cas toolkit by posting lab protocols, providing tips from experts in the field, and enabling access to multiple plasmids used for various applications. We hope this eBook will further facilitate the rapid use and development of CRISPR tools!

HISTORY OF CRISPR (CONT'D)

Further Reading

1. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. J Bacteriol. 1987 Dec;169(12):5429-33. PubMed PMID: 3316184.
2. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007 Mar 23;315(5819):1709-12. PubMed PMID: 17379808.
3. Atmos J. Diagram of the possible mechanism for CRISPR. Photo: Wikipedia. 2009 Sept 15. http://en.wikipedia.org/wiki/File:Crispr.png
4. IGM, Universite Paris-Sud. CRISPRdb. Last updated 2014 Aug 05. http://crispr.u-psud.fr/crispr/CRISPRdatabase.php
5. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJ, Charpentier E, Haft DH, Horvath P, Moineau S, Mojica FJ, Terns RM, Terns MP, White MF, Yakunin AF, Garrett RA, van der Oost J, Backofen R, Koonin EV. An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol. 2015 Nov;13(11):722-36. PubMed PMID: 26411297.
6. Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature. 2013 Jan 17;493(7432):429-32. PubMed PMID: 23242138.
7. Pawluk A, Bondy-Denomy J, Cheung VH, Maxwell KL, Davidson AR. A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. MBio. 2014 Apr 15;5(2):e00896. PubMed PMID: 24736222.
8. Bondy-Denomy J, Garcia B, Strum S, Du M, Rollins MF, Hidalgo-Reyes Y, Wiedenheft B, Maxwell KL, Davidson AR. Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. Nature. 2015 Oct 1;526(7571):136-9. PubMed PMID: 26416740.
9. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012 Aug 17;337(6096):816-21. PubMed PMID: 22745249.
10. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013 Jan; 339 (6121): 819–23. PubMed PMID: 23287718. PubMed Central PMCID: PMC3795411.
11. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. Science. 2013 Jan;339 (6121): 823–6. PubMed PMID: 23287722. PubMed Central PMCID: PMC3712628.
12. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol. 2014 Mar;32(3):279-84. PubMed PMID: 24463574. PubMed Central PMCID: PMC3988262.
13. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity Science. 2016 Jan 1;351(6268):84-8. PubMed PMID: 26628643.
14. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, Joung JK. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature. 2016 Jan 6. doi: 10.1038/nature16526. PubMed PMID: 26735016.
15. 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. *In vivo* genome editing using *Staphylococcus aureus* Cas9. Nature. 2015 Apr 9;520(7546):186-91. PubMed PMID: 25830891. PubMed Central PMCID: PMC4393360

HISTORY OF CRISPR (CONT'D)

16. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales AP, Li Z, Peterson RT, Yeh JR, Aryee MJ, Joung JK. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature. 2015 Jul 23;523(7561):481-5. PubMed PMID: 26098369.
17. 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. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell. 2015 Oct 22;163(3):759-71. PubMed PMID: 26422227.
18. Dickinson DJ, Pani AM, Heppert JK, Higgins CD, Goldstein B. Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette. Genetics. 2015 Aug;200(4):1035-49. PubMed PMID: 26044593.
19. Jao LE, Wente SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc Natl Acad Sci U S A. 2013 Aug 20;110(34):13904-9. PubMed PMID: 23918387.
20. Port F, Chen HM, Lee T, Bullock SL. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. Proc Natl Acad Sci U S A. 2014 Jul 22;111(29):E2967-76. PubMed PMID: 25002478.
21. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 2013 May 9;153(4):910-8. PubMed PMID: 23643243.
22. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic Screens in Human Cells Using the CRISPR/Cas9 System. Science. 2014 Jan 3;343(6166):80-4. PubMed PMID: 24336569. PubMed Central PMCID: PMC3972032.
23. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science. 2014 Jan 3;343(6166):84-7. PubMed PMID: 24336571. PubMed Central PMCID: PMC4089965.
24. 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. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell. 2014 Oct 8. PubMed PMID: 25307932. PubMed Central PMCID: PMC4253859.
25. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013 Jul 18;154(2):442-51. PubMed PMID: 23849981. PubMed Central PMCID: PMC3770145.
26. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell. 2013 Dec 19;155(7):1479-91. PubMed PMID: 24360272. PubMed Central PMCID: PMC3918502.
27. Fujita T, Fujii H. Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. Biochem Biophys Res Commun. 2013 Sep 13;439(1):132-6. PubMed PMID: 23942116.
28. Savic D, Partridge EC, Newberry KM, Smith SB, Meadows SK, Roberts BS, Mackiewicz M, Mendenhall EM, Myers RM. CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. Genome Research 2015 Oct;25(10):1581-9. PubMed PMID: 26355004.
29. Dalvai M, Loehr J, Jacquet K, Huard CC, Roques C, Herst P, Côté J, Doyon Y. A Scalable Genome-Editing-Based Approach for Mapping Multiprotein Complexes in Human Cells. Cell Rep. 2015 Oct 20;13(3):621-33. PubMed PMID: 26456817.
30. Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Rivera RM, Madhavan S, Pan X, Ran FA, Yan WX, Asokan A, Zhang F, Duan D, Gersbach CA. *In vivo* genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science. 2015 Dec 31. pii: aad5143. PubMed PMID: 26721684.
31. Tabebordbar M, Zhu K, Cheng JK, Chew WL, Widrick JJ, Yan WX, Maesner C, Wu EY, Xiao R, Ran FA, Cong L, Zhang F, Vandenberghe LH, Church GM, Wagers AJ. *In vivo* gene editing in dystrophic mouse muscle and muscle stem cells. Science. 2015 Dec 31. pii: aad5177. PubMed PMID: 26721686.
32. Long C, Amoasii L, Mireault AA, McAnally JR, Li H, Sanchez-Ortiz E, Bhattacharyya S, Shelton JM,

HISTORY OF CRISPR (CONT'D)

Bassel-Duby R, Olson EN. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science. 2015 Dec 31. pii: aad5725. PubMed PMID: 26721683.

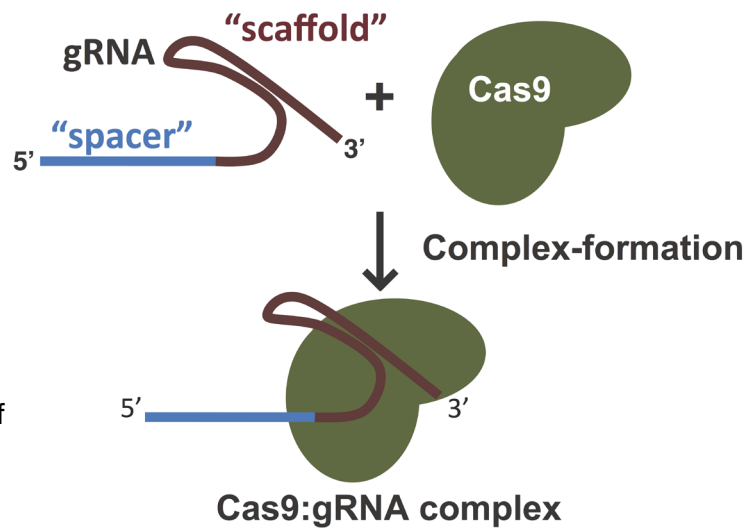
COMPONENTS OF CRISPR/CAS9

By Joel McDade, Tyler J. Ford | October, 2015

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. In this section we will go into a bit more detail about the components of these major players. After reading this section, 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. If you ever need to quickly look up any CRISPR associated terminology, be sure to hop down to the CRISPR glossary. Please note that, while this section is intended to provide a general overview of CRISPR components, new Cas9 variants are being discovered all the time. The requirements of these different systems can vary (for example, read our Cpf1 section for some of the interesting properties of this exciting new nuclease tool). You can keep up on all the exciting new developments in CRISPR research on our blog.

The Endonuclease

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 editing, the only CRISPR protein required is the Cas9 endonuclease or a variant thereof. This individual protein has all the components necessary to:



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)

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.

3. Cleave Target DNA Resulting in a Double-Strand Break (DSB)

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.

COMPONENTS OF CRISPR/CAS9 (CONT'D)

Cas9 Species/Variants and PAM Sequences

Species/Variant of Cas9	PAM Sequence
Streptococcus pyogenes (SP); SpCas9	NGG
SpCas9 D1135E variant	NGG (reduced NAG binding)
SpCas9 VRER variant	NGCG
SpCas9 EQR variant	NGAG
SpCas9 VQR variant	NGAN or NGNG
Staphylococcus aureus (SA); SaCas9	NNGRRT or NNGRR(N)
Neisseria meningitidis (NM)	NNNNGATT
Streptococcus thermophilus (ST)	NNAGAAW
Treponema denticola (TD)	NAAAAC
Cpf1 (from various species)	TTN
Additional Cas9s from various species	PAM sequence may not be characterized

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). Throughout the remainder of this eBook, 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 in a later section) and their preferred Cas9 variant (with the appropriate PAM sequence) to modify the desired genomic locus.

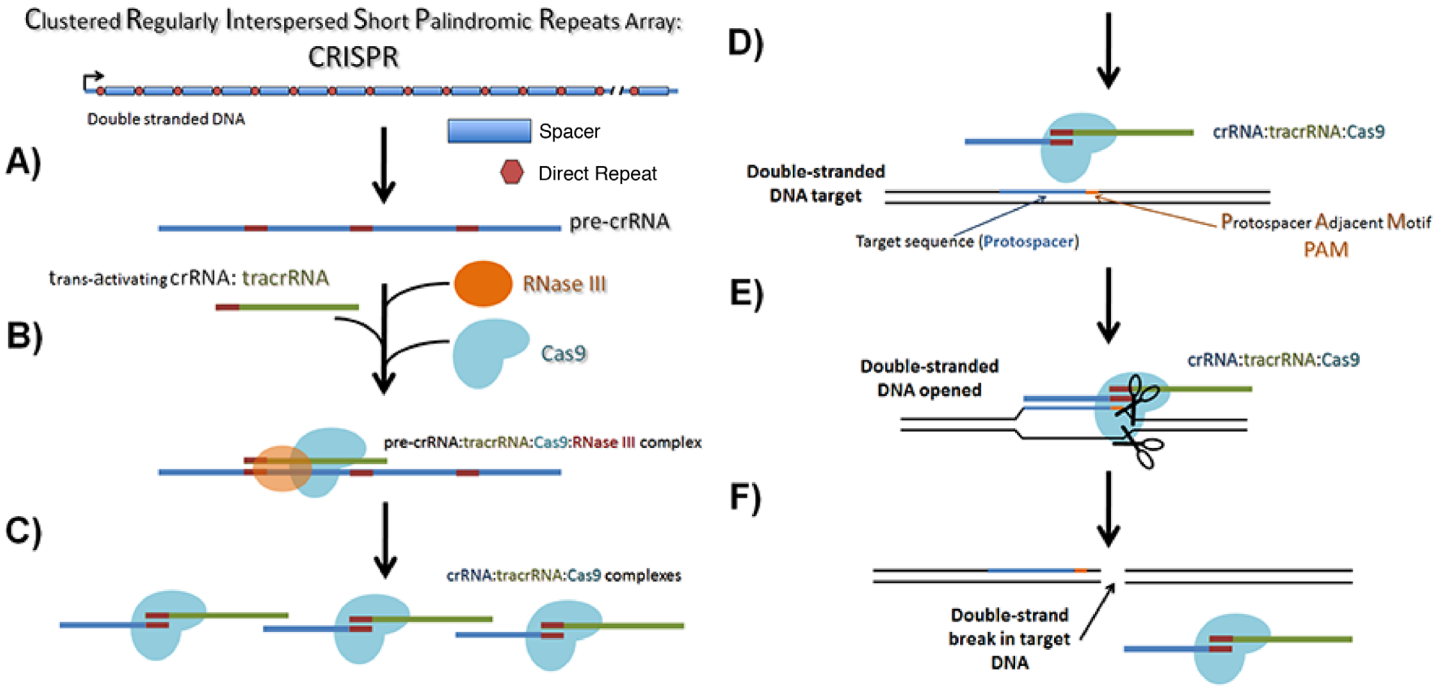
Some Historical Notes

CRISPR arrays in bacterial genomes consist of repeated elements separated by unique sequences. When researchers first discovered these arrays, they did not know their functions and simply called the repeated elements "direct repeats" and the unique stretches of DNA between them "spacers" (see figure on the next page).

After years of research, we now know that each direct repeat, combined with its adjacent spacer, ultimately encodes a single crRNA. The direct repeat regions contain sequences required for processing pre-crRNA into mature crRNA and tracrRNA binding. The spacer regions, on the other hand, are the unique, foreign DNA target sequences specific to each individual crRNA.

COMPONENTS OF CRISPR/CAS9 (CONT'D)

Bottom line, the "direct repeat region" combined with the tracrRNA forms the scaffold portion of a gRNA and the "spacer region" forms the target sequence.



Overview of native CRISPR arrays and their processing for cleaving foreign DNA. A) CRISPR arrays as found in a bacterial genomes are transcribed into pre-crRNAs containing both the spacer region and the direct repeat region. B) RNaseIII, the tracrRNA, and Cas9, bind to these transcripts and C) cleave them leaving mature crRNAs bound to the Cas9/tracrRNA complex. D) The mature crRNA is used to guide the Cas9 complex to the target DNA which is E) cleaved leaving a F) double-strand break. A "gRNA" is a researcher-designed hybrid of the tracrRNA and the crRNA. The "direct repeat region" combined with the tracrRNA forms the scaffold portion of a gRNA and the "spacer region" forms the target sequence.

CRISPR GLOSSARY

By Joel McDade | October, 2015

Below you can find a table defining many of the common terms used in articles and discussions about CRISPR research and CRISPR genome engineering. We hope that you'll find this glossary useful and maybe even print out a copy to keep at your desk for quick reference.

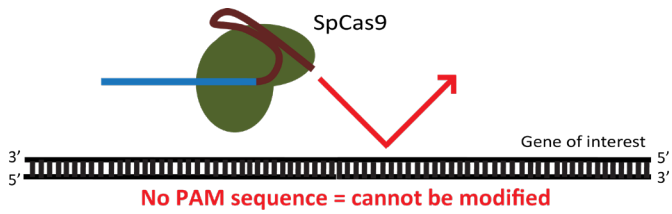
Term	Definition
Cas	CRISPR Associated Protein, the Cas9 nuclease is the active enzyme for the Type II CRISPR system
CRISPR	Clustered Regularly Interspaced Short Palendromic Repeat, a region in bacterial genomes used in pathogen defense
CRISPRa	CRISPR Activation, using a dCas9 or dCas9-activator with 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 Cas9
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, through the use of Cas9 or two Cas9-nickases targeting opposite strands
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
Genetic modification or manipulation	Any genetic perturbation, including genetic knock-out, gene activation, or gene repression
gRNA	Guide RNA, a synthetic fusion of the endogenous bacterial crRNA and tracrRNA; Provides both targeting specificity and scaffolding/binding ability for Cas9 nuclease; Does not exist in nature; Also referred to as "single guide RNA" or "sgRNA"
gRNA scaffold sequence	The sequence within the gRNA that is responsible for Cas9 binding; Does not include the 20bp spacer/targeting sequence that is used to guide Cas9 to target DNA

CRISPR GLOSSARY (CONT'D)

Term	Definition
gRNA targeting sequence	The 20 nucleotides that precede the PAM sequence in the genomic DNA; What gets put into a gRNA expression plasmid; 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 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 a double stranded DNA; Normally repaired by HDR
Nickase	Cas9 that has one of the two nuclease domains inactivated; Can be either the RuvC or HNH domain; Capable of cleaving only one strand of target DNA
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; Can be minimized by double-nick
On-target activity	Cas9 cleavage at a desired location due to gRNA targeting sequence with sufficient homology to recruit Cas9 to desired genomic locations
ORF	Open Reading Frame; The codons that make up a gene
PAM	Protospacer Adjacent Motif; Necessary for Cas9 to bind target DNA; Must immediately follow the target sequence
PCR	Polymerase Chain Reaction; Used to amplify a specific sequence of DNA
sgRNA	Single guide RNA; A synthetic RNA composed of a targeting sequence and scaffold sequence derived from endogenous bacterial crRNA and tracrRNA; Used to target Cas9 to a specific genomic locus in genome engineering experiments; Also referred to as a "gRNA"
Target sequence	Genomic target of the gRNA targeting sequence; The 20 nucleotides that are incorporated into the gRNA plus the PAM sequence; Not to be confused with the gRNA targeting sequence

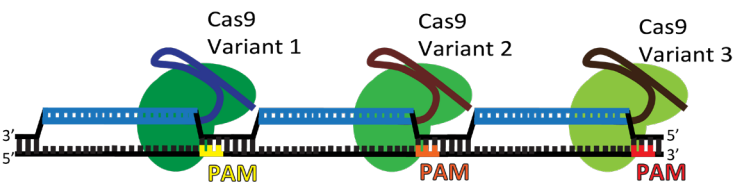
THE PAM REQUIREMENT AND EXPANDING CRISPR BEYOND SpCAS9

By Joel McDade | Nov 12, 2015



Diverse genomes and genomic targets require a variety of tools to engineer them effectively. Read on to learn how a variety of natural and engineered forms of Cas9 can be used to expand CRISPR's reach to new genomic loci.

The Need for More PAM Sequences



CRISPR/Cas9 can be used to modify any desired genomic target provided that sequence is unique compared to the rest of the genome and is located just upstream of a Protospacer Adjacent Motif (PAM sequence). A PAM sequence is a short stretch of DNA (typically 3-5 nucleotides) that serves as a binding signal for Cas9 and the presence of this sequence is a strict requirement for Cas9-mediated DNA cleavage. While PAM sequences for the commonly used *S. pyogenes* Cas9 (5'NGG'3) are abundant throughout the human genome, they are not always positioned correctly to target particular genes for Cas9-mediated knockout or modification. Furthermore, a target sequence may have high homology elsewhere in the genome. These additional sequences (so-called "off-targets") may be unintentionally mutated during attempts to use CRISPR mediated genome engineering on the gene of interest. The lack of suitable PAM sequences is of particular concern when trying to edit a gene using Homology Directed Repair (HDR), since HDR-mediated gene editing is most efficient when target sites are located in close proximity to the region to be edited. Therefore, the requirement for a single acceptable PAM sequence remains a technical hurdle in using CRISPR to modify genomic loci across the entire human genome. In this section, we will cover two ways to circumvent this limitation: 1) the use of novel *S. pyogenes* Cas9 variants with varying PAM sequences and 2) the use of Cas9 homologs derived from species other than *S. pyogenes*.

Synthetic *S. pyogenes* Cas9s

Kleinstiver et al. (1) 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. From these screens, they identified four novel SpCas9 variants with altered PAM binding specificity:

SpCas9 Variant	Mutations (relative to SpCas9)	PAM sequence (5' to 3')
D1135E variant	D1135E	NGG
VQR variant	D1335V, R1335Q and T1337R	NGAN or NGNG
EQR variant	D1135E, R1335Q and T1337R	NGAG
VRER variant	D1135V, G1218R, R1335E and T1337R	NGCG

Notably, the D1135E variant, which still recognizes the canonical *S. pyogenes* PAM sequence (5'NGG'3), is far more selective for the NGG PAM over the NGA PAM compared to wild-type SpCas9. This variant may therefore increase the specificity of genome modifications at DNA targets adjacent to NGG PAM sequences

THE PAM REQUIREMENT AND EXPANDING CRISPR BEYOND SpCAS9 (CONT'D)

when used in place of wild-type SpCas9. The remaining variants (VQR, EQR and VRER) recognize novel PAM sequences (shown above). Importantly, the VQR, EQR and VRER Cas9 variants are capable of cleaving genomic DNA when expressed in mammalian cells and zebrafish embryos and 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 for several genomic targets in human cells, indicating that the variants are likely just as selective as wild-type SpCas9. Inclusion of the VQR, EQR and VRER SpCas9 variants effectively doubles the targeting range of CRISPR/Cas9 within the human genome.

Characterization of Cas9 from Additional Bacterial Species

Over 20 additional Cas9 homologs derived from a variety of bacterial species have been isolated (many by Feng Zhang's lab at the Broad Institute) and the PAM binding specificity of at least 4 have been characterized (see table):

Cas9 Species	PAM sequence (5' to 3')
Streptococcus pyogenes (Sp)	NGG
Staphylococcus aureus (Sa)	NGRRT or NGRRN
Neisseria meningitidis (Nm)	NNNNGATT
Streptococcus thermophilus (St)	NNAGAAW
Treponema denticola (Td)	NAAAAC
~20 additional Cas9 species	PAM sequence may not be characterized

Non-SpCas9s bind a variety of PAM sequences, which may make them useful when no suitable SpCas9 PAM sequence is present within your gene of interest. Furthermore, non-SpCas9s may have other characteristics that make them more useful than SpCas9 for a specific application. For example, the coding sequence for Cas9 from *Staphylococcus aureus* (SaCas9) is about 1 kilobase smaller than SpCas9, which allows for packaging into adeno-associated virus (AAV). AAV-mediated delivery of SaCas9 has already been used to target liver cells of mice *in-vivo*, and the list of applications is certain to expand in the near-future. It is important to remember that non-SpCas9s are only compatible with the tracrRNA and crRNA (or synthetic gRNA) derived from the same species.

Characterization of Cas9 from Additional Bacterial Species

Isolation of novel CRISPR proteins or modification of existing CRISPR reagents has and will continue to dramatically increase the number of CRISPR applications. While *S. pyogenes* Cas9 has already been modified to enable gene knockout, repression and activation in a variety of cell types, it is certainly possible that novel Cas9s could be modified for similar approaches. Isolation of novel CRISPR proteins will also increase gene editing efficiencies and possibilities. An example is the recent identification of Cpf1, a non-Cas9 CRISPR nuclease that generates double strand breaks in target genes resulting in the formation of "sticky ends" rather than blunt ends (which is the case for all Cas9 proteins). As the number of CRISPR reagents continues to grow, so too will the number of reagents available through Addgene!

THE PAM REQUIREMENT AND EXPANDING CRISPR BEYOND SpCAS9 (CONT)

Further Reading

1. Kleinstiver, Benjamin P., et al. “Engineered CRISPR-Cas9 nucleases with altered PAM specificities.” Nature (2015). Pubmed PMID: 26098369.

CPF1: A CAS9 HOMOLOG

By Mary Gearing | October 14, 2015

CRISPR genome editing is no longer limited to Cas9 - there's a new nuclease in town! A recent paper from Feng Zhang's lab describes two genes from the Cpf1 family 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. For multiple reasons, Cpf1 may be even better suited for genome editing than Cas9 - read on to learn more about Cpf1.

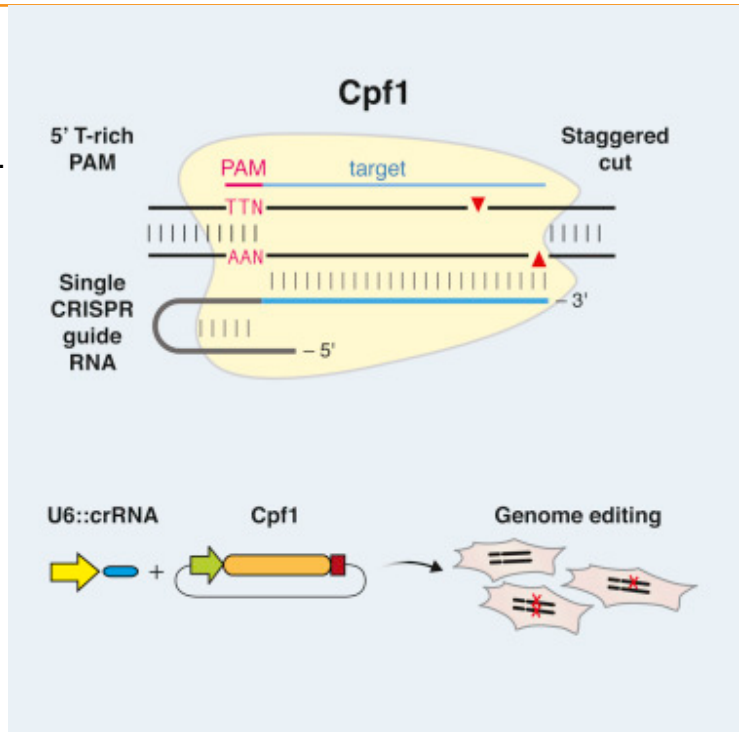


Image Attribution: Zetsche, Bernd, et al. "Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system." Cell (2015). PubMed PMID: 26422227.

How Was Cpf1 Found and Tested?

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 recently annotated in several genomes and is classified as a novel, 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' 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.

CPF1: A CAS9 HOMOLOG (CONT'D)

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, guide RNAs are only ~42 nt long. Direct synthesis of these gRNAs 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.

Zetsche et al. also suggest that Cpf1 may improve the frequency of HDR over non-homologous endjoining (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.

The application of Cpf1 to genome editing is exciting both in terms of basic science and translational applications. This putative type V CRISPR system proves we have a lot more to learn about CRISPR biology, and that future work may harness other nucleases superior for genome editing applications. Wherever the CRISPR journey leads us, Addgene will strive to provide all of the plasmids and resources you'll need to apply CRISPR technology to your own research!

Further Reading

1. Zetsche, Bernd, et al. "Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system." Cell (2015). PubMed PMID: 26422227.
2. Makarova, Kira S., et al. "An updated evolutionary classification of CRISPR-Cas systems." Nature Reviews Microbiology (2015). PubMed PMID: 26411297.

ENHANCING CRISPR TARGETING SPECIFICITY WITH eSpCAS9 AND SpCAS9-HF1

By Tyler J. Ford | Dec 16, 2015

As evidenced by all the CRISPR publications, press, and plasmids out there, it's obvious that CRISPR is a ground-breaking technology that's already had a huge impact on research and will be affecting our everyday lives very soon. Not only is CRISPR having effects on various biological disciplines, the base technology itself is constantly improving. Cas9 variants have been modified for genome editing, activating gene expression, visualizing genomic loci, and much more. Now, researchers from the Zhang and Joung labs have improved the on-target specificity of the Cas9 nuclease with two independently discovered CRISPR variants: eSpCas9 and SpCas9-HF1.

The Off-target Problem

It is well known that CRISPR/Cas9 genome editing can result in unwanted changes at non-target sites. Means to decrease these so-called "off-target effects" have included using a Cas9 nickase variant, lowering Cas9 expression, and truncating gRNA sequences used for targeting; however, these options can be cumbersome, can lower on-target efficiency, and sometimes even increase off-target effects, respectively. Recognizing these issues, researchers from the Zhang lab at the Broad Institute and the Joung lab at Harvard Medical School set out to decrease the off-target effects of CRISPR/Cas9 by altering the nuclease activity of the Cas9 nuclease itself (1).

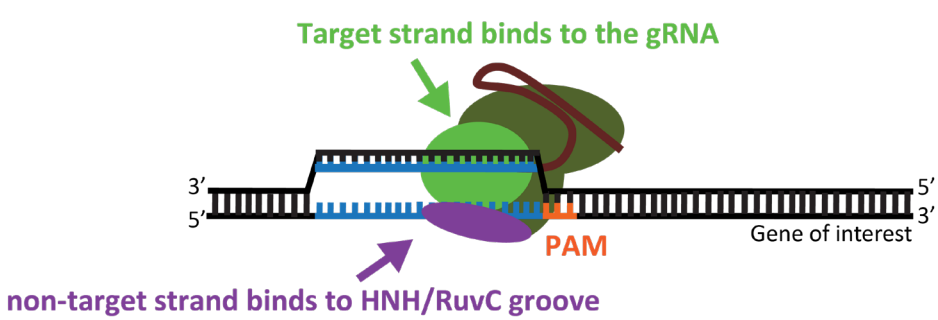
Design and Testing of eSpCas9

Looking at the structure of the Cas9 nuclease (PDB ID: 4OO8 and 4UN3), Slaymaker et al. hypothesized that Cas9 cutting efficiency increases when target DNA strand separation is stabilized (2,3). Stable strand separation is maintained by 2 sets of interactions :

1. Interactions between the non-target strand and a positively charged groove formed by the Cas9 HNH and RuvC nuclease domains

2. Interactions between the target strand and the gRNA (see figure below)

Because off-target sequences have less complementarity to the gRNA, they have a higher propensity to reform double helices and decrease Cas9 cutting efficiency. However, this is not always enough to keep off-target sites from being cut. Slaymaker et al. reasoned that, if they decreased the positive charge in the HNH/RuvC groove, then they could weaken interactions between the groove and the negatively charged DNA and therefore destabilize basal strand separation. This decreased strand separation combined with the weak separation afforded by gRNA binding at non-target sites would theoretically decrease off-target cutting.



Strand separation and target DNA binding to the Cas9 nuclease. Stable strand separation is maintained through interactions between the non-target strand and the HNH/RuvC groove and interactions between the target strand and the gRNA.

ENHANCING CRISPR TARGETING SPECIFICITY WITH eSpCAS9 AND SpCAS9-HF1 (CONT'D)

To decrease the electropositivity of the HNH/RuvC groove, Slaymaker et al. made a variety of alanine substitutions throughout the groove in 32 separate Cas9 mutants. When tested for their ability to cut the EMX1 locus in human embryonic kidney (HEK) cells, 11 of these initial mutants retained the on-target activity of wt Cas9 while decreasing off-target activity at known off-target loci. The authors went on to combine the best of these mutations and tested the combination mutants for their on-target and off-target activity at multiple genomic locations in HEK cells. Further testing of two of the mutants, SpCas9(K855A) and eSpCas9(1.1), followed by a technique for detecting genome-wide double strand breaks (BLESS), revealed that these mutants do not cause off-target effects at unanticipated sites and, as predicted, decrease off-target effects genome-wide.

Further work by Slaymaker et al. showed that these mutants are not toxic to HEK cells and that similar mutations can improve the specificity of Cas9 derived from S. aureus. This means you can likely apply similar rational mutations to your Cas9 homolog of choice and get improved on-target specificity.

Design and Testing of SpCas9-HF1

Kleinstiver et al. and the Joung Lab similarly thought that if they could weaken sequence independent interactions between Cas9 and DNA, then they could diminish off-target cutting (4). However, instead of weakening interactions between the non-target strand and Cas9, Kleinstiver et al disrupted interactions between Cas9 and the phophate backbone of the target DNA strand via mutations N497A, R661A, Q695A, and Q926A (SpCas9-HF1 contains all four mutations). Although the precise mutations were different than those found in eSpCas9, the effect on specificity was similar; SpCas9-HF1 generated fewer off-target cuts when compared to wt SpCas9 across a variety of genomic sites.

Future Possibilities

With their enhanced specificity, eSpCas9 and SpCas9-HF1 should enable researchers to make precise edits in mammalian cells and may decrease worries about off target effects in applied and/or therapeutic settings. Will combining the different mutations found in each of these great tools further enhance genome editing specificity? We hope one of you will be able to answer this question very soon! We look forward to seeing how the research community makes use of these tools and are excited to see continued improvements in genome engineering!

Further Reading

1. Slaymaker, Ian M., et al. "Rationally engineered Cas9 nucleases with improved specificity." Science (2015): aad5227. PubMed PMID: 26628643.
2. Nishimasu, Hiroshi, et al. "Crystal structure of Cas9 in complex with guide RNA and target DNA." Cell 156.5 (2014): 935-949. PubMed PMID: 24529477. PubMed Central PMCID: PMC4139937.
3. Anders, Carolin, et al. "Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease." Nature 513.7519 (2014): 569-573. PubMed PMID: 25079318. PubMed Central PMCID: PMC4176945.
4. Kleinstiver, Benjamin P., et al. "High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects." Nature (2016). PubMed PMID: 26735016.

SINGLE BASE EDITING WITH CRISPR

By Mary Gearing | August 16th, 2016

When we talk about CRISPR applications, one negative always comes up: the low editing efficiency of homology-directed repair (HDR). Compared to the random process of non-homologous end joining, HDR occurs at a relatively low frequency, and in non-dividing cells, this pathway is further downregulated. Like all CRISPR applications that use wild-type Cas9, editing by HDR also has some potential for off-target cleavage even when gRNAs are well-designed. Rather than try to improve HDR, Addgene depositor David Liu's lab created new Cas9 fusion proteins that act as "single base editors." These fusions contain dCas9 or Cas9 nickase and the rat cytidine deaminase APOBEC1, which can convert cytosine to uracil without cutting DNA. Uracil is subsequently converted to thymine through DNA replication or repair. Komor et al. estimate that hundreds of genetic diseases could be good targets for base editing therapy with these fusions, not to mention the potential basic and preclinical research applications. Read on to learn about this new way to make point mutations using CRISPR without double-stranded breaks.

dCas9 Fusion Proteins

When the CRISPR revolution first began, we were most excited about direct cleavage of target DNA. This application is still important, but catalytically dead dCas9 is almost equally as valuable as a targeting scaffold. While dCas9 can't cleave DNA, it can target a region of DNA with equal specificity to wild-type Cas9. dCas9 fusions have been used to regulate promoter activity, make epigenetic modifications, visualize genomic loci in living cells, and more.

A cytidine deaminase fusion to dCas9 makes a lot of sense when you take a closer look at the literature. Early work by Tsai et al. on CRISPR nickases showed that nickases could induce C->T mutations at a low frequency. They hypothesized that cytidine deaminases like APOBEC were the cause of these mutations, not Cas9. APOBEC and other DNA cytidine deaminases can only edit single-stranded DNA - a perfect scenario for Cas9 given that the target DNA and gRNA interact to displace a small portion of the nontargeted DNA strand. This specificity narrows down the number of cytosines potentially accessible to a dCas9-APOBEC fusion protein.

dCas9/Nickase Base Editing

At their very simplest, the requirements for using the Liu lab single base editors are:

- Cas9 fused to a cytidine deaminase
- A gRNA targeting Cas9 to a specific locus
- A target cytosine at position 4-8 in the non-targeted strand

Komor et al. created a first-generation base editor (BE1) using the rat cytidine deaminase APOBEC1 connected to dCas9 via a 16 base XTEN linker. To determine the editing capabilities of this protein, they incubated the purified base editor with DNA substrates containing Cs in positions 1-13 of the 20 base targeting sequence. After deamination, they used the uracil-specific USER enzyme to cleave deaminated DNA and determine deamination efficiency. BE1 displayed an activity window of approximately 5 nucleotides, from target positions 4 to 8. Across multiple targets, the average *in vitro* editing efficiency was 44%, with cytidines preceded by a T or a C edited at the highest rate. These editing rates are particularly impressive considering that when you modify only one strand of DNA, the maximum editing efficiency is 50%.

Once Komor et al. moved into a cell culture model, they noticed that editing efficiency dropped from 44% to 0.8-7.7%, likely due to base excision repair (BER) removing uracils from the DNA. Their second-generation BE2

SINGLE BASE EDITING WITH CRISPR (CONT'D)

system fuses an inhibitor of this process to dCas9, raising editing efficiency three-fold to a maximum of ~20%. For BE1 and BE2, indel formation is very low (<0.1%) since the DNA is not directly cleaved.

To increase base editing efficiency beyond 50%, you'd need a way to copy the edits into the opposite strand of DNA. To do so, Komor et al. turned dCas9 back into a nickase to simulate mismatch repair. 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.

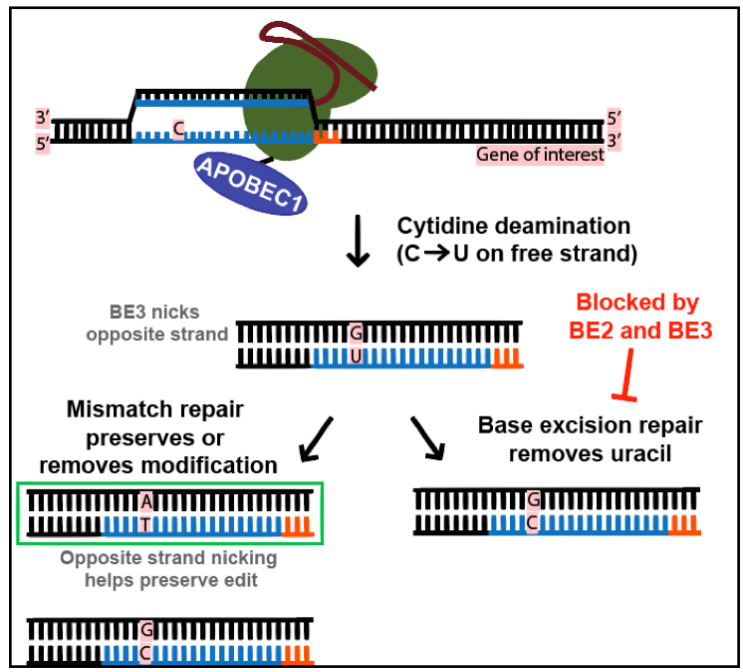


Figure 1: Potential outcomes of base editing. After cytidine deamination converts C to U, base excision repair may excise the U, eliminating the edit. Base editors BE2 and BE3 block base excision repair to encourage mismatch repair. With BE2, mismatch repair occurs randomly, propagating the edit about 50% of the time. BE3 employs Cas9 nickase to nick the strand opposite the edit, encouraging the cell to use the edited strand as a template and enhancing editing efficiency.

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; 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. Komor et al. also tested HDR and BE3 head-to-head for disease-relevant mutations in APOE4 and p53. Cas9-mediated HDR efficiency was very low (0.1-0.3%) or undetectable, but base editing efficiency ranged from 58-75% for APOE4 to 3-8% for p53.

Komor et al. also examined whether base editors can induce off-target mutations. Using previously characterized gRNAs with known off-target sequences and analyzing 10 off-target sequences per gRNA, they observed cytidine conversion in positions 4-8 for a subset of those off-targets. Unsurprisingly, the frequency of off-target deamination was highest at the sites where wild-type Cas9 induces the most off-target mutations, as previously determined using GUIDE-seq. The good news is that cytidines surrounding the gRNA sequences do not appear to be subject to off-target editing, indicating that off-target edits come from Cas9 activity, not APOBEC1 activity. Future base editors made using high-fidelity Cas9s could ameliorate the problem of off-target editing.

The Future of Base Editing

The work described in Komor et al. is incredibly creative and exciting, as it removes the need for DNA cleavage and repair according to a donor template. To really bring base editing to primetime, a few issues need to be worked out. The first, most obvious problem is the need for enzymes that can make conversions other than C->T, or for the opposite strand, G->A. Based on data from the NCBI ClinVar database, Komor et al. estimate that about 900 clinically relevant C->T or G->A variants are properly positioned near an NGG-PAM and could be edited using these first-generation editors. However, these SNPs represent a small fraction of the estimated 12,000 pathogenic SNPs in ClinVar, many more of which should be targetable with other base editing technologies. Luckily, given the number of DNA-modifying enzymes currently known, it seems like only a matter of time before new base editors are developed.

SINGLE BASE EDITING WITH CRISPR (CONT'D)

Comparison of the Gene Editing Technologies Studied in Komor et al.

Technique	HDR	BE1	BE2	BE3
Cas9	WT or Nickase	dCas9	dCas9	Nickase
Mutation Site	~10-100 bp from PAM (higher efficiency at shorter distances)	C in position 4-8 of targeting sequence on free strand	C in position 4-8 of targeting sequence on free strand	C in position 4-8 of targeting sequence on free strand
gRNA	Single (or paired with nickase)	Single	Single	Single
Repair Template	ssDNA or dsDNA; homology arm length varies with edit size	None	None	None
Edit Types	Insertions, deletions, point mutations	Point mutations (C->T or G->A)	Point mutations (C->T or G->A)	Point mutations (C->T or G->A)
Frequency of NHEJ at Target Locus	Indels may be more common at target site than HDR repair	Almost zero	Almost zero	Lower than HDR
Types of Off-target Effects	Indels at off-target loci	Conversion of other Cs in editing window and at off-target sites	Conversion of other Cs in editing window and at off-target sites	Conversion of other Cs in editing window and at off-target sites

The second potential hurdle for base editing is the small editing window available. For optimal editing, the targeted base should be at position 4-8 of the targeting sequence. Since PAM requirements limit the number of targetable sequences, using other Cas9 variants should expand targeting capacity. Other CRISPR enzymes like Cpf1 could also be adapted for base editing. Having more options for gRNAs would allow researchers to test for optimal gRNA on-target activity, likely increasing the frequency of base editing. This aspect will be especially important should base editing be used in clinical applications.

Recent work by Nishida et al. has further confirmed the utility of base editing while demonstrating the benefits of base editing variants. Nishida et al. created the Target-AID base editor using a cytidine deaminase from sea lamprey fused to Cas9 nickase. Target-AID acts similarly but not identically to BE3, modifying a 3-5 base window 18 bases upstream of the PAM instead of the 5 base window 15 bases upstream of the PAM seen with BE3. The authors suggest that the specific deaminase or the method of attachment to Cas9 may impact base editing action and efficiency.

Base editing efficiency will also likely increase as CRISPR delivery methods improve. In particular, ribonucleoprotein delivery may be a good option for base editing, as shorter BE expression could lower the editing frequency for other, non-preferred cytosines in the editing window, or for cytosines at off-target loci.

SINGLE BASE EDITING WITH CRISPR (CONT'D)

Base editing is the latest triumph for the CRISPR field, but it's important to remember that this advance builds on previous work. Structural and functional annotation of Cas9 has given researchers the knowledge to improve Cas9 beyond its original capabilities, as previously seen in the engineering of high fidelity Cas9s. New dCas9 fusions are popping up all the time, combining the benefits of CRISPR targeting with many new applications.

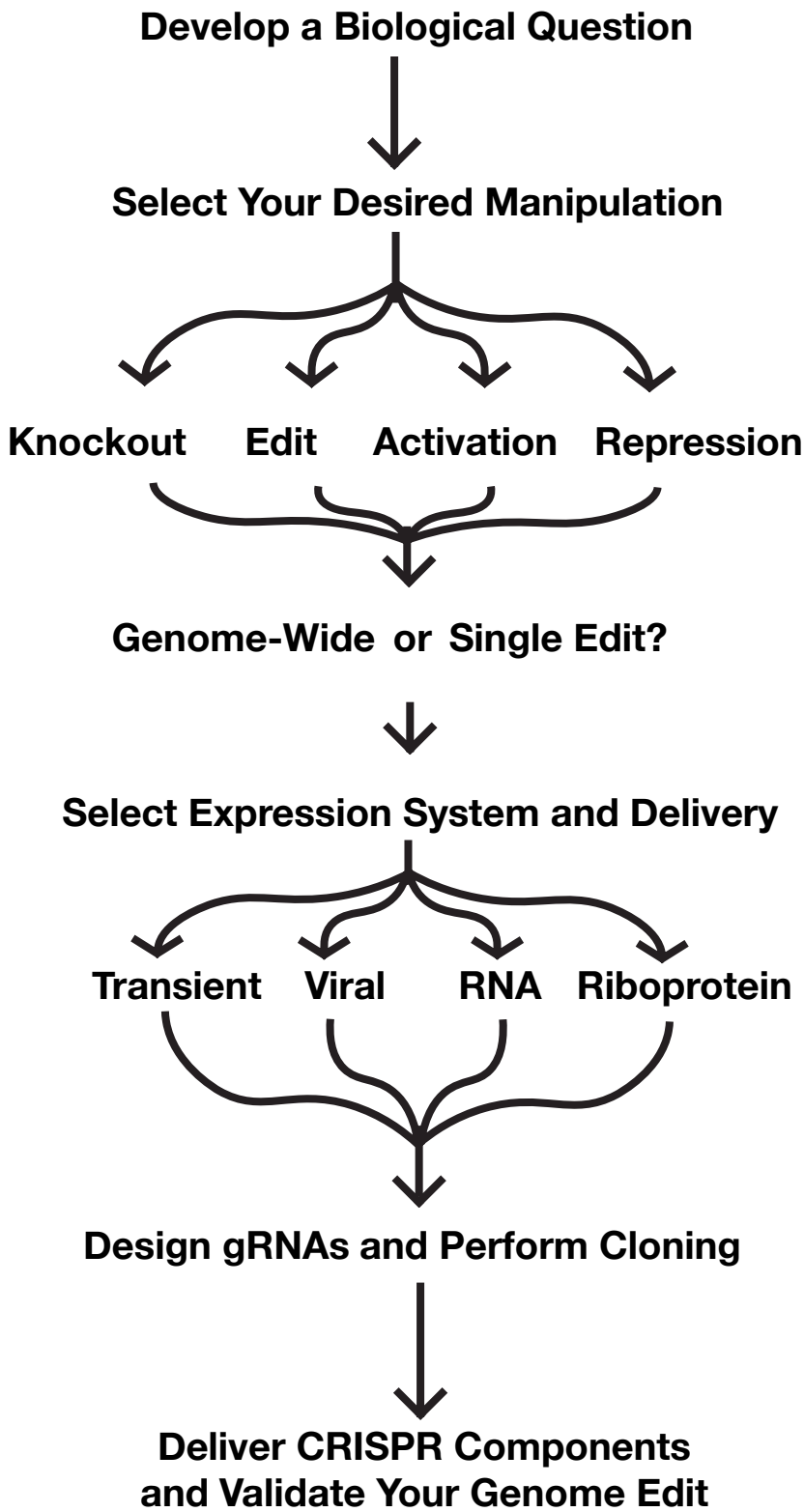
Further Reading

1. Komor, Alexis C., Yongjoo B. Kim, Michael S. Packer, John A. Zuris & David R. Liu. "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage." Nature (2016). doi:10.1038/nature17946 Epub 20 April 2016. PubMed PMID: 27096365.
2. Nishida, Keiji, et al. "Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems." Science (2016). doi:10.1126/science.aaf8729 Epub 4 August 2016. PubMed PMID: 27492474.
3. Tsai, Shengdar Q, et al. "Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing." Nat Biotechnol. 32(6):569-76 (2014). PubMed PMID: 24770325. PubMed Central PMCID: PMC4090141.
4. Tsai, Shengdar Q, et al. sGUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR–Cas nucleases." Nat Biotechnol. 33(2):187–197 (2015). PubMed PMID: 25513782. PubMed Central PMCID: PMC4320685.

USING CRISPR IN YOUR EXPERIMENTS

PLANNING YOUR CRISPR EXPERIMENT

By Joel McDade | Decmber, 2015



PLANNING YOUR CRISPR EXPERIMENT (CONT'D)

Getting Started

CRISPR/Cas9 is a powerful system that enables researchers to manipulate the genome of target cells like never before. This section will provide a general framework to get you started using CRISPR/Cas9 in your research. Although we will focus on using CRISPR in mammalian cells, many of these principles apply to using CRISPR in other organisms. We will go more in-depth into many of the items discussed here later in the eBook and provide links that allow you to jump to the appropriate chapters throughout this section. First, consider the genetic manipulation that is necessary to model your specific disease or process of interest. For instance, do you want to:

- Generate complete and permanent loss of gene expression or function (i.e. knock-out)?
- Express a mutated version of the gene (e.g. point mutant)?
- Increase or decrease expression of a target gene?

Once you have a clear understanding of what you are trying to do, you are ready to start navigating the different reagents that are available for your particular experiment.

Select Your Desired Genetic Manipulation

Different genetic manipulations require different CRISPR components. Selecting a specific genetic manipulation can be a good way to narrow down which reagents are appropriate for a given experiment. Make sure to check whether reagents are available to carry out your specific genetic manipulation in your specific model organism. There may not be a perfect plasmid for your specific application, and in such a case, it may be necessary to customize an existing reagent to suit your needs. You can learn more about selecting a Cas9 for your specific experiment here. Many CRISPR plasmids can be found organized by function on our website.

Genetic Manipulation	Application	Cas9	gRNA	Additional Considerations
Knock-out	Permanently disrupt gene function in a particular cell type or organism without regard for specific mutation	Cas9 (or Cas9 nickase)	Single (or dual) gRNA targeting 5′ exon or essential protein domains	Dual-nickase approach increases specificity but is less efficient. eSpCas9 can also be used to increase specificity.
Edit	Generate a specific user-defined sequence change in a particular gene, such as generating a point mutation or inserting a tag	Cas9 (or Cas9 nickase)	Single (or dual) gRNA targeting the region where the edit should be made	Requires DNA template for HDR; Efficiency is reduced compared to knock-out

PLANNING YOUR CRISPR EXPERIMENT (CONT'D)

Genetic Manipulation	Application	Cas9	gRNA	Additional Considerations
Repress or Interfere (CRISPRi)	Reduce expression of a particular gene(s) without permanently modifying the genome	dCas9 or dCas9-repressor (such as dCas9-KRAB)	gRNA(s) targeting promoter elements of target gene	dCas9-KRAB is more effective than dCas9 alone for mammalian cell lines
Activate (CRISPRa)	Increase expression of an endogenous gene(s) without permanently modifying the genome	dCas9-activator (such as dCas9-VP64)	gRNA(s) targeting promoter elements of target gene	Many different activators exist, all derived from S. pyogenes dCas9

Select Expression System

To use the CRISPR system, you will need to express both a gRNA and Cas9 in your target cells. The expression system you need will depend upon your specific application. For example, certain cell types are easier to transfect than others. For easy-to-transfect cell types (e.g. HEK293 cells), transfection with standard transfection reagents may be sufficient to express both your gRNA and Cas9. For more difficult cells (e.g. primary cells), viral delivery of CRISPR reagents may be more appropriate. See Chapter 4 on CRISPR expression systems and delivery methods for more information

The table below summarizes the major expression systems and variables for using CRISPR in mammalian cells. Some of the variables include:

- Species of Cas9 and gRNA
- Species of promoter and expression pattern of promoter for Cas9 and gRNA
- Presence of a selectable marker (drug or fluorophore)
- Delivery method

Expression System	Components of System	Application
Mammalian expression vector	• Cas9 promoter can be constitutive (CMV, EF1alpha, CBh) or inducible (Tet-ON); U6 promoter is typically used for gRNA • May contain reporter gene (e.g. GFP) to identify and enrich positive cells, or selection marker to generate stable cell lines (for details, see Addgene's Plasmids 101: Mammalian vectors)	Transient or stable expression of Cas9 and/or gRNA in a mammalian cell line that can be transfected at high efficiency

PLANNING YOUR CRISPR EXPERIMENT (CONT'D)

Expression System	Components of System	Application
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) or selection marker to identify and enrich positive cells • Packaging and envelope plasmids provide the necessary components to make lentiviral particles (for details about lentivirus, see Addgene's Lentivirus guide)	• 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/Cas9
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 (for details, see Addgene's AAV guide)	• 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, then delivered to target cells (i.e. 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 riboprotein 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

PLANNING YOUR CRISPR EXPERIMENT (CONT'D)

Additional Addgene Resources:

- Depositor protocols for plasmids at Addgene
- JOVE Video Reprint: Protocol for genomic deletions in mammalian cell lines

Select Your Target Sequence and Design Your gRNA

Once you have selected your CRISPR components and method of delivery, you are ready to select a target sequence and design your gRNA. A general overview of how to design a gRNA is presented below. More details on gRNA design can be found in our How to Design Your gRNA section. An overview of CRISPR software tools for gRNA design can be found in the CRISPR Software Matchmaker section.

A. Know your cell line and genome sequence

The cell line you choose determines a variety of factors related to your experiment. The genomic sequence used to design gRNAs will depend upon the gene in question and the species from which your cells were derived. When possible, you should sequence the region you are planning to modify prior to designing your gRNA, as sequence variation between your gRNA targeting sequence and target DNA may result in reduced cleavage. The number of alleles for each gene may vary depending on the specific cell line or organism, which may affect the observed efficiency of knock-out or knock-in using CRISPR/Cas9.

B. Select gene and genetic element to be manipulated

In order to manipulate a given gene using CRISPR/Cas9, you will have to identify the genomic sequence for the gene you are trying to target. However, the exact region of the gene you target will depend on your specific application. For example:

- To activate or repress a target gene using dCas9-activators or dCas9-repressors, gRNAs should be targeted to the promoter driving expression of your gene of interest.
- For genetic knock-outs, gRNAs commonly target 5′ constitutively expressed exons, which reduces the chances that the targeted region will be removed from the mRNA due to alternative splicing. Exons near the N-terminus are also often targeted since frameshift mutations here will increase the likelihood that a nonfunctional protein product is produced.
- Alternatively, gRNAs can be designed to target exons that code for known essential protein domains. The benefit of this approach is that non-frameshift mutations such as insertions or deletions (which commonly result from DSBs) are more likely to alter protein function when they occur in protein domains that are essential for protein function.
- For gene editing experiments using HDR, it is essential that the target sequence be close to the location of the desired edit. In this case, it is necessary to identify the exact location where the edit should occur and select a target sequence nearby.

C. Select gRNAs based on predicted "on-target" and "off-target" activity

A PAM sequence is absolutely necessary for Cas9 to bind target DNA. As such, one must identify all PAM sequences within the genetic region to be targeted (PAM is 5′ NGG 3′ for SpCas9). If there are no PAM sequences within your desired target, you may want to consider using a Cas9 from a different species or a *S. pyogenes* Cas9 variant that binds a PAM sequence present in your genomic target (see additional Cas9 variants

PLANNING YOUR CRISPR EXPERIMENT (CONT'D)

and PAM sequences). Once all possible PAM sequences and putative target sites have been identified, it is time to choose which site is likely to result in the most efficient on-target cleavage.

Clearly, the gRNA targeting sequence needs to match the target site, but it is also critical to ensure that the gRNA targeting sequence does NOT match additional sites within the genome. In a perfect world, your gRNA targeting sequence would have perfect homology to your target with no homology elsewhere in the genome. Realistically, a given gRNA targeting sequence will have additional sites throughout the genome where partial homology exists. These sites are called "off-targets" and should be considered when designing a gRNA for your experiments. In general, off-target sites are not cleaved as efficiently when mismatches occur near the PAM sequence, so gRNAs with no homology or those with mismatches close to the PAM sequence will have the highest specificity. Cas9 variants eSpCas9 and SpCas9-HF1, were designed by the Zhang lab at the Broad Institute to reduce "off-target activity."

In addition to "off-target activity", it is also important to consider factors that maximize cleavage of the desired target sequence ("on-target activity"). It is now understood that two gRNA targeting sequences, each having 100% homology to the target DNA may not result in equivalent cleavage efficiency. In fact, cleavage efficiency may increase or decrease depending upon the specific nucleotides within the selected target sequence. For example, gRNA targeting sequences containing a G nucleotide at position 20 (1 bp upstream of the PAM) may be more efficacious than gRNAs containing a C nucleotide at the same position in spite of being a perfect match for the target sequence. Therefore, close examination of predicted on-target and off-target activity of each potential gRNA targeting sequence is necessary to design the best gRNA for your experiment.

Several gRNA design programs have been developed that are capable of locating potential PAM and target sequences and ranking the associated gRNAs based on their predicted on-target and off-target activity (see The CRISPR Software Matchmaker Section). Additionally, many plasmids containing "validated" gRNAs are now available through Addgene. These plasmids contain gRNAs that have been used successfully in genome engineering experiments and have been published in peer-reviewed journals. Using validated gRNAs can save your lab valuable time and resources when carrying out experiments using CRISPR/Cas9.

Browse Plasmids: Validated gRNAs

D. Synthesize and clone desired gRNAs

Once you have selected your target sequences, it is time to design your gRNA oligos and clone these oligos into your desired vector. In many cases, targeting oligos are synthesized, annealed, and inserted into plasmids containing the gRNA scaffold using standard restriction-ligation cloning. However, the exact cloning strategy will depend on the gRNA vector you have chosen, so it is best to review the protocol associated with the specific plasmid in question (see CRISPR protocols from Addgene depositors).

E. Deliver Cas9 and gRNA

Your method of delivery will be determined by the specific expression system that you have chosen. It may be necessary to optimize a protocol for delivery of gRNA and Cas9 to your target cells, as transfection efficiency will vary based on the method of delivery and the cell type tested. Read the Chapter 4 on CRISPR expression systems and delivery methods for more information.

F. Validate genetic modification

PLANNING YOUR CRISPR EXPERIMENT (CONT'D)

Once you have successfully delivered the gRNA and Cas9 to your target cells, it is time to validate your genome edit. Delivery of Cas9 and gRNA to wild-type cells will result in several possible genotypes within the resulting "mutant" cell population. A percentage of cells may be wild-type due to either (1) a lack of gRNA and/or Cas9 expression, or (2) a lack of efficient target cleavage in cells expressing both Cas9 and gRNA. Cells that have undergone modification (mutant cells) can be homozygous (modification of both alleles) or heterozygous (modification of a single allele). Furthermore, in mutant cells containing two mutated alleles, each mutated allele may be different owing to the error-prone nature of NHEJ. Even for gene editing experiments using HDR, not every mutated allele will contain the desired edit as a large percentage of DSBs are still repaired by NHEJ. Therefore, the end result of most experiments is a heterogeneous population of cells containing a wide variety of mutations or edits within target genes.

How do you determine that your desired edit has occurred? The exact method necessary to validate your edit will depend upon your specific application, and in some cases, new approaches must be devised. However, there are several common ways to verify that your cells contain your desired edit, including, but not limited to:

1. Mismatch-cleavage assay (for NHEJ repaired DSBs): Provides a semi-quantitative readout of the percentage of alleles that have been mutated within a mixed cell population. The region of interest is PCR amplified, PCR products are denatured, renatured, treated with a nuclease that cleaves DNA heteroduplexes, and run on an agarose gel to identify DNA fragments.

2. PCR and restriction digest (for HDR repaired DSBs): For small nucleotide edits that introduce a novel restriction site. The region of interest is PCR amplified, digested with the appropriate restriction enzyme and run on an agarose gel to identify DNA fragments.

3. PCR amplification and gel electrophoresis (for HDR or NHEJ): For large deletions or insertions, the region of interest can be PCR amplified using primers that (A) flank the region of interest (deletions or small insertions) or (B) span the genome-insert boundary (insertions only). The PCR product is then run on an agarose gel to determine whether the edit was successful.

4. PCR amplification, subcloning and Sanger sequencing (for HDR or NHEJ): Provides semi-quantitative assessment of targeting frequency and exact sequence of targeted alleles. Involves PCR amplification of targeted region from DNA, subcloning into a plasmid, and screening individual clones.

5. PCR amplification and next-generation sequencing (for HDR or NHEJ): Provides quantitative assessment of the genome edits in your target sequence and can also be used to examine off-target effects.

6. More information on each of these techniques can be found in the Validating your Genome Edit section.

WHICH CAS9 DO I CHOOSE FOR MY CRISPR EXPERIMENT?

By Joel McDade | Jan 19, 2016

The advent of CRISPR/Cas9 made it easier than ever to efficiently make precise, targeted genome modifications. Cas9 has been modified to enable researchers to knock out, activate, repress or even image your favorite gene. But, with such a wide variety of Cas9-based reagents available, how do you choose which Cas9 is right for your particular experiment? This section will help familiarize you with the wide array of Cas9s available through Addgene's repository and make it easy to select the Cas9 reagent that is right for you.

The first thing to do in any CRISPR experiment is identify what, exactly, you are trying to accomplish. Are you trying to permanently knock-out a gene in a cell type or organism? Are you trying to reduce expression of a particular gene without permanently modifying the genome? Does it make more sense to try and activate at a particular locus? What about modifying the epigenome at a particular location? As you might expect, the answer to these questions will substantially affect your decision about which Cas9 you need for your experiment. Below is a brief summary of a few of the common genetic manipulations one can carry out using Cas9 and the specific Cas9s that can be used for each.

Repair of Cas9-Induced Breaks by NHEJ

Carrying out a genetic knockout using standard SpCas9

Knock-out cells or animals are created by co-expressing a gRNA specific to the gene to be targeted and the endonuclease Cas9. The genomic target can be any ~20 nucleotide DNA sequence, provided it meets two conditions: 1) The sequence is unique compared to the rest of the genome. 2) The target is present immediately upstream of a Protospacer Adjacent Motif (PAM). SpCas9 is a common choice when there is a suitable target site and little concern for off-target effects (e.g. optimal gRNA design with minimal homologous sites throughout the genome). Addgene carries a wide variety of SpCas9 containing plasmids that have been optimized for carrying out genome engineering experiments in bacteria, yeast, worms, *drosophila* and mammals.

Concerns about specificity? Consider using high-specificity Cas9 variants.

The specificity of Cas9-mediated cleavage can be enhanced by properly designing your gRNA sequence. That is, choosing a target sequence that has minimal homology elsewhere in the genome. A variety of approaches have been used to further enhance Cas9 specificity. For example, Cas9-nickase (Cas9n) takes advantage of the fact that Cas9 makes double-strand breaks (DSBs) through the combined activity of two nuclease domains, RuvC and HNH. Converting one of the two critical enzymatic residues to an alanine (D10A or H840A) generates a "nicking" Cas9 that cannot generate a double-strand break. Thus, two properly targeted Cas9n molecules are required to efficiently create DSBs at the target locus, which greatly enhances specificity compared to wild-

WHICH CAS9 DO I CHOOSE FOR MY CRISPR EXPERIMENT? (CONT'D)

type SpCas9.

While Cas9n is certainly more specific than wild-type SpCas9, DSBs are still detectable at target sites when only one gRNA is expressed with Cas9n. This means it's possible for Cas9n to bind and cause an indel at one of the off-target sites for either of its gRNAs. One can overcome this limitation through the use of nuclease dead Cas9 (dCas9) fused to the non-specific endonuclease FokI. FokI only cleaves target DNA when dimerized. Therefore, dCas9-FokI essentially requires proper targeting of two dCas9-FokI molecules at the target site before any cleavage occurs; dCas9-FokI is much less likely to cut at an off target specified by a single gRNA than Cas9n.

An obvious limitation of the Cas9n or dCas9-FokI approach is that they both necessitate two suitable target sequences in close proximity in order to efficiently generate a DSB. Several labs, including Feng Zhang's lab at the Broad Institute and Keith Joung's group at MGH, have used structural biology to identify key residues that mediate Cas9's ability to cleave off-target sites. So-called "enhanced Cas9" (Zhang) or "high-fidelity Cas9" (Joung) have comparable cleavage activity to wild-type SpCas9 at target loci but have greatly reduced off-target activity. These Cas9 variants enhance specificity without requiring two or more adjacent target sites within the target locus. More information on the enhanced specificity Cas9 variants can be found in our section entitled "Enhancing CRISPR Targeting Specificity with eSpCas9 and SpCas9-HF1".

What do I do when there is no 5'NGG'3 PAM sequence present?

- **Synthetic Cas9s with novel PAM recognition**

Through a series of positive selection screens in bacteria, Keith Joung's group identified mutants of S. Pyogenes Cas9 (VQR, EQR and VRER Cas9 variants) that recognize novel non-NGG PAM sequences. Importantly, the VQR, EQR and VRER Cas9 variants are capable of modifying genomic loci that cannot be modified using wild-type SpCas9, and their specificity for the PAM variants is similar to wild-type SpCas9 for several genomic targets in human cells. Including the VQR, EQR and VRER SpCas9 variants effectively doubles the targeting range of CRISPR/Cas9 within the human genome. More information on the various synthetic Cas9s available through Addgene can be found in our section entitled "The PAM Requirement and Expanding CRISPR Beyond SpCas9".

- **Non *S. pyogenes* Cas9s**

Additional Cas9 homologs have been isolated from a wide variety of bacterial species and many bind PAM sequences other than the typical NGG PAM sequence. So called "non-Sp" Cas9s may be more suitable for your experiment for reasons other than the PAM sequence. For example, the coding sequence for Cas9 from *Staphylococcus aureus* (SaCas9) is about 1 kilobase smaller than SpCas9, which allows for packaging into adeno-associated virus (AAV), the current gold standard for gene therapy. It is important to remember that non-SpCas9s are only compatible with the tracrRNA and crRNA (or synthetic gRNA) derived from the same species. More information on the various non-SpCas9s that are available through Addgene can be found in our section entitled "The PAM Requirement and Expanding CRISPR Beyond SpCas9".

WHICH CAS9 DO I CHOOSE FOR MY CRISPR EXPERIMENT? (CONT'D)

- **Non-Cas9 CRISPR endonuclease, Cpf1**

A new CRISPR endonuclease was recently identified by the Zhang lab at the Broad institute. This non-Cas9 CRISPR endonuclease, termed Cpf1, is a class II CRISPR endonuclease that is capable of cleaving target DNA in an RNA dependent manner. What makes Cpf1 so interesting is not the similarities to Cas9 (it has many) but the differences. The PAM sequence for Cpf1 is 5' TTN 3' and is located 5' to the target site (in contrast to the Cas9 PAM which is 3' to the target site). The endonuclease contains two enzymatic residues similar to Cas9 (D917 and E1006), but both residues are located in the RuvC domain (Cpf1 lacks an HNH domain) and mutation of either residue completely abolishes DNA cleavage. In contrast, SpCas9 can be converted into a nickase by mutating a single enzymatic residue. Intriguingly, Cpf1 cleavage results in a 5 nucleotide 5' overhang 18 base pairs from the PAM sequence. This is different from Cas9 cutting, which results in blunt DNA ends 3 base pairs distal to the PAM sequence. Whether or not the staggered cutting pattern of Cpf1 will allow researchers to avoid using the inefficient HDR-dependent pathway for gene editing remains to be demonstrated. Regardless, it is likely that the most exciting applications of Cpf1 are yet to come. You can find the Cpf1 plasmids and links to the publication on the Addgene website and read a little more about it in the CPf1 section of this eBook.

Activation and Repression of Target Genes Using Cas9

One unique aspect of Cas9 is its ability to bind target DNA independently of its ability to cleave target DNA. In other words, Cas9 containing mutations that disrupt DNA cleavage (D10A and H840A for SpCas9), is still capable of binding target DNA. Further, so-called nuclease dead Cas9 or "dCas9" can be used as a platform to deliver various cargo to specific DNA loci by fusing them directly to dCas9. This property of dCas9 has been exploited to localize a diverse array of proteins to target genes, including transcriptional activators and transcriptional repressors (see our section on Cas9-based gene regulation).

- **Repressing target genes using dCas9-based repressors**

Early experiments using dCas9 in bacteria demonstrated that targeting dCas9 to transcriptional start sites was sufficient to "repress" or "knock-down" transcription by blocking transcription initiation. In mammalian cells, robust repression requires targeting dCas9 tagged with transcriptional repressors to the promoter region of the gene of interest. You may want to consider using dCas9-based repressors including dCas9-KRAB when knocking out your target gene is toxic to cells. Plasmids for repressing target genes in a variety of species and cell types can be found on Addgene's website.

- **Activating target genes using dCas9-based activators**

The array of dCas9-based activators is quite diverse. The simplest activator consists of dCas9 fused to a single transcriptional activation domain (typically VP64). A second generation of activators has recently been developed and these alter gene expression using a few different approaches. For example, the "SunTag" system facilitates recruitment of multiple activation domains to the same genetic locus through co-expression of epitope-tagged dCas9 and antibody-activator fusion proteins. The Synergistic Activation Mediator complex consists of a dCas9-VP64 fusion and a modified gRNA that is capable of interacting with an additional RNA-binding transcriptional activator. Additional plasmids for activating target genes in a variety of species and cell types can be found on Addgene's website.

WHICH CAS9 DO I CHOOSE FOR MY CRISPR EXPERIMENT? (CONT'D)

OK, I've Selected a Cas9! What Do I Do Next?

You have selected the Cas9 that is right for your experiment. What do you do next? Here are some considerations that will help you move your CRISPR experiment forward.

Designing or selecting a gRNA - Addgene may already carry a "validated gRNA" for your favorite gene from your favorite species. Validated gRNAs have been used successfully in experiments and published in peer reviewed journals and will save your lab time and money when starting a CRISPR experiment. If you need to generate a new gRNA for your experiment, you can select from a variety of "empty gRNA vectors" and design your gRNA targeting sequence using one of the many freely available gRNA design programs.

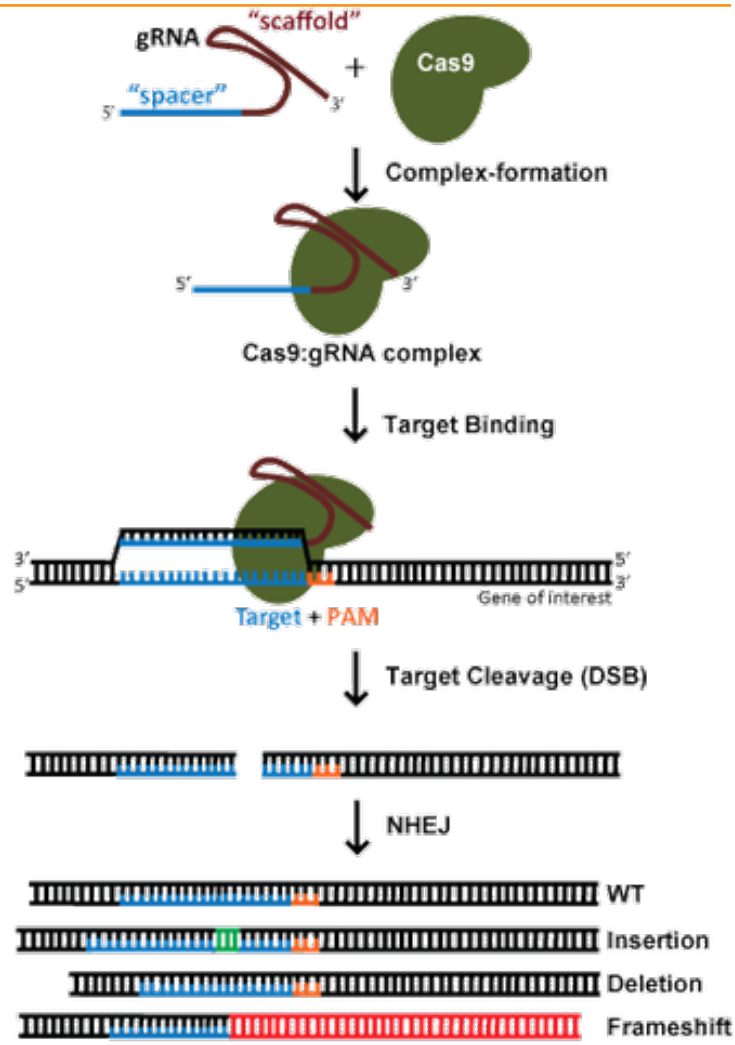
What type of expression system or delivery method should I use? Selecting the appropriate Cas9 reagent for your experiment is only half the battle - now you have to select the appropriate expression system and deliver Cas9 to your target cells! Addgene carries a variety of Cas9 containing plasmids that have been optimized for expression in different species and cell types including (but not limited to) bacteria, yeast, plants, *drosophila*, worms and mammals (Browse plasmids by model organism, here). For difficult to transfect cell types, you may want to consider using lentivirus to deliver Cas9 to your target cells. More information on the various delivery systems for mammalian cells can be found in our section entitled, "Mammalian Expression Systems and Delivery Methods".

Validating your genome edit - Once you have delivered Cas9 and a gRNA to your target cells, it is time to confirm that your target sequence has been modified! The exact protocol may vary depending on your specific experiment, but a broad overview of the various ways in which you can verify your genome edit can be found in section entitled "Validating Your Genome Edit".

USING CAS9 TO GENERATE KNOCK-OUTS VIA NHEJ

By Joel McDade, David Wyatt, Dale Ramsden, Tyler J. Ford | Compiled on November 17th, 2015

This post was contributed by David Wyatt and Dale Ramsden, UNC at Chapel Hill.



CRISPR/Cas9 can be used to generate knock-out cells or animals by co-expressing a gRNA specific to the gene to be targeted and the endonuclease Cas9. The genomic target can be any ~20 nucleotide DNA sequence, provided it meets two conditions:

1. The sequence is unique compared to the rest of the genome.
2. The target is present immediately upstream of a Protospacer Adjacent Motif (PAM).

The PAM sequence is absolutely necessary for target binding and the exact sequence is dependent upon the species of Cas9 (5′ NGG 3′ for *Streptococcus pyogenes* Cas9). We will focus on Cas9 from *S. pyogenes* as it is currently the most widely used in genome engineering. Once expressed, the Cas9 protein and the gRNA form a riboprotein complex through interactions between the gRNA "scaffold" domain and surface-exposed positively-charged grooves on Cas9. Cas9 undergoes a conformational change upon gRNA binding that shifts the molecule from an inactive, non-DNA binding conformation, into an active DNA-binding conformation. Importantly, the "spacer" sequence of the gRNA remains free to interact with target DNA. The Cas9-gRNA complex will bind any genomic sequence with a PAM, but the extent to which the gRNA spacer matches the target DNA determines whether Cas9 will cut. Once the Cas9-gRNA complex binds a putative DNA target, a "seed" sequence at the 3′ end of the gRNA targeting sequence begins to anneal to the target DNA. If the seed and target DNA sequences match, the gRNA will continue to anneal to the target DNA in a 3′ to 5′ direction.

Cas9 will only cleave the target if sufficient homology exists between the gRNA spacer and target sequences. The "zipper-like" annealing mechanics may explain why mismatches between the target sequence in the 3′ seed sequence completely abolish target cleavage, whereas mismatches toward the 5′ end are permissive for target cleavage. The Cas9 nuclease has two functional endonuclease domains: RuvC and HNH. Cas9 undergoes a second conformational change upon target binding that positions the nuclease domains to cleave opposite strands of the target DNA. The end result of Cas9-mediated DNA cleavage is a double strand break (DSB) within the target DNA (~3-4 nucleotides upstream of the PAM sequence).

The resulting DSB is then repaired by one of two general repair pathways:

1. The efficient but error-prone Non-Homologous End Joining (NHEJ) pathway
2. The less efficient but high-fidelity Homology Directed Repair (HDR) pathway

USING CAS9 TO GENERATE KNOCK-OUTS VIA NHEJ (CONT'D)

The NHEJ repair pathway is the most active repair mechanism, capable of rapidly repairing DSBs, but frequently results in small nucleotide insertions or deletions (Indels) at the DSB site. Here we will provide some of the details behind NHEJ.

Non-Homologous End Joining

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, homologue, or exogenously provided DNA) 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.

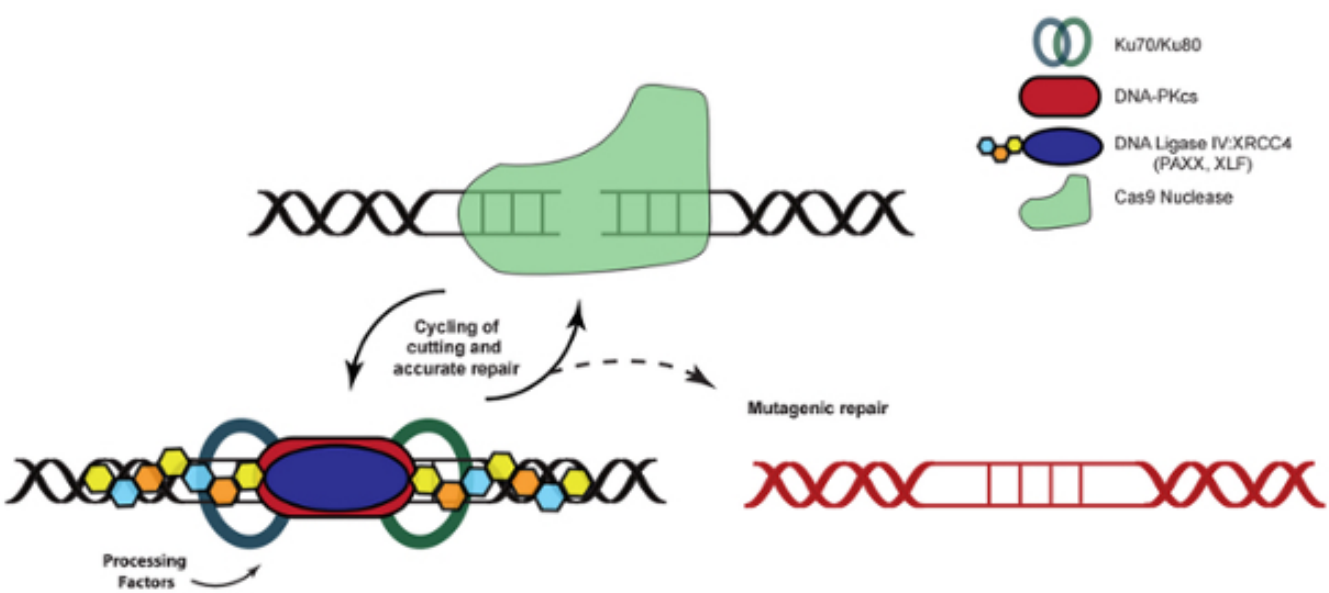


Image courtesy of D. Wyatt and D. Ramsden, UNC at Chapel Hill

USING CAS9 TO GENERATE KNOCK-OUTS VIA NHEJ (CONT'D)

Repair of Cas 9-Induced Breaks by NHEJ

As illustrated above, NHEJ-mediated repair of Cas9-generated breaks is useful if the intent is to make a null allele ("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 flanking sequence ("microhomologies").

It must be emphasized that 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.

NHEJ can also be engaged by variants of the canonical Cas9 approach. A pair of CRISPR guides that flank regions of hundreds of base pairs or more 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 direct insertion of 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, it's presumed that this results in a DSB with long overhangs. This "double nickase" strategy vastly reduces breaks and mutations at off-target sites, but it is not yet clear how NHEJ engages this class of breaks (see the section on nickase for more information).

USING CAS9 TO GENERATE KNOCK-OUTS VIA NHEJ (CONT'D)

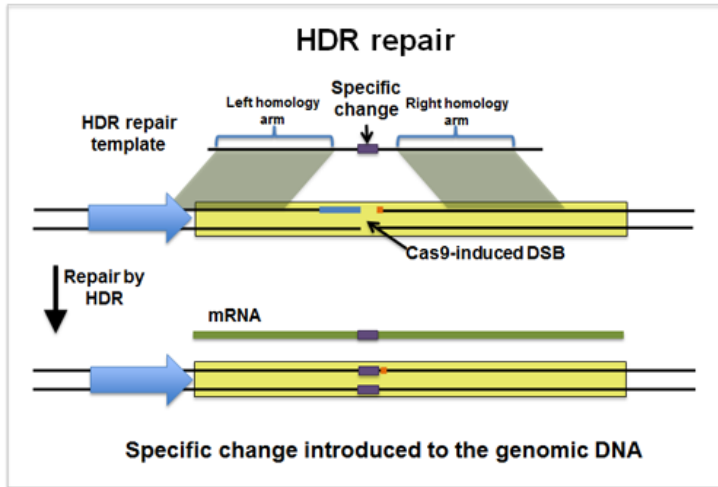
Further Reading

1. Bennardo, Nicole, et al. "Limiting the persistence of a chromosome break diminishes its mutagenic potential." PLoS Genet 5.10 (2009): e1000683. PubMed PMID: 19834534. PubMed Central PMCID: PMC2752804.
2. Bétermier, Mireille, et al. "Is non-homologous end-joining really an inherently error-prone process." PLoS Genet 10.1 (2014): e1004086. PubMed PMID: 24453986. PubMed Central PMCID: PMC3894167.
3. Waters, Crystal A., et al. "The fidelity of the ligation step determines how ends are resolved during non-homologous end joining." Nature Communications 5 (2014). PubMed PMID: 24989324. PubMed Central PMCID: PMC4107315.
4. Ghezraoui, Hind, et al. "Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining." Molecular Cell 55.6 (2014): 829-842. PubMed PMID: 25201414. PubMed Central PMCID: PMC4398060.
5. Ran, F. Ann, et al. "Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity." Cell 154.6 (2013): 1380-1389. PubMed PMID: 23992846. PubMed Central PMCID: PMC3856256.

EDITING WITH HOMOLOGY DIRECTED REPAIR

By Chari Cortez | March 12, 2015

DNA lesions are defined as sites of structural or base-pairing damage in DNA. Perhaps the most nocuous 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.



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

Mechanisms to Repair DNA Double-Strand Breaks

Genome stability necessitates 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 (discussed in the previous section) is the canonical homology-independent pathway as it involves the alignment of only one to a few complementary bases at most for the re-ligation of two ends, whereas HDR uses longer stretches of sequence homology to repair DNA lesions. This section focuses on HDR, which is considered to be 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 very specific mutations into the damaged DNA.

The three central steps of the HDR pathways are listed as follows:

1. 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.

2. The invasive strand can then displace one strand of the homologous DNA duplex and pair with the other; this results in the formation of the hybrid DNA referred to as the displacement loop (D loop).

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

The invasion of the 3' single-stranded DNA (ssDNA) into the homologous DNA duplex (Step 2) is the defining point of HDR. There are four different HDR pathways that can be employed to repair DSBs and the specific mechanisms used in Steps 2 and 3 define the individual pathways as described below.

Mechanisms to Repair DNA Double-Strand Breaks

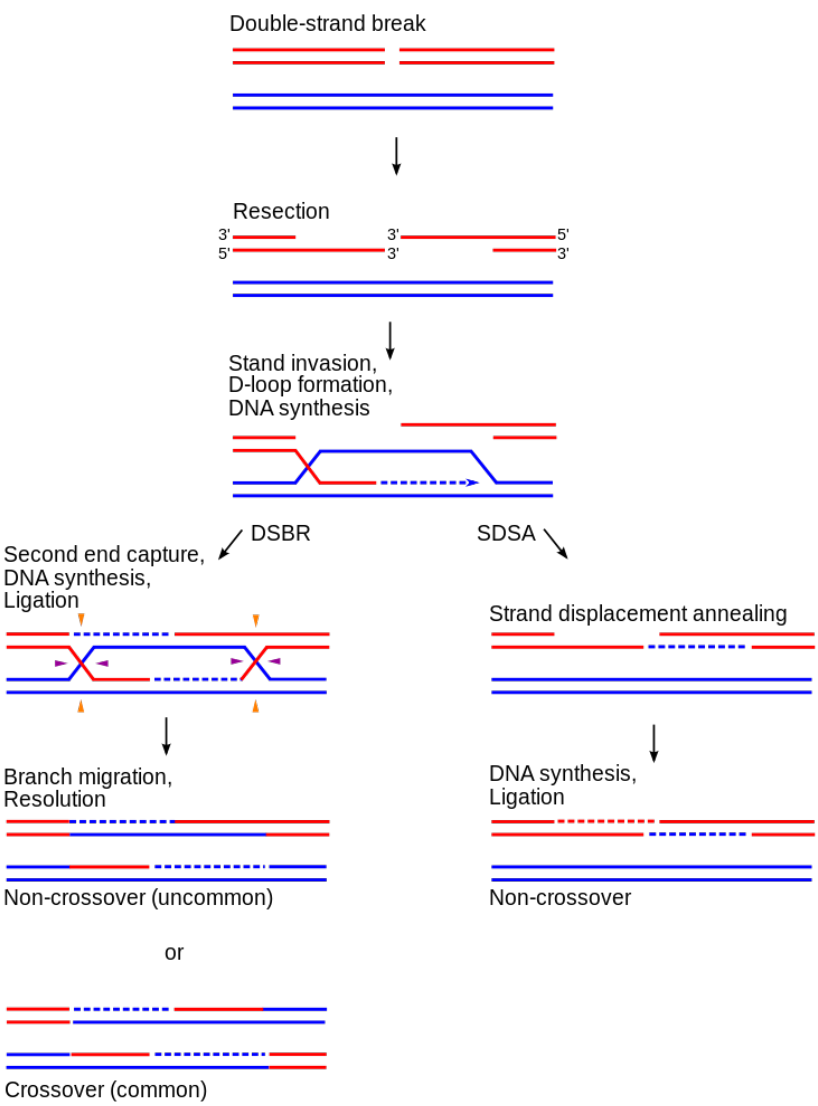
HDR can occur either non-conservatively or conservatively. The non-conservative method is composed

EDITING WITH HOMOLOGY DIRECTED REPAIR (CONT'D)

of the single-strand annealing (SSA) pathway and, in the interest of space, will not be discussed here. The conservative methods, characterized by the accurate repair of the DSB by means of a homologous donor (e.g., sister chromatid, plasmid, etc), is composed of three pathways: the classical double-strand break repair (DSBR), synthesis-dependent strand-annealing (SDSA), and break-induced repair (BIR).

Classical Double-Strand Break Repair (DSBR)

In the classical DSBR pathway, the 3' ends invade an intact homologous template, then serve as primers 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 below, 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.



"HR schematic diagram" by Emw2012. Licensed under CC-BY-SA-3.0 via Wikimedia Commons.

Synthesis-Dependent Strand-Annealing Pathway (SDSA)

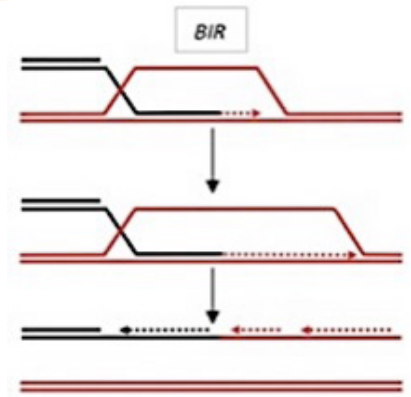
As illustrated on the right branch in the figure to the left, SDSA is conservative, and results exclusively in noncrossover events. This means all newly synthesized sequences are present on the same molecule. Unlike DSBR, following strand invasion and D loop formation in SDSA, the newly synthesized portion 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

Although BIR is not as well characterized as either DSBR or SDSA, 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

EDITING WITH HOMOLOGY DIRECTED REPAIR (CONT'D)

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.



Repair of DSBs and Genome Engineering

The advent of plasmid-based methods to induce DSBs engendered numerous technologies that integrated homologous recombination into genome engineering efforts. Early meganuclease-based technologies established the foundation for plasmid-based genome engineering tools such as Zinc Finger Nucleases (ZFNs) and TAL Effector Nucleases (TALENs). ZFNs and TALENs can both be used to direct an endonuclease to a specific DNA locus targeted for modification. On the surface, the discovery and development of RNA-guided CRISPR/Cas9 technology may just appear to be a new method for directing a nuclease to cause a specific DSB; however, the ease of creating guides, the speed of the system, and the overall versatility in application has not just reinvigorated genome engineering, but has 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; however, there may be some wiggle room. In general, the insertion sites of the modification should be no more than 100bp away from the DSB, ideally less than 10bp away if possible, and the overall length of the homology arm is an important factor to consider when designing these (more on this below). Longer distances of up to 200bp may work, but the efficiency will likely be lower and you may need to introduce a selection marker to ensure the modification is present.

One important point to note is that the CRISPR/Cas9 system does not stop once a DSB is introduced and repaired. As long as the gRNA target site/ PAM site remain intact, the Cas9 endonuclease will keep cutting and the DSB will keep getting repaired through either NHEJ or HDR. This could be problematic if you are trying to introduce a very specific mutation or sequence. To get around this, one may consider designing the HDR template in such a way that will ultimately block further Cas9 targeting after the initial DSB is repaired. For example, the PAM could be mutated such that it is no longer present, but the coding region of the gene is not affected (i.e. a silent mutation).

The efficiency of HDR is generally low (<10% of modified alleles) even in cells that express Cas9, gRNA, and an exogenous repair template. For this reason, many laboratories are attempting to artificially enhance HDR by synchronizing the cells within the cell cycle stage (S-phase) when HDR is most active, or by chemically or genetically inhibiting genes involved in NHEJ. The low efficiency of HDR has several important practical implications. First, since the efficiency of Cas9 cleavage is relatively high and the efficiency of HDR is relatively low, a portion of the Cas9-induced DSBs will be repaired via NHEJ. In other words, the resulting population of cells will contain some combination of wild-type alleles, NHEJ-repaired alleles, and/or the desired HDR-edited allele. Therefore, it is important to confirm the presence of the desired edit experimentally, and if necessary, isolate clones containing the desired edit (see our validation section in Plan Your Experiment).

EDITING WITH HOMOLOGY DIRECTED REPAIR (CONT'D)

What Makes the Best Template: Plasmid DNA or Single-stranded Donor Oligonucleotide (ssODN)?

The size of your intended mutation is a big factor in deciding on a single- or double-stranded DNA repair template. Historically, plasmids have been used as dsDNA templates when creating gene-targeting vectors; however, ssDNA templates (ssODNs) have come into common use for smaller modifications as they tend to have a higher efficiency. As a basic guideline, small mutations of up to ~50bp or single point mutations can successfully be introduced using ssODN templates, while dsDNA plasmid-based templates should be used for larger inserts such as fluorescent proteins or selection cassettes.

For ssODNs, the templates should be as long as possible with the Cas9-induced break point centered within the template. Scientists have been successful with template lengths of ~100-200bp in total, with at least 40bp (but usually closer to 50-80bp) homology arms on either side of your intended mutation. Because target sequence placement is PAM-dependent, it is not always possible to have the insertion site right next to the cut site; however, they should be reasonably close (within ~20bp) to each other.

For larger inserts, dsDNA encompassing homology arms of 800bp each or larger should be used. Plasmids are the most common source for providing dsDNA targets. This webpage is a great resource for designing a gene targeting vector.

Further Reading

1. FAQ on Addgene's website: https://www.addgene.org/crispr/zhang/faq/#HR
2. CRISPR Forum Moderated by the Zhang Lab https://groups.google.com/forum/#!searchin/crispr/HDR
3. Maria Jasin and Rodney Rothstein. "Repair of strand breaks by homologous recombination." Cold Spring Harbor Perspectives in Biology 5.11 (2013): a012740. PubMed PMID: 24097900. PubMed Central PMCID: PMC3809576.
4. Pardo B, Gomez-Gonzalez B, and Aguilera A. "DNA repair in mammalian cells: DNA double-strand break repair: how to fix a broken relationship." Cell Mol Life Sci Mar;66(6) (2009): 1039-56. PubMed PMID: 19153654.
5. Shrivastav, Meena, Leyma P. De Haro, and Jac A. Nickoloff. "Regulation of DNA double-strand break repair pathway choice." Cell Research 18.1 (2008): 134-147. PubMed PMID: 18157161.
6. Pfeiffer, Petra, Wolfgang Goedecke, and Günter Obe. "Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations." Mutagenesis 15.4 (2000): 289-302. PubMed PMID: 10887207.
7. Pâques, Frédéric, and James E. Haber. "Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae." Microbiology and Molecular Biology Reviews 63.2 (1999): 349-404. PubMed PMID: 10357855. PubMed Central PMCID: PMC98970.
8. Ran, F. Ann, et al. "Genome engineering using the CRISPR-Cas9 system." Nature Protocols 8.11 (2013): 2281-2308. PubMed PMID: 24157548. PubMed Central PMCID: PMC3969860.
9. Singh, Priti, John C. Schimenti, and Ewelina Bolcun-Filas. "A mouse geneticist's practical guide to CRISPR applications." Genetics 199.1 (2015): 1-15. PubMed PMID: 25271304. PubMed Central PMCID: PMC4286675.
10. Doudna JA and Sontheimer EJ. "The use of CRISPR/Cas9, ZFNs, and TALENs, in generating site-specific genome alterations." Methods in Enzymology 546. (2014). PubMed PMID: 25398356.
11. Li, Kai, et al. "Optimization of genome engineering approaches with the CRISPR/Cas9 system." PLoS One (2014): e105779. PubMed PMID: 25166277. PubMed Central PMCID: PMC4148324.
12. Peng, Ying, et al. "Making designer mutants in model organisms." Development 141.21 (2014): 4042-

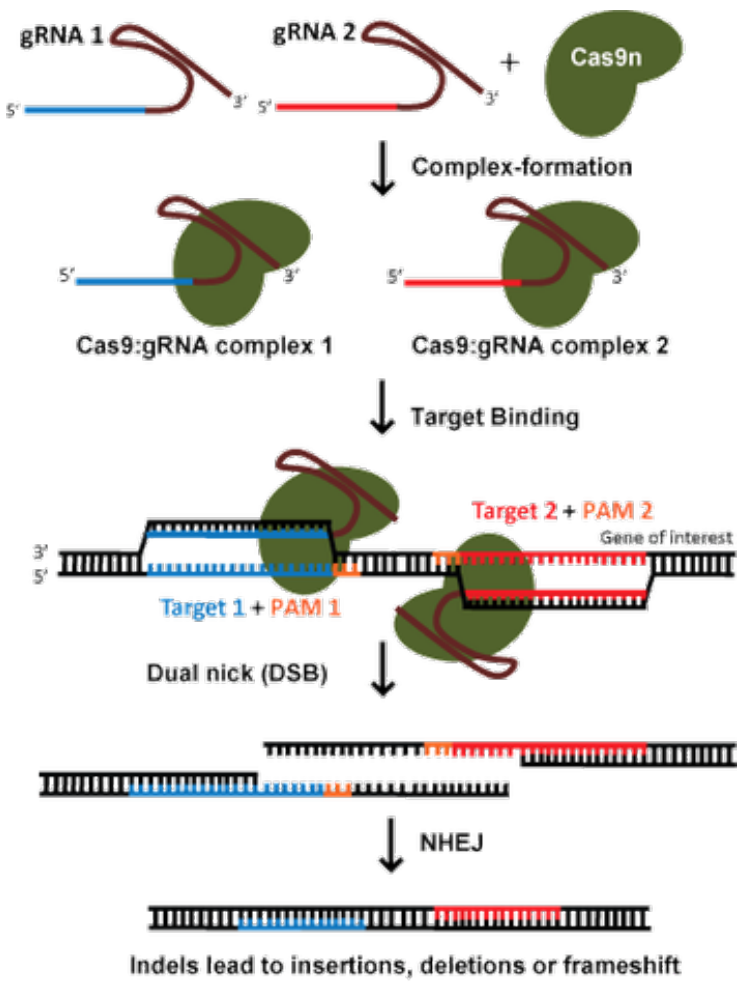
EDITING WITH HOMOLOGY DIRECTED REPAIR (CONT'D)

4054. PubMed PMID: 25336735. PubMed Central PMCID: PMC4302887.

USING CAS9 NICKASES FOR ENHANCED SPECIFICITY

By Joel McDade | October, 2015

CRISPR/Cas9 is highly specific when gRNAs are designed correctly, but specificity is still a major concern, particularly as CRISPR is being developed for clinical use. The specificity of the CRISPR system is determined in large part by how specific the gRNA targeting sequence is for the genomic target compared to the rest of the genome. Ideally, a gRNA targeting sequence will have perfect homology to the target DNA with no homology elsewhere in the genome. Realistically, a given gRNA targeting sequence will have additional sites throughout the genome where partial homology exists. These sites are called "off-targets" and need to be considered when designing a gRNA for your experiment. Learn how to design a gRNA in our "How to Design Your gRNA for CRISPR Genome Editing" section.



In addition to optimizing gRNA design, specificity of the CRISPR system can also be increased through modifications to Cas9 itself. As discussed previously, Cas9 generates double strand breaks (DSBs) through the combined activity of two nuclease domains, RuvC and HNH. The exact amino acid residues within each nuclease domain that are critical for endonuclease activity are known (D10A for RuvC and H840A for HNH in *S. pyogenes* Cas9) and modified versions of the Cas9 enzyme containing only one active catalytic domain (called "Cas9 nickase") have been generated. Cas9 nickases still bind DNA based on gRNA specificity, but nickases are only capable of cutting one of the DNA strands, resulting in a "nick", or single strand break, instead of a DSB. DNA nicks are rapidly repaired by HDR (homology directed repair) using the intact complementary DNA strand as the template (jump to our HDR section for more details). Thus, two nickases targeting opposite strands are required to generate a DSB within the target DNA (often referred to as a "double nick" or "dual nickase" CRISPR system). This requirement dramatically increases target specificity, since it is unlikely that two off-target nicks will be generated within close enough proximity to cause a DSB. Therefore, if specificity and reduced off-target effects are crucial, consider using the dual nickase approach to create a double nick-induced DSB. The nickase system can also be combined with HDR-mediated gene editing for highly specific gene edits.

Further Reading

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. Trevino, Alexandro E., and Feng Zhang. "Genome Editing Using Cas9 Nickases." Methods Enzymol. 546 (2014): 161-74. PubMed PMID: 25398340.
3. Shen, Bin, et al. "Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target

USING CAS9 NICKASES FOR ENHANCED SPECIFICITY (CONT'D)

effects." Nature Methods 11.4 (2014): 399-402. PubMed PMID: 24584192.

HOW TO DESIGN YOUR gRNA FOR CRISPR GENOME EDITING

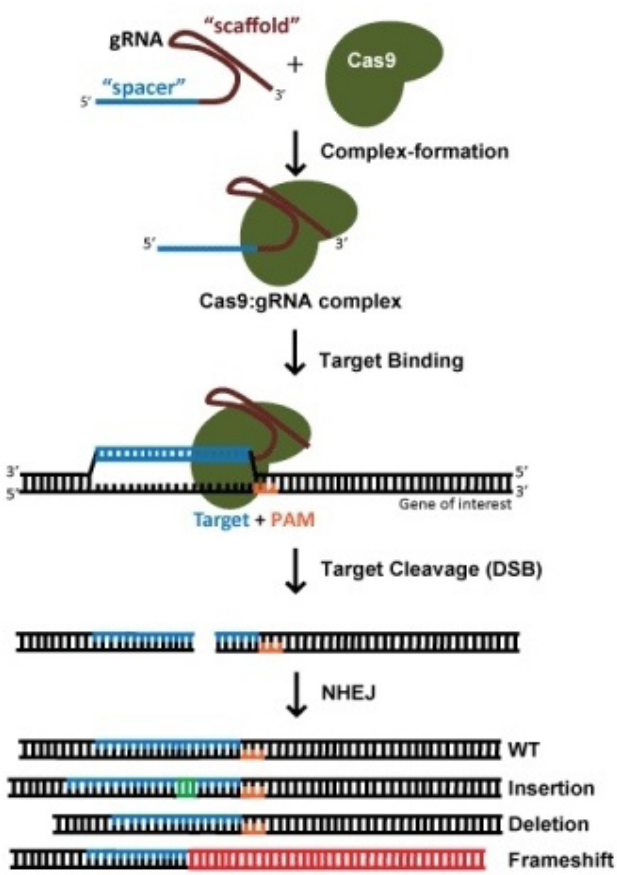
By John Doench | May 3, 2017

This post was contributed by guest blogger, Addgene Advisory Board member, and Associate Director of the Genetic Perturbation Platform at the Broad Institute, John Doench.

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 blog post will discuss differences between these approaches, as well as 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.

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 tying 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.



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, 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).

The Jigsaw: Editing by HDR

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

HOW TO DESIGN YOUR gRNA FOR CRISPR GENOME EDITING (CONT'D)

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.

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 (1). This means that, for gene editing, there are usually very few potential gRNAs. 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 (2). Thus, for gene editing, location is the most critical design parameter.

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 was recently shown that the FANTOM database, which relies on CAGE-seq to directly capture the mRNA cap, provides the most accurate mapping (3). In this case, location 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 On-Target Activity

Whether one's goal is gene disruption or gene editing, of one gene or genome-wide, being able to distinguish effective from ineffective gRNAs can greatly streamline an experiment and simplify interpretation of results. Previously, we had examined sequence features that enhance on-target activity of gRNAs by creating all possible gRNAs for a panel of genes and assessing, by flow cytometry, which sequences led to complete protein knockout (4). By examining the nucleotide features of the most-active gRNAs from a set of 1,841 gRNAs, we derived scoring rules and built a website implementation of these rules to design gRNAs against genes of interest. We then expanded our dataset and improved our computational modeling to derive Rule Set 2 for prediction of gRNA efficacy (5). We measured the activity of more than 2,000 additional gRNAs to further strengthen the statistical power, and confirm the generalizability, of activity predictions. In collaboration with Microsoft Research, we explored the use of more-powerful computational modeling approaches. While our initial model (Rule Set 1) was based on a

HOW TO DESIGN YOUR gRNA FOR CRISPR GENOME EDITING (CONT'D)

fairly simple classification model, we found that the use of regression models in general, and gradient-boosted regression trees in particular, greatly improved the power of our predictions. Web-based implementation of Rule Set 2 is now available from both the Broad and Microsoft and independent publications have shown its predictive value (6).

Decreasing Off-Target Effects

Avoiding off-target effects of Cas9, that is, cutting at other, unintended sites in the genome, is an important step in designing gRNAs. Merely glancing through the literature shows that different groups have come to wildly different conclusions as to the specificity of gRNAs. To take two examples, compare these titles:

"Low Incidence of Off-Target Mutations in Individual CRISPR-Cas9 and TALEN Targeted Human Stem Cell Clones Detected by Whole-Genome Sequencing" (7)

"High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells" (8)

It is reasonable to ask, well, which is right? As usual, the truth lies in the details, which is another way of saying that you can't judge a journal article by its title! Indeed, both titles are correct within the confines of each study, but the generalizability is what matters most. For sure, some differences in these reports (and many others) likely relate to differences in experimental systems, but probably most importantly, both of these papers examined small numbers of gRNAs. Are there some really promiscuous gRNAs? For sure! Are there quite specific ones? You bet! Of course, the same could be said for essentially any targeting technology – there are both really specific and really non-specific TALENs, siRNAs, antibodies, and small molecules.

Generalizability, then, needs to come from sampling from large numbers, and indeed, rules governing off-target effects are beginning to be understood in more detail. First, direct physical detection of off-target sites though techniques like GUIDE-Seq have shown that some gRNAs have dozens of detectable off-target sites, but that same study also found 1 gRNA, of 10 examined, that had zero off-target sites by their technique (9). Further, they showed that existing heuristics to find and score off-targets in fact miss many sites. They compared GUIDE-Seq results to two prediction algorithms from Feng Zhang's lab and Michael Boutros's lab and "discovered that neither program identified the vast majority of off-target sites found by GUIDE-seq." Of course, at the time of launch, these servers were based on the best-available information at the time, and the perfect should not be the enemy of the good.

More recently, we have examined off-target sites at much larger scale than previous studies and developed the CFD score to predict off-target sites with better sensitivity and specificity than previous heuristics (5). In the course of this study, we also found that the search algorithm itself plays a perhaps-under-appreciated role in arriving at the right result. Because of its ease of implementation and speed, many have used bowtie2 to perform scans of the genome to find off-target sites that contain small numbers of mismatches, but the bowtie algorithm was not designed for quite this purpose, and in fact misses many potential off-target sites, especially sites with more than 1 mismatch. Thus, both the search metric and the scoring metric are critical for

HOW TO DESIGN YOUR gRNA FOR CRISPR GENOME EDITING (CONT'D)

a comprehensive view of potential off-target sites (10).

For gene editing approaches, where the goal is to introduce a specific change at a specific site, the choice of gRNAs is often quite limiting and thus sometimes all your gRNAs will have poor off-target properties. One method to decrease off-target effects with CRISPR technology is the use of two gRNAs in combination with a mutated "nickase" version of Cas9. This approach has the benefit of increased specificity and thus a reduced rate of off-target dsDNA breaks. One downside of this approach, though, is that the requirement for two target sites will mean some specific locations are not suitable for creating a dsDNA break. When possible, though, this is the preferred approach for gene editing (learn more about nickase and specificity here). Another approach to decrease off-target effects is the use of Cas9 variants with engineered mutations that result in decreased binding energy between the protein, the RNA, and the DNA (11, 12). As a result, mismatched (i.e. off-target) sites can generally no longer serve as substrates for cutting.

Genome-Wide Pooled gRNA Libraries

We have implemented our on- and off-target scoring rules to create genome-wide pooled libraries. Our first attempts were named Avana (a grape used for making wine) for human and Asiago (a cheese) for mouse, and we compared performance to the GeCKO library, which was developed before these rules were available. For both positive and negative selection screens, we found that these new libraries were able to identify more hits with greater statistical confidence, due to the increased consistency of different gRNAs targeting a gene, that is, more of the gRNAs in the library were efficacious.

While it is of course true that more gRNAs per gene provide more information, this comes at the cost of screening and sequencing more cells, which puts some cellular models and experimental systems out of reach. Thus for many researchers, a primary screen that uses a smaller, high-activity genome-wide library will be desirable. Towards this end, we have made new libraries, named Brunello for human (again, a wine-making grape… you can see where we're going with this) and Brie for mouse, that take into account both our newest on-target designs and avoidance of off-target sites. These libraries are available from Addgene as both plasmid pools and ready-to-use lentivirus. Our (as-yet-unpublished... but also not-yet-rejected!) data show that the improvement from Avana to Brunello is approximately equal to the improvement we saw in going from GeCKOv2 to Avana. By one analysis approach, we see that the use of just a single gRNA in the Brunello library outperforms the use of all 6 gRNAs in the GeCKOv2 library. Additionally, we have designed libraries for CRISPRa (Calabrese and Caprano) and CRISPRi (Dolcetto and Dolomiti) using optimized design rules, which will soon be available via Addgene as well. A publication describing these libraries is likewise working its way through The System.

Delivery Options

Once a target site has been identified, it is important to consider delivery options. For conducting genetic screens in pooled format, the use of an integrating virus (e.g. lentivirus) is critical to the entire process. However, for generating a cellular model, long-term expression of CRISPR components is not desirable, due to the potential for accumulation of off-target lesions. Transient expression options are the most appropriate choices for the creation of a stable cell line. These can include the transfection or electroporation of plasmid DNA, mRNA, or Cas9 protein pre-complexed to in vitro transcribed or synthesized gRNA, or the use of non-integrating viruses such as AAV or Adenovirus.

If performing HDR, the repair template can be either a long, dsDNA (e.g. a plasmid) or a single-stranded oligo-

HOW TO DESIGN YOUR gRNA FOR CRISPR GENOME EDITING (CONT'D)

nucleotide co-delivered with the Cas9 and gRNA. The choice between the two templates is largely dictated by the size of the intended change; small (< ~40 nt) changes can be introduced with synthesized oligonucleotides of ~100 – 200 nts in length. These can simply be purchased commercially. Large inserts, such as the introduction of GFP to tag a protein, require a template with much larger homology arms.

Conclusions

In sum, 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 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 as a hit.

This is 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.

Further Reading

1. Yang, L. et al. Optimization of scarless human stem cell genome editing. Nucleic Acids Research 41, 9049–9061 (2013). PubMed PMID: 23907390. PubMed Central PMCID: PMC3799423.
2. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 1–17 (2016). PubMed PMID: 27096365. PubMed Central PMCID: PMC4873371.
3. Radzisheuskaya, A., Shlyueva, D., Müller, I. & Helin, K. Optimizing sgRNA position markedly improves the efficiency of CRISPR/dCas9-mediated transcriptional repression. Nucleic Acids Research 44, e141–e141 (2016). PubMed PMID: 27353328. PubMed Central PMCID: PMC5062975.
4. Doench, John G., et al. "Rational design of highly active gRNAs for CRISPR-Cas9-mediated gene inactivation." Nature biotechnology (2014). PubMed PMID: 25184501. PubMed Central PMCID: PMC4262738.
5. Doench, J.G. et al. Optimized gRNA design to maximize activity and minimize off-target effects for genetic screens with CRISPR-Cas9. Nat Biotechnol. PubMed PMID: 26780180. PubMed Central PMCID: PMC4744125.
6. Haeussler, M. et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol 17, 1–12 (2016). PubMed PMID: 27380939. PubMed Central PMCID: PMC4934014.
7. Veres, Adrian, et al. "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 15.1 (2014): 27-30. PubMed PMID: 24996167. PubMed Central PMCID: PMC4082799.
8. 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.
9. 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.

HOW TO DESIGN YOUR gRNA FOR CRISPR GENOME EDITING (CONT'D)

10. Bae, S., Park, J., Park, J. & Kim, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30, 1473–1475 (2014). PubMed PMID: 24463181. PubMed Central PMCID: PMC4016707.
11. Slaymaker, I. M. et al. Rationally engineered Cas9 nucleases with improved specificity. Science 351, 84–88 (2016). PubMed PMID: 26628643. PubMed Central PMCID: PMC4714946.
12. Kleinstiver, B. P. et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529, 490–495 (2016). PubMed PMID: 26735016. PubMed Central PMCID: PMC4851738.

THE CRISPR SOFTWARE MATCHMAKER

By Cameron MacPherson | November 3, 2015

This section was contributed by guest writer Cameron MacPherson at the Institut Pasteur.

CRISPR Software and the Piñata Effect

Two years ago I was a part of a group (Biology of Host-parasite Interactions, Institut Pasteur, Paris) that changed genome editing in the malaria community for the better (1). Given the timing, it shouldn't be a surprise that the CRISPR system was involved. Today, that same laboratory enjoys a successful edit rate of over 90% in their work editing the genome of *Plasmodium falciparum* (the parasite that causes malaria). I attribute their success to technical expertise, thoughtful gRNA design, and the abnormally low GC content of the *Plasmodium falciparum* genome. To put this last point into perspective, the *Plasmodium falciparum* genome contains only 0.66 million targetable NGG PAM sites whereas the human genome has about 300 million. With such a sparsely targetable genome, off-targeting is less of a worry and on-targeting likely more efficient. These insights are hard to appreciate without computational support. Indeed, rational gRNA design is not possible without relying on some kind of pre-analysis. At the end of 2014, I began developing software to make gRNA design accessible to all. At the time I thought there was room for improvement and a year later it became quite clear that others thought the same. Since January 2013 there have been 33 CRISPR software tools published and documented by OMICtools. It must be confusing for a newcomer to decide on the right software to use. With so much choice the question I have is, how did we get here in the first place?

How CRISPR Innovation Affected Software Development

CRISPR, like the miRNA bubble before it, offers an interesting view into how rapid innovation affects the research community. I've likened it to the celebration surrounding the piñata. From the perspective of those building the piñata, their role is to fill the prop with something of value. They are reviewers of content. On the other hand, the goals of the piñata beating party goers are altogether different. They are filled with blind anticipation. Their trust in the quality of the content is implicitly defined by their trust in the piñata builders - the reviewers. After some physical exertion on the part of the party goers, the piñata bursts and its contents spill to the ground. There is something akin to a frenzy as the crowd tries to weigh and determine the value of every sweet or toy strewn across the floor. In the final wake it all settles down and the value of each item changes from what the reviewers so painstakingly assessed to something a bit more representative of the communities' opinion. The trouble is, with so much choice, in those few moments after the piñata burst, value was subjectively ascribed relative to the basic needs of each individual; the community as a whole has yet to settle on any true kind of value. I've dubbed this the "Piñata Effect", it is the disconnect arising between those reviewing the contents of the piñata and the audience blindly receiving them; it results in highly similar content and the absence of iterative design; it is caused by not having the time to assess public response; and, it is the stage I think we are currently at in the CRISPR software space. The CRISPR research community hasn't had a chance to develop a consensus on the best CRISPR software tools yet.

THE CRISPR SOFTWARE MATCHMAKER (CONT'D)

As a genome editing tool, the CRISPR/Cas9 technology has been surrounded by a whirlwind of research activity and development since its inaugural year in 2012. That's just 3 to 4 years ago. It is a bubble of innovation (a piñata), and software development has caught up with it. The peer review process has been challenged with many closely spaced CRISPR software submissions. With few prior publications to go on, a submitted manuscript could only have been viewed as a significant improvement. As a result, the software piñata has been filled with slight variations on a theme without many resources to review them. Collectively, the pool of CRISPR software embodies solutions that facilitate most experimental applications in CRISPR engineering. When compared to that collective utility, it's easy to fault the lack of features in a single application. The question is then, should we use all tools as some kind of a meta/Frankenstein app? Or perhaps there is a clear winner, is there something that we can invest our time into learning and extract the most value? This is only one of many decisions faced by those engaging in CRISPR design, but it is a significant hurdle made worse by an ever increasing marketplace. Since 2013 about 11 tools have been added every year.

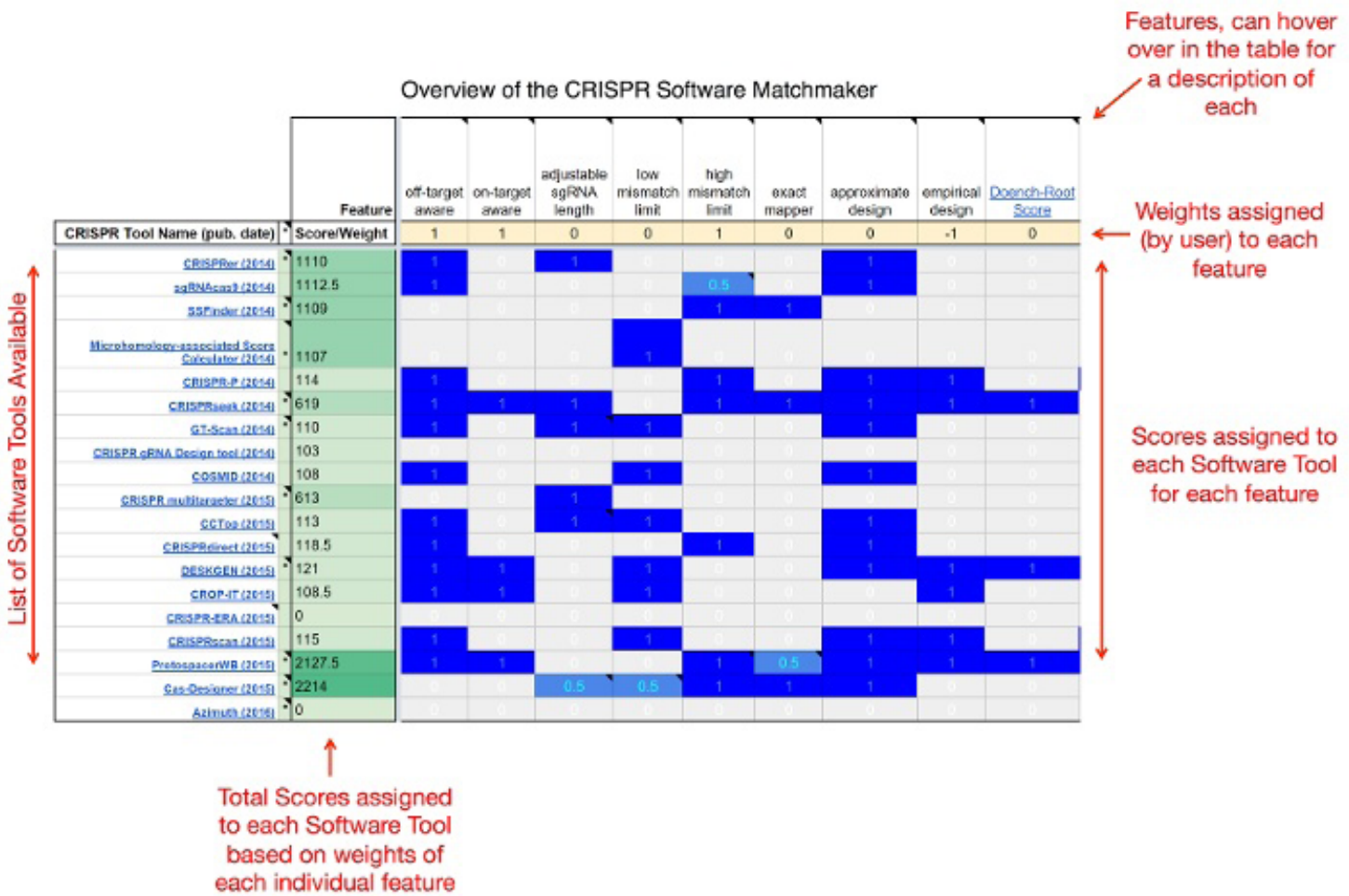
Breaking Down the Barriers: My View on Current gRNA Design Tools

The goal of this post is to provide some insight into available CRISPR software tools, what problems various tools are trying to solve, and finally how we might proceed in the future. When we think of designing experiments using CRISPR, there are two major areas where software can help. The foremost handles the design of gRNA and represents the lion's share of the currently available tools (we will focus on these tools in this section). The second major area deals in post-experiment quality control, a good example is CRISPR-GA. These tools assess repair events by type (NHEJ/HDR) and track indels at a single locus, usually the target site. Multi-locus assessment would be required to fully evaluate off-targeting, but currently no software has been designed for this purpose. This second area is extremely important and given the lack of attention to it, it is an obvious place to focus further development. Quality control is, however, not a focus of this section.

I separate gRNA design tools into database and *de novo* solutions. Database tools allow us to view and get a sense of which gRNA designs have previously worked and under what conditions. Such resources could be highly valuable for automated rational design. The first of only three such databases, EENdb (published 1 Jan 2013), is a simple catalogue of reported gRNA designs. The second database, CrisprGE (published 27 Jun 2015), is noteworthy for its broad scope, curated content, and ease of access. The third, WGE (published 15 Sep 2015), is more similar to EENdb than CrisprGE, but is also a good example of how database tools can be used to aid design. These databases are still young and will require a collaborative effort from the community in order to succeed.

With respect to *de novo* gRNA design, I again separate it into two categories. On the one hand we have software tools that apply some approximate or sensible rules to determine the value of one gRNA design over another. On the other hand some software tools make use of empirically derived and/or prior knowledge to inform on the qualities of what determines good gRNA design. Incidentally, the first software for gRNA design (by Hsu et al., 2013, [2]) used a combination of both empirical and approximate methods. The details of their scoring function can be found in their online tool's documentation. Since Hsu et al., 2013 (2), other tools have incorporated new parameters such as the Doench-Root score (by Doench et al., 2014, [3]) into their scoring functions. Many tools have also opted to mix-and-match different scoring algorithms and parameters. This tapestry of ideas is what makes choosing the right tool confusing and it means that there is no one-size-fits-all approach. You will have to select a tool based on your project and not general opinion.

THE CRISPR SOFTWARE MATCHMAKER (CONT'D)



CLICK HERE TO USE THE CRISPR SOFTWARE MATCHMAKER

I could spend a lot of time running through each tool in detail, but that is already covered in the published articles. Instead, I have broken down the tools into the individual features that make each tool unique. The CRISPR Software Matchmaker is composed of these features and enables you to select the tool(s) based on your project needs. There are 8 major categories found in the table, describing everything from basic to advanced functionality as well as how the user is expected to interact with the tool (a screen shot and overview of the table can be found above). These categories are further discussed below (definitions for all terms are also available in the table):

- **Basic functions**: These functions embody the main goals of the software and should be the first place you look to determine a tool's suitability. Functions of this category are the most common between tools. Examples: "single-target design", "multi-target design", "off-target aware", "high mismatch limit", "approximate design", "empirical design", "single-PAM design", and "multi-PAM design". Trends: By the end of 2014, tools began moving away from designing exclusively for NGG PAM targets and began allowing for arbitrary PAM definitions. These newer tools may still prove somewhat useful for one of the latest NTT targeting Cas protein, Cpf1. Separately, on-target efficacy is still a concept that only a few

THE CRISPR SOFTWARE MATCHMAKER (CONT'D)

tools are trying to solve.

• **Advanced functions**: These functions are not entirely necessary but should be considered extremely useful, depending on design goals. Examples: "feature aware", "SNP aware", "secondary structure aware", and "microhomology aware". Trends: Fewer tools are being released with advanced functionality, it seems these kinds of features are delegated to other more suitable software tools such as ApE or commercial workbenches such as MacVector or Benchling.

• **Utility functions**: These functions help speed up the gRNA design process by removing repetitive tasks and/or by providing features to help with the post-design process such as primer and plasmid design. Examples: "multiplex design", "multi-method design", and "single-method design". Trends: The most common utility functions are batch design or multiplex features. However, tools aiding in primer design are more and more common.

• **User interaction**: Software design elements that fall into this category describe how the user is expected to interact with the software. This category is important for users wishing to select tools based on their comfort level with operating a computer. Examples: "offline", "online", "CLI", and "GUI". Trends: Online tools dominate the software space but generally rely on less powerful algorithms to detect off-targets. The few offline tools available mostly work on any computer but generally require working knowledge of the command line or scripting.

• **Input flexibility**: Software tools require some kind of data input in order to generate results. While different types of input are more user-friendly than others, this is of less concern than data output as different input types can be readily converted. Examples: "organism", "sequence", "identifier", and "load". Trends: Sequence based input is by far the most common with most tools also requiring that the user specify an organism.

• **Output diversity**: Different tools provide results to the user in different formats. This can have a large impact on downstream results. Examples: "HTML", "visual track", "plot", "tabular", "interactive", and "save". Trends: Most tools provide tabular HTML formatted output. Surprisingly, only 2 tools provide FASTA output. A few notable tools allow the user to save the results and return later or share them.

• **Community exclusivity**: Different communities develop software for their own needs. This often results in software design that is applicable to only one, or a subset of organisms. In only a few cases the software is organism agnostic. Most tools are exclusive to one or a subset of organisms. Examples: "organism agnostic", "model organisms", "small genome", "large genome", and "organism biased scoring." Trends: Many tools can handle genomes of any size and are only limited by the time it takes authors to add a genome to the tool's repertoire. Organism agnostic tools are the better solution but are often packaged as offline tools requiring expertise in one programming language or another.

• **Supported organisms**: This section lists the genomes supported by each tool. The genomes are denoted by organism name and genome assembly version. The names in the table appear as they do in the software. Most gRNA design tools require the user to specify an organism. They do this because they rely on pre built indices or databases in order to find and present the results to you as soon as possible. Some tools don't require the organism to be specified and are truly organism agnostic. Other tools require an organism to be specified but also allow you to build your own database; these tools will be marked as organism agnostic and have currently available organisms listed in this section.

THE CRISPR SOFTWARE MATCHMAKER (CONT'D)

CRISPR Software Advice for Specific Users

You will find that some tools cater to specific use cases through the way their algorithms were developed or by their focus on specific organisms. For this reason I generally recommend that laboratories that have already attempted several CRISPR experiments evaluate the predictive power of each tool they are interested in by comparing their past results to the designs suggested by each tool. For those who are entering gRNA design for the first time, it is best to choose a tool based on immediate needs and re-evaluate after several experiments. Alternatively, the databases mentioned above or, even better, first-hand knowledge from laboratories employing CRISPR on the same organism could be used as a proxy for this last approach.

- **Top tool for first timers**: You need to learn the language and what the parameters represent. The best learn-by-doing tool is E-CRISP. The authors have made a huge effort so that their tool is both didactic and functional. It is also the only tool I have come across to offer classes of parameters based on the type of CRISPR experiment being pursued. You should also look at the glossary section of the CRISPR Software Matchmaker, it provides a list of terms I think are important to define for the CRISPR software space. It is unlikely that you will find these terms anywhere else, because I developed them for this post, but they should give you an idea of what to look out for.

- **Top tools for bioinformaticians**: For flexibility in your own analysis, you need access to raw data, the target sites, off-target sites, scores and the statistics that go into calculating them. For this, some online tools such as Cas-OFFinder will suffice. Cas-OFFinder provides a bare-bones data dump of ALL strings in a genome using any IUPAC encoded pattern at any edit distance. It won't, however, compute any score other than the edit distance. Added functionality is left to you to implement. You can also download the software for use on any machine with OpenCL enabled hardware. This OpenCL dependency is a severe limitation, but the tool is fast enough that, if you're engaged in heavy CRISPR design (I'm talking screens and/or very large genomes), then it is worthwhile buying a dedicated machine. On the offline front, there is a Python script called SSFinder, but I find its implementation to be too slow for practical use. The R based package CRISPRseek offers great utility and coupled with Bioconductor should be the first choice for anyone already familiar with R. For the Java enthusiasts, take a look at sgRNAcas9.

- **Top tools for transferring CRISPR technology to a new organism**: For obvious reasons you can rule out any tool that limits you to a subset of specific genomes. These are most of the tools, but the good news is that you only need one. ProtospacerWB was developed with the purpose of applying CRISPR technologies to new organisms and will help you whether you have a full assembly or not. It is an offline tool, but comes with a graphical user interface.

- **Best strategy for labs**: Define what you need the software tool to do before going shopping for one. Use the CRISPR Software Matchmaker to select the best tool based on your needs. Refine your criteria. Repeat until you have found the best tool. Do several experiments and use the results to re-evaluate all tools. Report your findings to help everyone else.

CRISPR technology has reached into so many different communities that it is all but impossible to fairly judge individual tools. I believe at this time it is only possible to objectively break the tools into their individual offerings and allow you to select them based on your needs. We will likely begin to see a contraction of available tools with more and more features being integrated into so-called genome editing workbenches such as Benchling (commercial, online), DESKGEN (commercial, online), or ProtospacerWB (academic, offline). While it remains to be seen, I believe the future of CRISPR software is promising. Rapid innovation in this space has

THE CRISPR SOFTWARE MATCHMAKER (CONT'D)

put us in a good position to step back and cherry pick the features that will actually make a difference in our experiments and lives. To do this effectively, there needs to be good communication between developers and end users.

Further Reading

1. Ghorbal, Mehdi, et al. "Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system." Nature Biotechnology 32.8 (2014): 819-821. PubMed PMID: 24880488.
2. Hsu, Patrick D., et al. "DNA targeting specificity of RNA-guided Cas9 nucleases." Nature Biotechnology 31.9 (2013): 827-832. PubMed PMID: 23873081. PubMed Central PMCID: PMC3969858.
3. Doench, John G., et al. "Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation." Nature Biotechnology 32.12 (2014): 1262-1267. PubMed PMID: 25184501. PubMed Central PMCID: PMC4262738.

GENOME-WIDE SCREENING USING CRISPR/CAS9

By Joel McDade | August 18, 2015

What genes are important for your phenotype? Many scientists (possibly you!) 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 and could aid in the development of novel therapies. While CRISPR/Cas9 is certainly not the first means to carry out these so-called "forward genetic screening experiments", it is certainly the most robust. In this blog post, we will discuss how CRISPR libraries are being used to perform genome-wide screens and highlight some of the reagents that have been made available through Addgene.

What Makes CRISPR So Special?

A major advantage of CRISPR/Cas9 over previous genome editing techniques is its simplicity and versatility. The CRISPR/Cas9 system 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 user defined, and 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), can result in a loss-of-function mutation within the target gene. CRISPR/Cas9 can also be used to activate or repress target genes without permanently modifying the genome by using nuclease dead Cas9 (dCas9) and dCas9-activators/repressors.

Figure 1: Like this jar of jellybeans, a pooled CRISPR library is a complex mixture. Pooled libraries are composed of many different gRNA-containing plasmids, each with a different genomic target. The goal of a CRISPR screen is to use Cas9 and a pool of gRNAs to identify genes that are essential for a given phenotype. Image Credit: https://www.flickr.com/photos/72005145@N00/5600978712.

What CRISPR 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/Cas9 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/Cas9 genome wide-screening experiments commonly use lentivirus to deliver a pooled population of gRNAs to Cas9 expressing cells. Pooled lentiviral CRISPR libraries (heretofore 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 (see Figure 1). Individual gRNAs

GENOME-WIDE SCREENING USING CRISPR/CAS9 (CONT'D)

are designed *in-silico* using publicly available gRNA design software and synthesized. Pooled gRNAs are then cloned into a lentiviral transfer vector, resulting in the final CRISPR library. CRISPR libraries have been adapted to knock-out, activate or repress target genes by combining a gRNA library with the aforementioned derivatives of Cas9. Several CRISPR libraries are now available through Addgene.

Choosing the Library that Is Right for You

There are several factors to consider when selecting a CRISPR library for your experiments, such as: 1) what species are your cells derived from? Currently, Addgene carries CRISPR libraries that target mouse, human, fly, and *T. gondii* genes. 2) What is the genetic modification you are trying to make? Addgene carries CRISPR libraries for gene knock-out, activation and repression. 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. A well-developed biological question is absolutely necessary to ensure that you select the correct CRISPR library for your experiment.

What Are the Steps Involved in a CRISPR Screen?

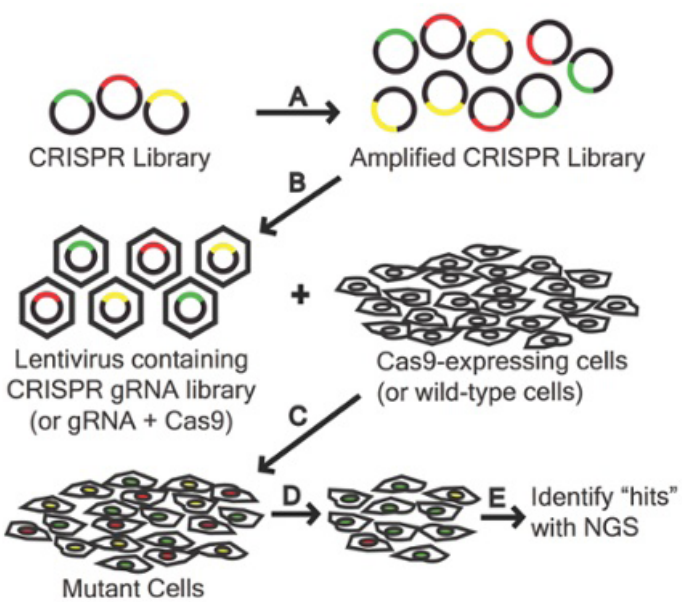


Figure 2: The CRISPR library must be amplified (A) to be used to generate lentivirus (B). Cas9-expressing cells or wild-type cells are treated with lentivirus containing the gRNA library or gRNA library plus Cas9 to generate mutant cells, respectively (C). Mutant cells are screened (D) and "hits" are identified using next-generation sequencing (E).

Performing a forward genetic screen using CRISPR libraries is a multi-step process (see figure at left). 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 be used to generate lentivirus. Be sure to check the representation of your library (percentage of each gRNA pre- and post-amplification) using next-generation sequencing. Cas9-expressing cells are then transduced with lentivirus containing the CRISPR library to generate a heterogeneous population of mutant cells, with each cell or sets of cells containing a mutation in a different gene. 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. Here, mutant cells are treated with 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.

GENOME-WIDE SCREENING USING CRISPR/CAS9 (CONT'D)

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 gRNAs per target gene, and maintaining the distribution of each gRNA within the population ("representation") is absolutely essential. 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. Furthermore, 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.

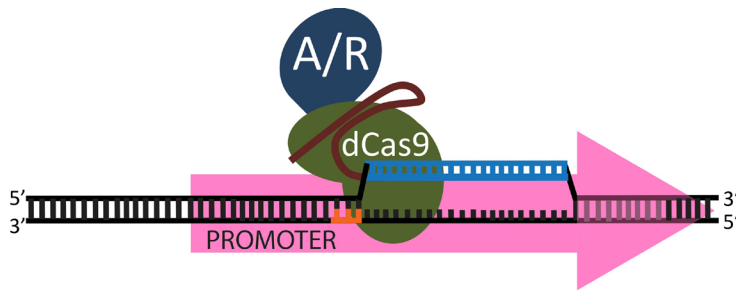
REGULATING GENE EXPRESSION WITH CRISPR

By Tyler J. Ford and Joel McDade | January, 2016

In previous sections we've highlighted how CRISPR has made it easy for researchers to delete genes from or edit the genome, but a variety of scientists have cleverly fused Cas9 to protein domains that allow them to do much more than alter DNA sequence. In this section we'll cover how researchers have engineered CRISPR to both repress and activate gene expression. See later sections for more great CRISPR applications!

Enabling New Functions with dCas9

The CRISPR/Cas system is a remarkably flexible tool for genome manipulation. A unique feature of Cas9 is its ability to bind target DNA independently of its ability to cleave target DNA. Specifically, both RuvC- and HNH- nuclease domains can be rendered inactive by point mutations (D10A and H840A in SpCas9), resulting in a nuclease dead Cas9 (dCas9) molecule that cannot cleave target DNA. The dCas9 molecule retains the ability to bind to target DNA based on the gRNA targeting sequence. The first experiments using dCas9 in bacteria demonstrated that targeting dCas9 to transcriptional start sites was sufficient to "repress" or "knock-down" transcription by blocking transcription initiation (figure to the right) (1).



Cas9 can be fused to an activator (A) or repressor (R) to alter expression of a gene whose promoter is targeted by the gRNA.

Qi et al. showed that targeting dCas9 to transcription factor binding sites in the promoter region of a gene or to regions downstream of the transcription start site resulted in robust repression in *E. coli* with few off-target effects. This is only effective, however, if dCas9 is targeted to the non-template strand and appears to function by dCas9 physically blocking RNA polymerase. Qi et al. further showed that dCas9 alone can modestly repress transcription in mammalian cells. Robust repression in mammalian cells and dCas9-based gene activation requires further engineering.

Further Engineering dCas9 for Gene Repression and Activation

Although dCas9 alone can moderately repress gene expression in mammalian cells, these repressive effects can be enhanced by fusing dCas9 to the KRAB repressor domain. Gilbert et al. showed that, when targeted to promoter regions, dCas9-KRAB fusions effectively repress expression of a variety of genes in mammalian and yeast cells and have few off target effects. Notably however, not all promoter-targeting gRNAs facilitate robust repression; the rules governing dCas9-based repression in mammalian cells are not well understood yet and any studies designed using these systems should use multiple gRNAs (2).

Unlike repression, dCas9-based transcription activation requires a dCas9 fusion. In *E. coli* this has been accomplished by fusing dCas9 to the omega subunit of RNAP (3). In mammalian cells, a variety of transcriptional activators have been fused to dCas9 including VP64 and p65. Similarly to the repressors, the level of activation using these fusions is dependent upon the location targeted as well as the specific gRNA used and the rules for effective activation are not yet well understood. It is again recommended that researchers test multiple gRNAs when attempting to activate or repress gene expression using one of these dCas9 fusions.

REGULATING GENE EXPRESSION WITH CRISPR (CONT'D)

Recruiting Multiple Transcriptional Activators Simultaneously

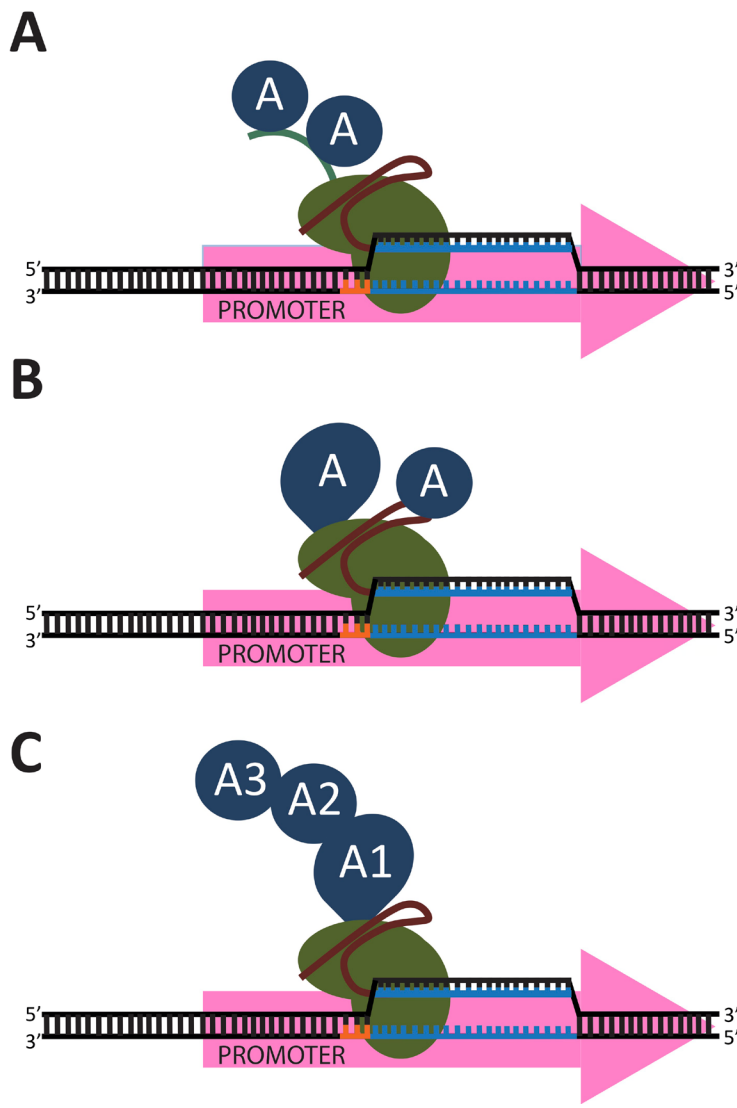
Scientists have recently developed a second generation of activators that alter gene expression by combining the effects of multiple transcriptional activators. For instance, the "SunTag" system recruits multiple activation domains to the same genetic locus through co-expression of epitope tagged dCas9 and antibody-activator fusion proteins. Here, the antibody fusion proteins bind to the epitope-tagged Cas9 and activate gene expression at the promoter targeted by a co-expressed gRNA (4). In a similar vein, the Synergistic Activation Mediator (SAM) complex consists of a dCas9-VP64 fusion and a modified gRNA that is capable of interacting with an additional RNA-binding transcriptional activator. Here the modified gRNA has a dual role - it both directs the entire complex to the promoter of interest and recruits a transcriptional activator resulting in greater activation of gene expression than with dCas9-VP64 alone (5). The Zhang Lab has developed a SAM-based pooled library that can be used to screen the effects of activating gene expression genome-wide. Finally, it's possible to simply express Cas9 fused to multiple transcriptional activators as in dCas9-VPR from the Church lab (6).

These various methods of altering gene expression should make it easier than ever for researchers to study the function of their gene of choice without knocking it out and should be especially useful to scientists studying genes with essential functions.

You can find many plasmids with dCas9 based transcription activators and repressors on the Addgene website:

Find Activators Here

Find Repressors Here



Some possible ways to use dCas9 to recruit multiple transcriptional activators domains to a promoter of interest. A) dCas9 is fused to a series of epitope tags each of which can recruit an activator fusion protein. B) Cas9 is directly fused to a transcriptional activation domain while a modified gRNA recruits an additional transcriptional activator. C) Cas9 itself is directly fused to multiple activators.

Further Reading

1. 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.
2. 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.

REGULATING GENE EXPRESSION WITH CRISPR (CONT'D)

3. Bikard, David, et al. "Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system." Nucleic Acids Research 41.15 (2013): 7429-7437. PubMed PMID: 23761437. PubMed Central PMCID: PMC3753641.

4. 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.

5. Konermann, Silvana, et al. "Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex." Nature (2014). PubMed PMID: 25494202. PubMed Central PMCID: PMC4420636.

6. Chavez, Alejandro, et al. "Highly efficient Cas9-mediated transcriptional programming." Nature Methods 12.4 (2015): 326-328. PubMed PMID: 25730490. PubMed Central PMCID: PMC4393883.

CAS9 ACTIVATORS: A PRACTICAL GUIDE

By Marcelle Tuttle and Alex Chavez | August 8th, 2016

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).

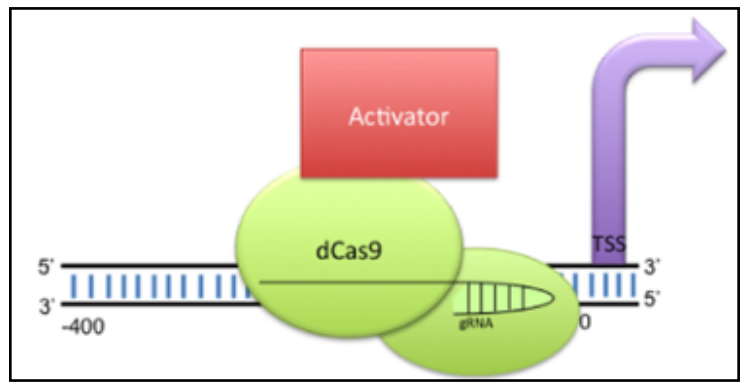


Figure 1: Generic activation experiment. dCas9 binds upstream of the promoter of a gene and drives transcription by either having an activator fused to it or by recruitment of an activator. In order to achieve maximal activation, we recommend designing your gRNA so that it targets a region 0 - 400 bp upstream of the transcription start site.

When to Use and When Not to Use Cas9 Activators

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

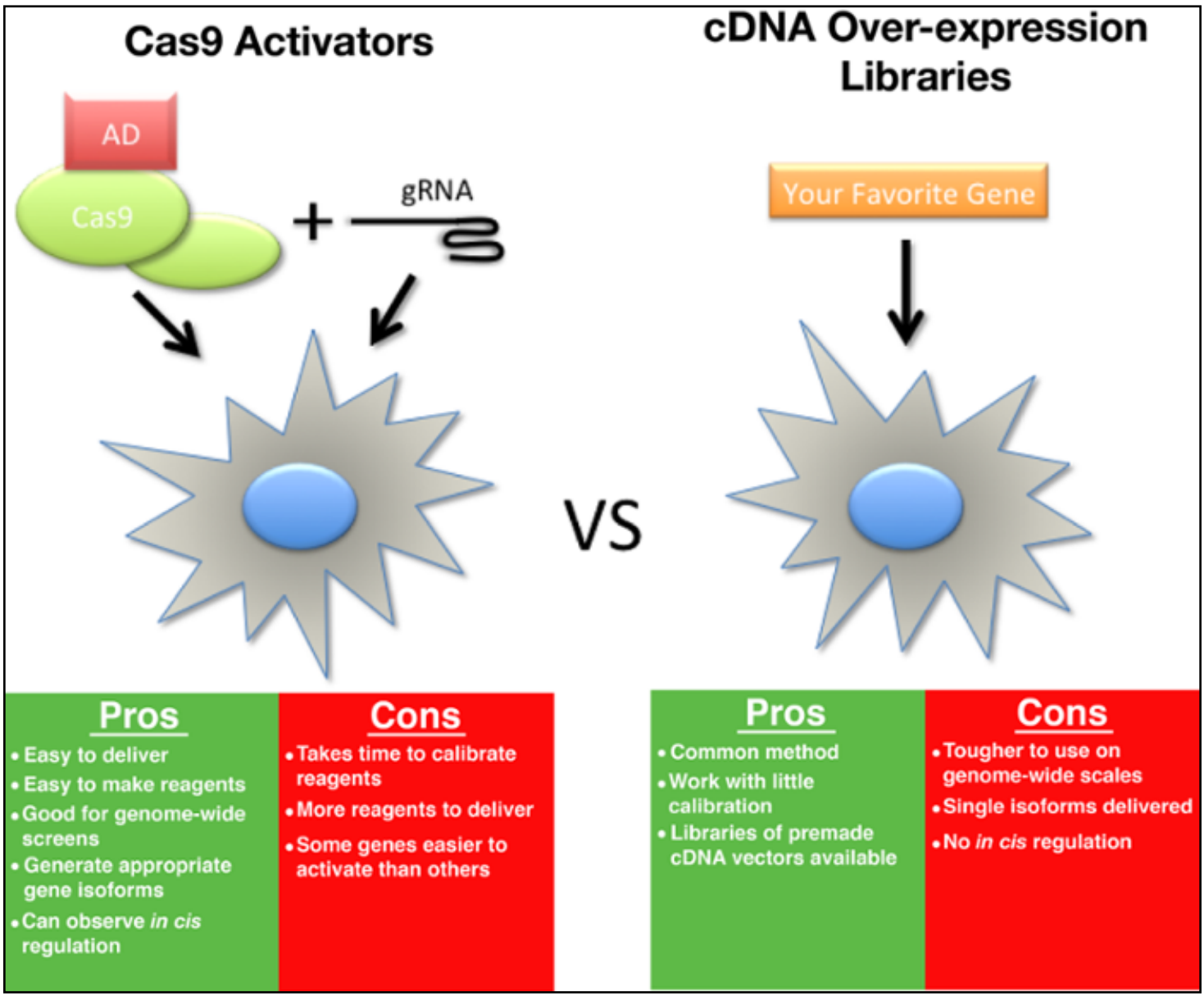
cDNA libraries, which consist of plasmids that over-express coding sequences from a given cell type or organism, have been used in a similar manner to Cas9 activators. However, these can be difficult to construct and deliver when compared to gRNAs. Additionaly, cDNAs cannot be used to study in cis regulation 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.

While Cas9 activators can be enormously powerful tools, they're not a good fit for every application. For example, Cas9 activators are particularly bad for experiments wherein you only want to activate a single gene; as 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 Cas9 activator; 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.

All of the above concerns with Cas9 activators are, for the most part, less of an issue in the setting of large scale genome-wide screens as the only consequence of a gene not being properly up-regulated in this setting is a potentially missed hit rather than complete experimental failure. 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.

Finally, we would like to point out that, while Cas9 activators are very useful for *in vitro* experiments, the technology is not exactly there yet for *in vivo* experiments. Most Cas9 activators are simply too large to fit all components into the most promising delivery vector, Adeno-associated virus (AAV). In the future, however,

CAS9 ACTIVATORS: A PRACTICAL GUIDE (CONT'D)



smaller Cas9 orthologs and transcription factor components could lead to an activator small enough to fit into the AAV chassis and yet still retain the ability to potently induce targeted gene expression.

Which Cas9 Activator Should I Choose

There are a wide variety of activators you can use for your experiments. We have found that SAM (1), Suntag (2, 3), and VPR (4) are good choices across multiple cell lines (HEK293T, MCF7, U2-OS, Hela, N2A, 3T3) and organisms (5). Our general advice, however, is to use whichever activator is most accessible to you and which you are most familiar with.

CAS9 ACTIVATORS: A PRACTICAL GUIDE (CONT'D)

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 (1, 4, 5) 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 (6) or our lab's sgRNA scorer 1.0 (7) 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.

Best of luck on your experiments!

Further Reading

1. Konermann, Silvana, et al. "Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex." Nature (2014). PubMed PMID: 25494202. PubMed Central PMCID: PMC4420636.
2. 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.
3. Gilbert, Luke A., et al. "Genome-scale CRISPR-mediated control of gene repression and activation." Cell 159.3 (2014): 647-661. PubMed PMID: 25307932. PubMed Central PMCID: PMC4253859.
4. Chavez, Alejandro, et al. "Highly efficient Cas9-mediated transcriptional programming." Nature methods 12.4 (2015): 326-328. PubMed PMID: 25730490. PubMed Central PMCID: PMC4393883.
5. Chavez, Alejandro, et al. "Comparison of Cas9 activators in multiple species."Nature methods (2016). PubMed PMID: 27214048.
6. Wong, Nathan, Weijun Liu, and Xiaowei Wang. "WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system." Genome biology16.1 (2015): 1. PubMed PMID: 26521937. PubMed Central PMCID: PMC4629399.
7. 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.

TRUNCATED GRNAS FOR REGULATING GENE EXPRESSION

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

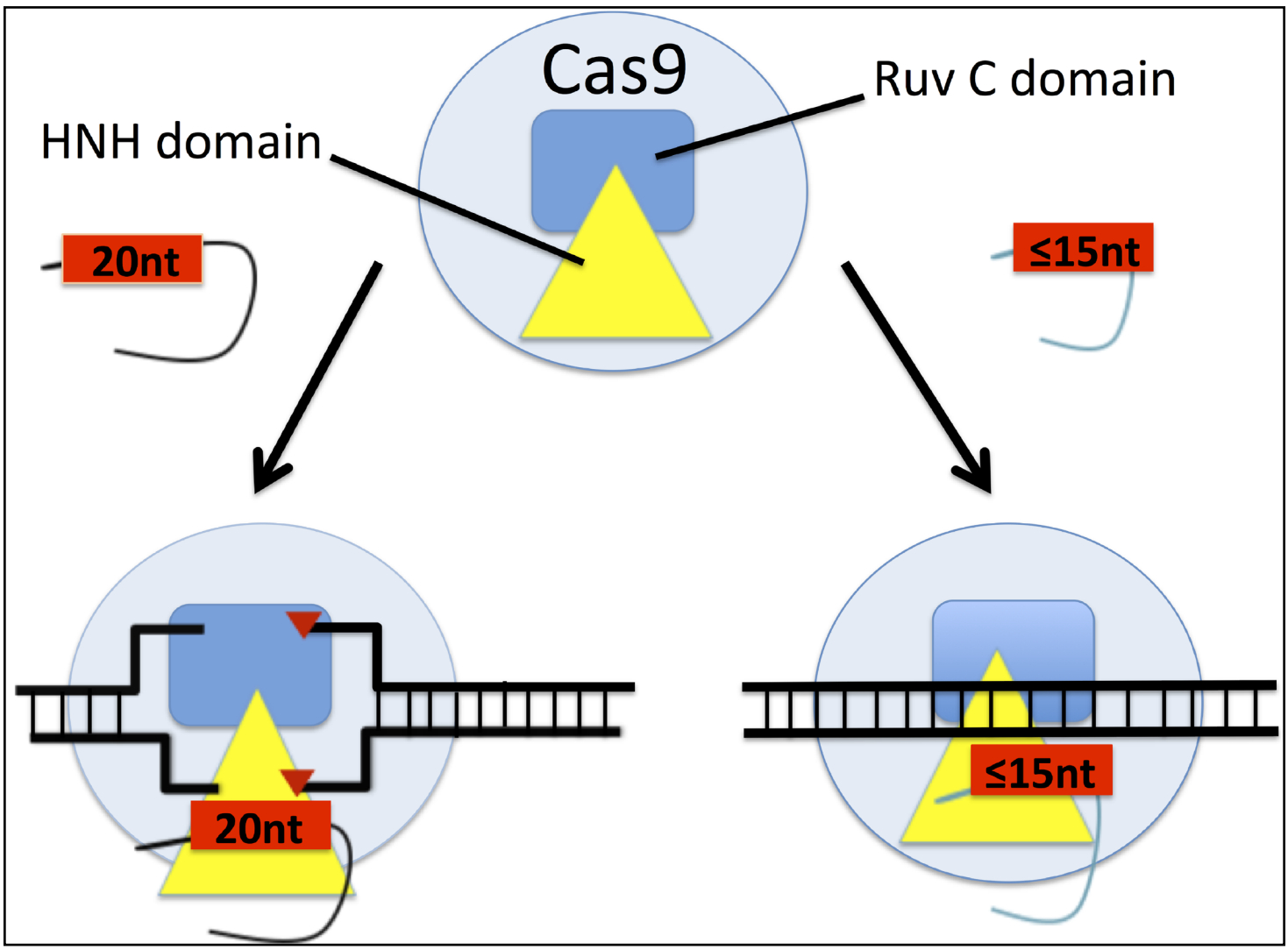


Figure 1: Schematic of native (nuclease-competent) Cas9 interacting with either a full-length gRNA (20 nt complementarity to target site) or truncated gRNA (less than or equal to 15 nt complementarity to target site). When complexed with a full-length gRNA, native Cas9's two catalytic domains, HNH and 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.

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.

TRUNCATED GRNAS FOR REGULATING GENE EXPRESSION (CONT'D)

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 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 ≤15 nt 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).

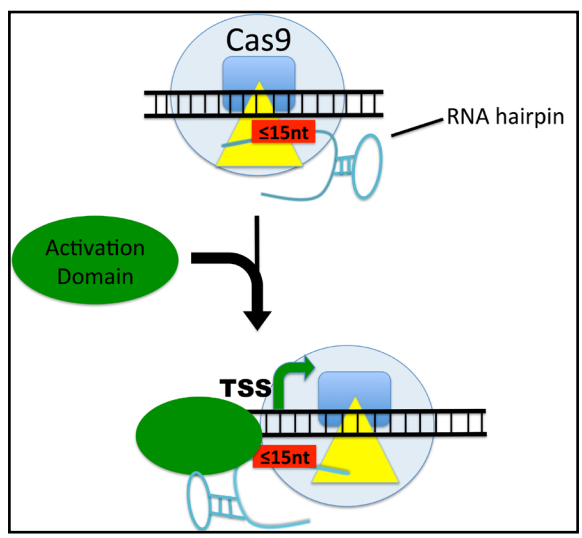


Figure 2: Gene activation with nuclease-competent Cas9. An RNA hairpin incorporated into the truncated gRNA recruits a transcriptional activator to the Cas9-gRNA complex. When this complex is directed to a target upstream of the transcriptional state site (TSS) of a gene of interest, it induces potent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

TRUNCATED GRNAS FOR REGULATING GENE EXPRESSION (CONT'D)

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 20 nt 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 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.

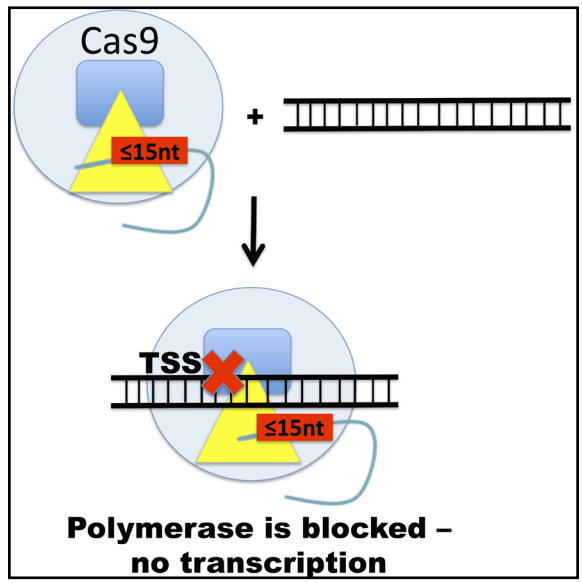


Figure 3: Gene repression with nuclease-competent Cas9. When complexed with a truncated gRNA, Cas9 is still able to bind DNA but is not longer competent to cut DNA. When directed to a site downstream of the transcriptional start site of a gene of interest, Cas9 forms a steric block that prevents RNA polymerase from efficiently transcribing the gene. This results in gene silencing with no observed genetic alterations at the target site.

Further Reading

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. PubMed Central PMCID: PMC5292764.
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.

PLASMIDS FOR EDITING THE EPIGENOME

By Mary Gearing | February 14, 2017

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.

Maeder et al. 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. Mendenhall et al. additionally fused a TAL effector to LSD1 histone demethylase to demethylate enhancer regions. 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

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 the recently published pLV-dCas9-p300-P2A-PuroR for lentiviral expression.

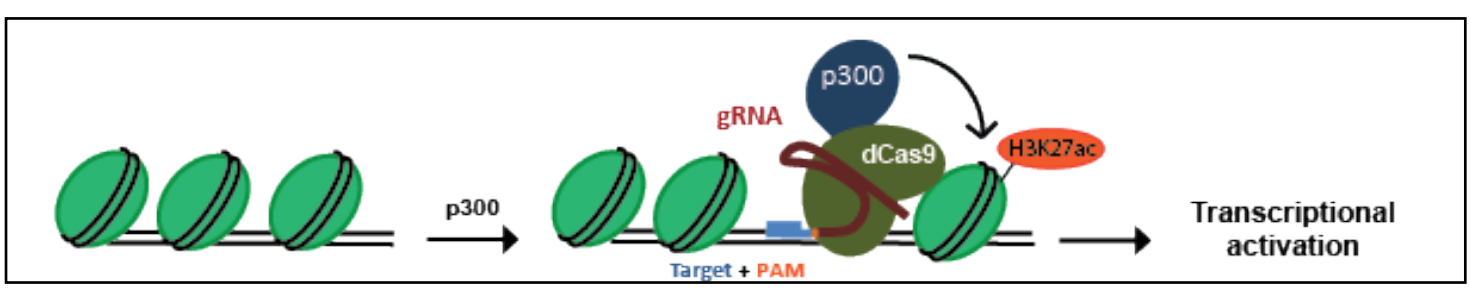


Figure 1: dCas9-p300 adds H3K27ac marks to histones in promoter and enhancer regions. These marks are associated with transcriptional activation.

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.

PLASMIDS FOR EDITING THE EPIGENOME (CONT'D)

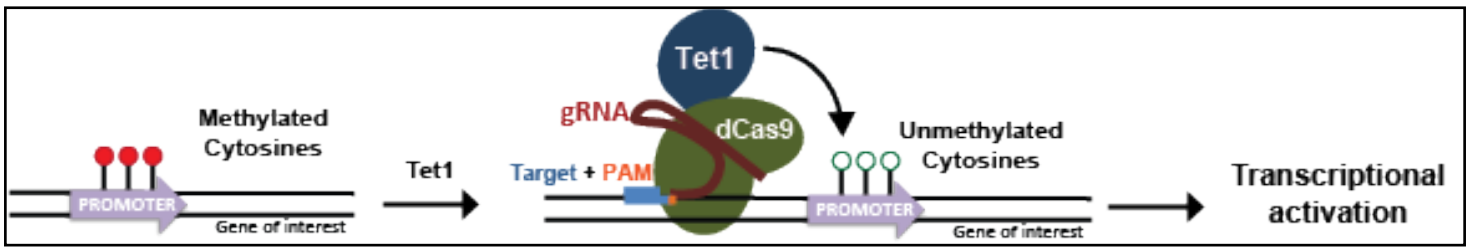


Figure 2: dCas9-Tet1 demethylates cytosines at promoters and enhancers. This targeted demethylation is associated with increased transcription.

Repression

DNA Methyltransferase 3 Alpha

Vlatka Zoldoš' lab has deposited pdCas9-DNMT3A-EGFP and pdCas9-DNvMT3A-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.

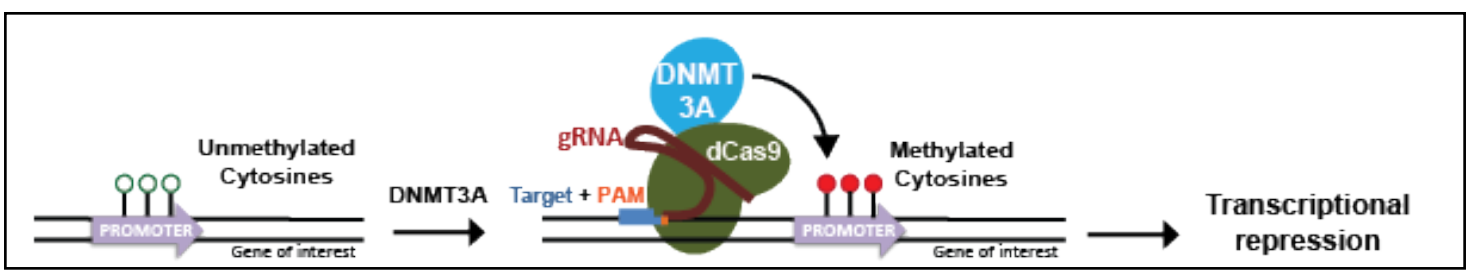


Figure 3: dCas9-DNMT3A methylates cytosines at promoters and enhancers. This targeted methylation is associated with transcriptional repression.

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. Stolzenburg et al. compared ZFN-KRAB and ZFN-DNMT3A, finding that KRAB induced silencing was transient and quickly reversed in 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 - Hilton et al. found that

PLASMIDS FOR EDITING THE EPIGENOME (CONT'D)

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. Let us know about your experience with these constructs by emailing us at blog@addgene.org.

Further Reading

1. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. 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. Nat Biotechnol. 2013 Oct 9. PubMed PMID: 24108092. PubMed Central PMCID: PMC3858462.
2. Locus-specific editing of histone modifications at endogenous enhancers. Mendenhall EM, Williamson KE, Reyon D, Zou JY, Ram O, Joung JK, Bernstein BE. Nat Biotechnol. 2013 Sep 8. PubMed PMID: 24013198. PubMed Central PMCID: PMC3858395. PubMed Central PMCID: PMC3858395.
3. A CRISPR-based approach for targeted DNA demethylation. Xu X, Tao Y, Gao X, Zhang L, Li X, Zou W, Ruan K, Wang F, Xu GL, Hu R. Cell Discov. 2016 May 3;2:16009. PubMed PMID: 27462456. PubMed Central PMCID: PMC4853773.
4. Editing DNA Methylation in the Mammalian Genome. Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, Shu J, Dadon D, Young RA, Jaenisch R. Cell. 2016 Sep 22;167(1):233-247.e17. PubMed PMID: 27662091.
5. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA. Nat Biotechnol. 2015 May;33(5):510-7. doi: 10.1038/nbt.3199. PubMed PMID: 25849900. PubMed Central PMCID: PMC4430400.
6. CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. Klann TS, Black JB, Chellappan M, Safi A, Song L, Hilton IB, Crawford GE, Reddy TE, Gersbach CA. Nat Biotechnol. 2017 Apr 3. doi: 10.1038/nbt.3853. PubMed PMID: 28369033.
7. 7. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Vojta A, Dobrinic P, Tadic V, Bockor L, Korac P, Julg B, Klasic M, Zoldos V. Nucleic Acids Res. 2016 Mar 11. PubMed PMID: 26969735. PubMed Central PMCID: PMC4937303.
8. Stable oncogenic silencing in vivo by programmable and targeted de novo DNA methylation in breast cancer. Stolzenburg S, Beltran AS, Swift-Scanlan T, Rivenbark AG, Rashwan R, Blancafort P. Oncogene. 2015 Oct;34(43):5427-35. PubMed PMID: 25684141. PubMed Central PMCID: PMC4633433.
9. Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. McDonald JI, Celik H, Rois LE, Fishberger G, Fowler T, Rees R, Kramer A, Martens A, Edwards JR, Challen GA. Biol Open. 2016 May 11. PubMed PMID: 27170255. PubMed Central PMCID: PMC4920199.

HOW TO EXPERIMENT WITH ENHANCERS AND EPIGENOMICS

By Aneesh Karve | Aug 9, 2016

This section was contributed by guest writer, Aneesh Karve, CTO at Qult Data. This section was originally published on the Quilt Genomics Blog and is republished here with permission.

Quilt is a collaborative database for genomics. In this section, Quilt CTO Aneesh Karve, shows how to design experiments that work anywhere in the genome. Aneesh's research interests include proteomics, machine learning, and visualization for big biology.

GPS for the Genome

Technique	What it Locates
ChIP-seq	Proteins (for our example later, histones)
ChiRP-seq	RNA
Hi-C	DNA (genome-to-genome interactions)
DNase-seq	DNA (regions that are accessible for binding)

We can think of the human genome as a map with three coordinates: chromosome, start, and stop. For instance (chr3, 1, 10) indicates a stretch of DNA at the very beginning of the third chromosome, ten base pairs in length. An emerging family of sequencing techniques function as a kind of "GPS for the genome" to compute coordinates for genetic elements like protein, RNA, and DNA (Table 1). As with GPS in the real world, coordinates alone aren't very useful. We'll need something like Google Maps to help us identify and visualize addresses. That's where enhancers and genome math come in. They help us to transform raw genomic coordinates into meaningful experiments.

Google Maps: Enhancers and Genome Math

Suppose that you wish to use Google maps to find all coffee shops near your house, excluding Starbucks. Taking a nerdy perspective, you might denote your search as follows:

(my_house + coffee) - starbucks

See how the notation works? The + operator denotes intersection and the – operator denotes set difference. That's the intuition for how genome math helps us to locate interesting addresses in the genome. Let's now examine how we can locate powerful stretches of DNA known as enhancers with the help of genome math.

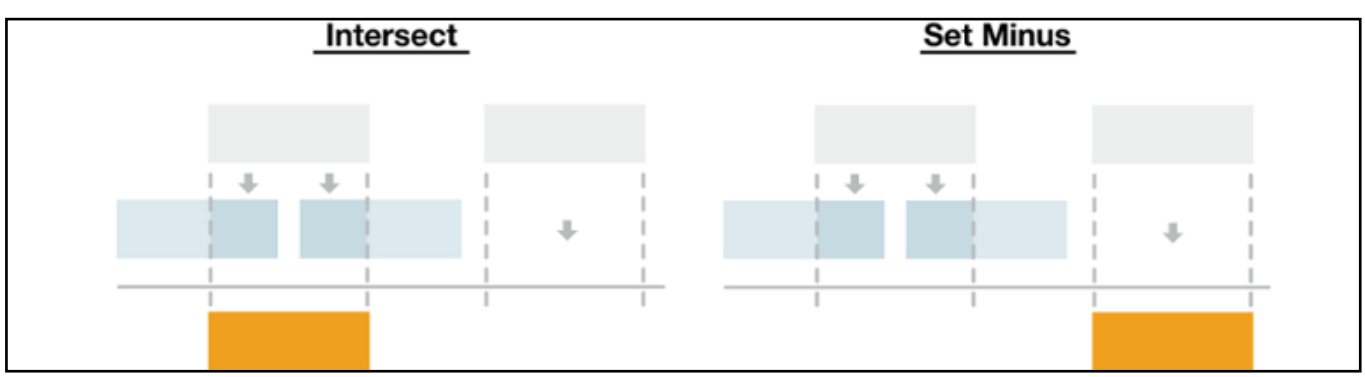


Figure 1: Genome math illustrated. Intersect (left) and set minus (right).

Enhancers are regions of DNA that demonstrate "spooky action at a distance." Through the marvel of DNA

HOW TO EXPERIMENT WITH ENHANCERS AND EPIGENOMICS (CONT'D)

compaction, an enhancer can increase the expression of a gene that is millions of base pairs away. (For details on DNA compaction and the structural proteins that make it possible, see the appendix: DNA is a 3D fractal).

Enhancer biology is a complex and dynamic field. We're going to focus on a tried and true method of finding enhancers by isolating genomic regions that are bound to modified proteins called histones. We can detect modified histones with "GPS for proteins," ChIP-seq from Table 1. Because of DNA's 3D geometry and the chemical properties of modified histones, a genomic region that has mono-methylated and acetylated histones, but not tri-methylated histones, functions as an enhancer. We can therefore denote enhancers as follows:

(mono_methylation + acetylation) - tri_methylation

In the next section we'll apply the above formula to a real-world experiment. We'll start with ChIP-seq data from the ENCODE project, find enhancers in embryonic stem cells, and conclude with a targeted CRISPR screen that can disrupt these enhancers.

A Real-World Experiment

Suppose you run a ChIP-seq experiment (think "GPS for proteins") for NANOG, an essential transcription factor in embryonic stem cells (ESCs). Your ChIP-seq finds just over 13,000 significant binding peaks for NANOG in the human genome. But not all of those 13,000 regions are important for maintaining ESCs. So which of these 13,000 regions are critical? One hypothesis: the enhancers! This leads us to a three-step approach for designing an experiment to identify the critical NANOG binding sites:

1. Find Enhancers that have NANOG binding sites

2. Design a CRISPR screen to target and disrupt the NANOG enhancers

3. CRISPR out the enhancers from step 2. See which ESCs die or differentiate

Step 3 reveals which NANOG-related genes are critical to stem cell survival. Knowing which genes influence the survival of our cell culture is the foundation of modern drug discovery and therapeutics. We'll have more to say about clinical applications of CRISPR in the next section.

In order to denote the NANOG enhancers from step 1 with genome math, we'll need a bit of shorthand from the field of epigenomics:

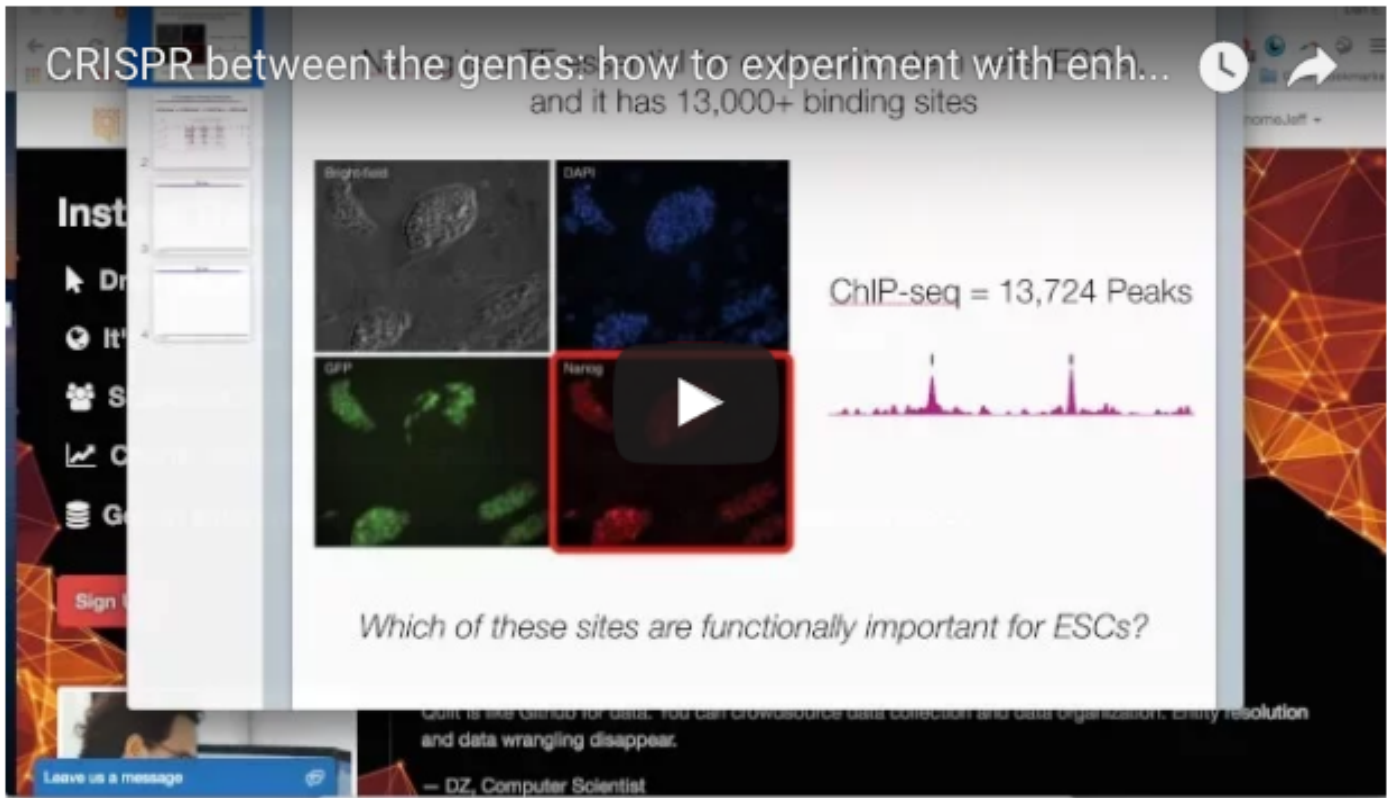
- H3 – one of NANOG's associated histone proteins
- K4 and K27 – locations of the amino acid lysine in H3
- me1, me3, and ac – denoting mono-methylation, tri-methylation, and aceylation, respectively (these are chemical modifications, or functional groups, found on lysine)

Putting it all together, we get the following expression for step 1:

(H3K4me1 + H3K27ac) – H3K4me3

The video linked on the next page demonstrates how anyone can find enhancers with Quilt.

HOW TO EXPERIMENT WITH ENHANCERS AND EPIGENOMICS (CONT'D)



Armed with the genome math expression for NANOG enhancers, we're ready to design a CRISPR screen to disrupt these enhancers. This is a cutting edge application of CRISPR. If you try it, we want to hear from you :-). Until now, the vast majority of CRISPR screens have targeted protein-coding genes. We're walking into terra incognita by targeting epigenomic elements, like enhancers.

The third and final step is to conduct our CRISPR screen. We start by infecting millions of embryonic stem cells (ESCs) with a lentiviral vector, an attenuated retrovirus in the same family as HIV. By design, our lentivirii are genetically programmed to CRISPR out the enhancers we identified in step 2. The result is a heterogenous population of stem cells, usually housed in a single flask. Through a bit of stochastic magic and Poisson statistics, each sub-population has, on average, one distinct enhancer disrupted. As our ESCs die and differentiate over time, we periodically use next-generation sequencing to measure the relative proportion of guide RNAs (gRNAs) across the population. Recall that guide RNAs are the targeting mechanism for CRISPR. Therefore if a gRNA drops or disappears over time, we infer that the enhancer it targets is a "pillar of function" for our stem cells. Remove this pillar and the ESC dies.

If you're interested in designing your own CRISPR screens for enhancers, check out the Appendix.

Conclusion: Shedding Light on Dark Matter

Precise knowledge of which stretches of the genome are pillars for stem cells, or metastasized tumor cells, or alzheimers-affected neurons, or [your cell line of interest], is the foundation of precision medicine. We can apply this knowledge to create targeted disease therapies with minimal side-effects on healthy cells and maximal effect on unhealthy cells.

HOW TO EXPERIMENT WITH ENHANCERS AND EPIGENOMICS (CONT'D)

Until recently, the human genome has been full of dark matter: enhancers, lncRNAs, repetitive elements, repressors, insulators, and more. We know that this matter exists, but traditional approaches to study its function have been prohibitively difficult. CRISPR, in combination with the techniques from Table 1, provides us with powerful GPS-like techniques to explore dark matter in the genome. There are countless unknown regions yet to explore. I hope that this brief guide can help you do just that.

To design your own enhancer screens, try the Universal Guide RNA designer to search for guide RNAs by gene name or by gene coordinate.

If you're working with scientific data, we built Quilt for you. Try it and tell us what you think.

Good luck, and always keep going.

Appendix

DNAse Hypersensitive Sites and Intergenic CRISPRs

Yours truly generated the DHS data set by starting with all of the DHS sites from 125 different human cell types from the ENCODE Project. DHS sites are the most inclusive markers of regulatory regions in the genome, including enhancers, promoters, insulators, and more. I then identified valid gRNA sequences without off-target effects for the 2+ million DHS sites. See the appendix on gRNA selection and filtering for off-target effects for further details.

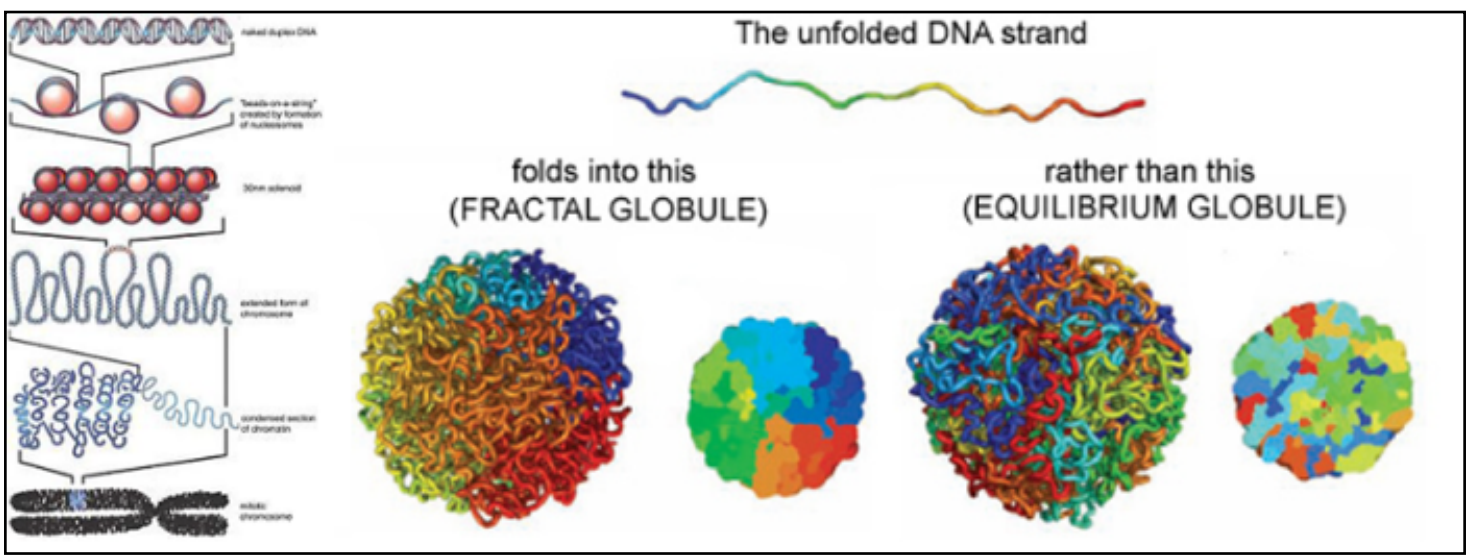


Figure 2: DNA Compaction. Graphics courtesy of beyondthedish (left) and biobabel (right)

People commonly think of DNA as a linear polymer of A,T,G, and C nucleotides arranged a double helix. People are wrong :-). In reality, cellular DNA is a complex three-dimensional globule of chromatin. Chromatin is a combination of coiled DNA and structural proteins called histones. Chromatin folds into secondary structures

HOW TO EXPERIMENT WITH ENHANCERS AND EPIGENOMICS (CONT'D)

(loops) and tertiary structures (globules) to achieve exquisite compaction—on the order of 700 terabytes per gram—and form the X-shaped chromosomes we all know and love.

To understand compaction, suppose you have a piece string that's 10 meters long. You twist and roll the string into a tight ball. Now the two ends of the string, instead of being 10 meters apart, are mere millimeters apart. Similarly, cells compact DNA in a way that brings linearly distant regions close together.

gRNA Selection and Filtering for Off-target Effects

We first generated a multi-fasta file of the hg19 genome using Bedtools getFasta. These regions and their reverse complements were parsed for spCas9 PAM sites (NGG) and then filtered based on two main criteria: no TTTTT allowed (this is a polymerase terminator), and no off-target effects for the identified 23-mer gRNA. Off-target determination was established with Bowtie2 using the parameters first described in Kearns et al.:

```
bowtie2 -f -x HG19_GENOME --local -f -k 10 --very-sensitive-local -L 9 -N 1 -U
GRNA_23MERS -S GRNA_HITS.sam
```

TAGGING GENES WITH CRISPR

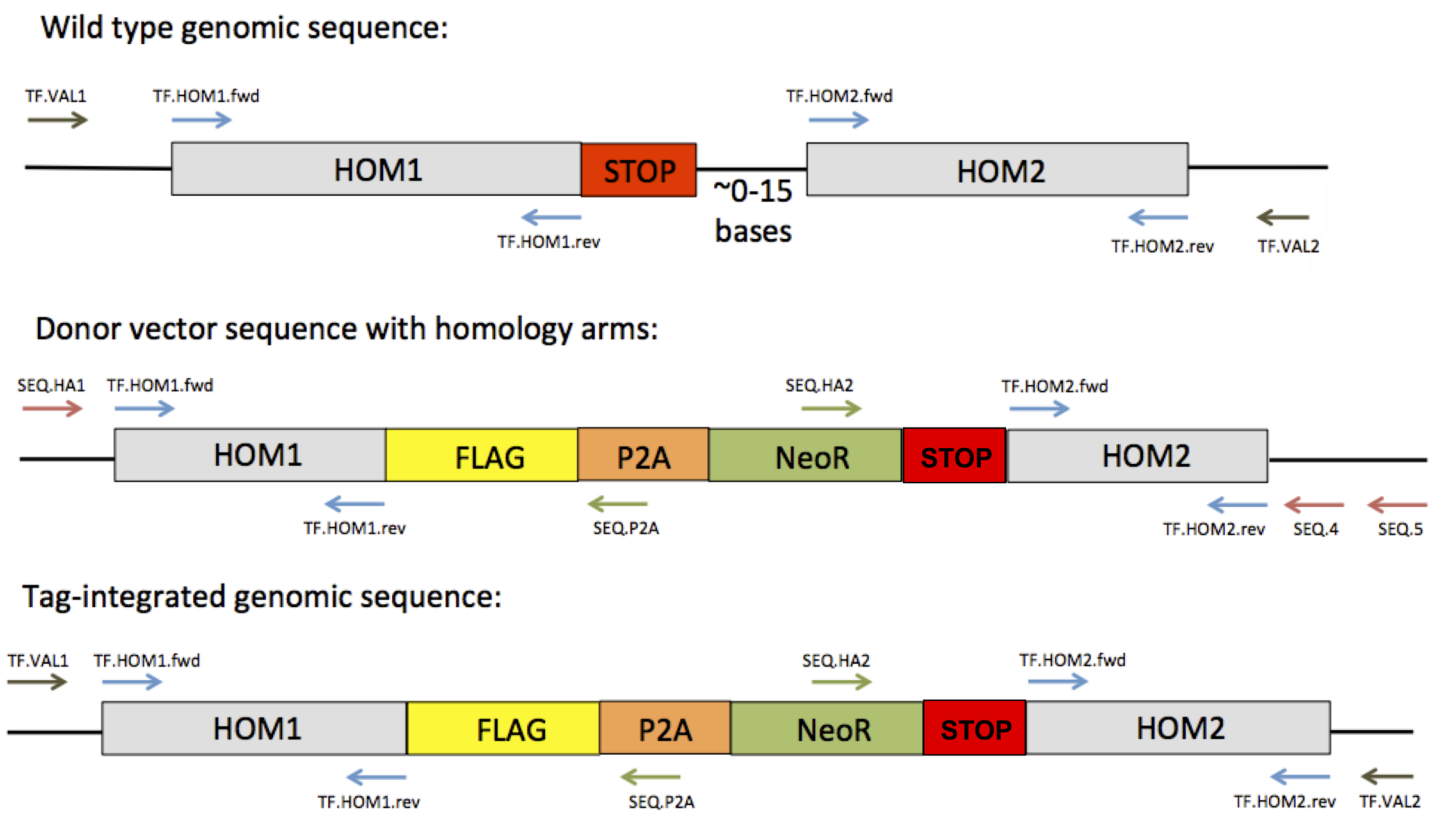
By Mary Gearing | December, 2015

First described in the 1980s, protein tags are now one of the most useful items in a scientist's toolbox. As we've covered in Plasmids 101, tags can help you determine localization of a protein of interest, purify it, or determine its expression level without the need for a custom antibody. CRISPR has made it easier than ever to tag endogenous proteins, allowing researchers to track how proteins bind to DNA or to other proteins.

The Basics of CRISPR Tagging

Addgene depositors Eric Mendenhall and Richard Myers have developed a method to easily insert FLAG tags into the C-termini of endogenous loci. To tag your protein of interest, first design a seed sequence targeting the locus you'd like to tag, and then insert it into a plasmid containing Cas9 and the gRNA scaffold. Ideally the gRNA should cut within -5 to +15 base pairs from the stop codon.

Second, create a repair template according to the guide below. Mendenhall and Myers recommend using IDT gBlocks to specify the homology arms of the repair template. Plasmid pFETCh_Donor contains the 3X-FLAG tag and a neomycin resistance gene; the homology arms can be cloned into the vector using Gibson Assembly. Once you introduce the construct into cells, you'll select using neomycin. Isolating clonal populations is not necessary; Mendenhall and Myers use pooled neomycin-positive cells for experiments.



A schematic for FLAG tagging endogenous proteins using plasmid pFETCh_Donor. Left and right homology arms are cloned into the FLAG-P2A-NeoR containing destination vector using Gibson Assembly. Homology-directed repair removes the stop codon and adds a 3X-FLAG tag and neomycin resistance. NeoR is co-transcribed with the FLAG-tagged TF. The P2A self cleaving linker causes the ribosome to skip a peptide bond between the Flag-tagged protein and the antibiotic resistance cassette resulting in two separate protein products.

TAGGING GENES WITH CRISPR (CONT'D)

Application: Multiplexed Transcription Factor Tagging

Mendenhall and Myers are members of the ENCyclopedia of DNA Elements Project (ENCODE), working to define where the ~1,500 transcription factors of the human genome bind. Since fewer than 10% of antibodies are suitable for the ChIP-seq analysis commonly used to map transcription factor binding, they created CETCh-seq (CRISPR epitope tagging ChIP-seq of DNA-binding proteins) to tag transcription factors and analyze their binding in a scalable, global approach.

To test the universality of CETCh-seq, Savic et al. selected five DNA-binding proteins expressed at different levels in HepG2 cultured cells. They designed gRNAs to target the 3' UTR and used an EMM0021-based repair template to add a FLAG tag to the C-terminus of each TF. They screened for homologous recombination using PCR and subsequently verified tag insertion via Western blotting and Sanger sequencing. From the five TFs targeted in parallel, Savic et al. successfully tagged four TFs.

Savic et al. subsequently conducted CETCh-seq using a FLAG antibody. In cells that did not contain a tagged TF, they did not observe binding events, showing that the method has low background. To validate their results, they compared binding from CETCh-seq to datasets obtained using ChIP-seq with verified antibodies, finding an average of ~85% overlap in binding sites between the datasets. Technical and biological CETCh-seq replicates were also highly concordant (rho=0.92-0.98), indicating the robustness and specificity of the technique. Using RNA-seq, Savic et al. verified that TF tagging does not alter the transcriptome, further strengthening the case for CETCh-seq as new method to profile TF binding.

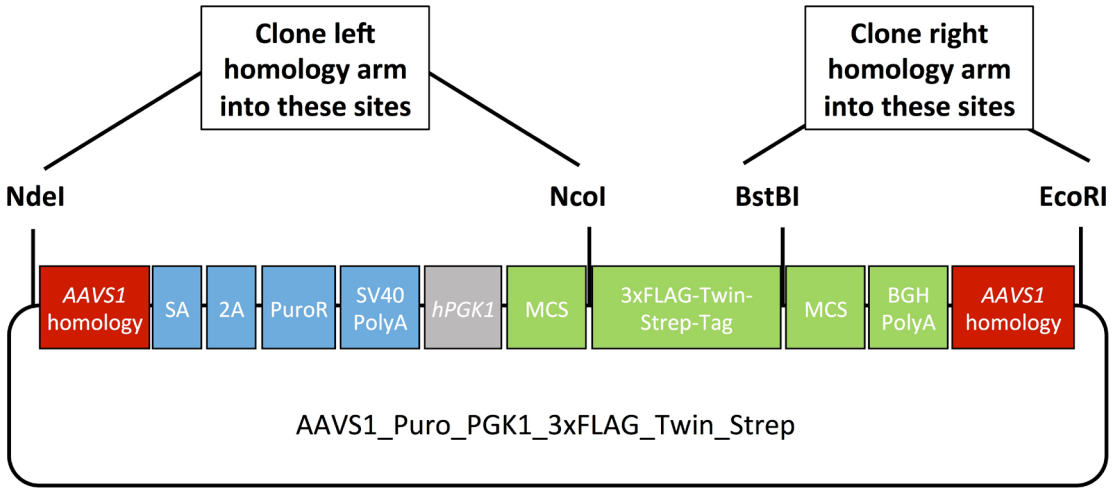
After completing the initial experiments in HepG2 cells, Savic et al. turned to MCF7 cells to verify that CETCh-seq is robust in multiple cell types. For their targeted locus, RAD21, they again found good technical and biological reproducibility, as well as concordance with validated ChIP results. CETCh-seq was also successful in murine embryonic stem cells, opening up the possibility of generating transgenic mice through this tagging approach!

Application: Affinity Purification Tagging for Protein Complex Isolation

Addgene depositor Yannick Doyon is interested in isolating native protein complexes to study biochemical interactions. Dalvai et al. designed a system to add 3X-FLAG-2X-STREP tags to the N- or C-termini of endogenous loci, enabling gentle, high-yield purification of protein complexes. Tagging endogenous loci gets rid of many problems associated with protein overexpression, such as nonphysiological binding. The system can also be used to insert a cDNA into the AAVS1 locus, a "safe harbor" locus that does not interrupt the function of other genes.

To tag genes using the Doyon system, you need a gRNA that cleaves near the terminus of the locus you'd like to tag, as well as a repair template based on Addgene AAVS1_Puro_PGK1_3xFLAG_Twin_Strep. After constructing the homology arms via PCR or with IDT gBlocks, sequentially digest the vector to insert the left arm, then the right arm. A basic schematic is shown below, but further details are available in the supplemental methods of Dalvai et al. If you'd like to use this plasmid to insert a cDNA into AAVS1, just use the multiple cloning site to insert the sequence downstream of the PGK1 promoter.

TAGGING GENES WITH CRISPR (CONT'D)



A schematic for 3X-FLAG-2X-STREP protein tagging. To use plasmid AAVS1_Puro_PGK1_3xFLAG_Twin_Strep for endogenous protein tagging, digest with NdeI and NcoI to insert the left homology arm, then digest that construct with BstBI and EcoRI to insert the right arm. If inserting a tagged cDNA into AAVS1, clone the cDNA using the MCS downstream of hPGK1.

Advantages and Tips

These plasmid systems promise to make endogenous protein tagging much easier and faster than ever before. Using gBlocks speeds up cloning, and CRISPR greatly increases the frequency of homologous recombination. Although these systems are FLAG- and STREP tag-based, they can be adapted to other tags, allowing the tagging of multiple loci within a cell population.

One potential pitfall of CRISPR tagging is that the gRNA must bind close to the targeted terminus. If you can't find a gRNA that works for your locus using SpCas9, it may make sense to try alternative Cas9s or Cpf1, which have different PAM requirements.

Ready to start tagging? Plasmids for FLAG and FLAG-STREP tagging are available at Addgene.

Further Reading

1. Savic, Daniel, et al. "CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins." Genome Research 25 (2015): 1581–1589. PubMed PMID: 26355004.
2. Dalvai, Mathieu, et al. "A Scalable Genome-Editing Based Approach for Mapping Multiprotein Complexes in Human Cells." Cell Reports 13 (2015): 621-633. PubMed PMID: 26456817.

USING CRISPR TO VISUALIZE THE GENOME

By Mary Gearing | December, 2015

The Basics of CRISPR Tagging

CRISPR may be best known for genome editing and transcriptional activation, but it's also helping researchers visualize the genome and its organization within the nucleus, also called the 4D nucleome. Visualizing specific loci has historically been difficult, as techniques like fluorescent *in situ* hybridization (FISH) and chromosome capture suffer from low resolution and can't be used *in vivo*. Some researchers have used fluorescently tagged DNA-binding proteins to label certain loci, but this approach is not scalable for every locus...unlike CRISPR!

Using dCas9 to Visualize Genomic Loci *In Vivo*

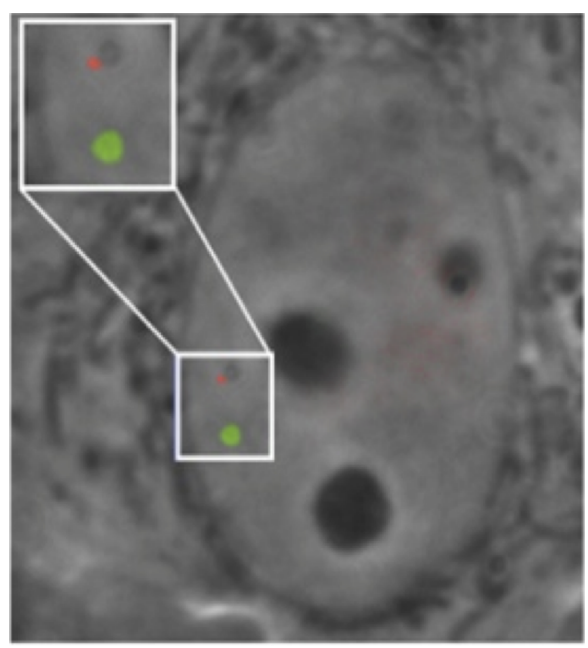
The flexibility of CRISPR has improved our ability to target almost every genomic locus, and to do so *in vivo*. All you need is fluorescently labeled dCas9 and a gRNA specifying your locus of interest. This technique was developed in the lab of Addgene depositor Stanley Qi and used in Chen et al. to track telomere dynamics throughout the cell cycle. To optimize the signal to noise ratio of the system, they modified GFP-dCas9 and the gRNA scaffold to enhance complex assembly, thereby decreasing the amount of background fluorescence from unbound GFP-dCas9. With these improvements, labeling efficiency was similar to that of a comparable FISH method. Importantly, they did not observe labeling when a) no gRNA was present or b) when a gRNA binding a non-mammalian sequence (GAL4) was supplied.

In addition to labeling repetitive telomeric sequences, Chen et al. successfully labeled protein-coding genes with both intronic and exonic gRNAs. In fact, the method is specific and sensitive enough to detect gene copy number based on the number of fluorescent puncta observed. By labeling two genes simultaneously, they could also monitor the spatial relationship of the two genes over time.

Labeling Multiple, Intrachromosomal Loci

Building on the work of Chen et al., Thoru Pederson's lab has used CRISPR to label multiple loci in distinct colors. To create a colorful Cas9 toolbox, Ma et al. turned to SpCas9 and its orthologs NmCas9 and St1Cas9. Each ortholog was fused to a different fluorescent protein to create three distinct colors. The specificity of these orthologs is key: since each ortholog requires a different PAM sequence, a gRNA designed for one dCas9 should be specific to that ortholog and not cross-talk with the other orthologs.

Ma et al. tested their dCas9 variants using gRNAs specific for telomeric sequences and showed that different fluorescently labeled dCas9s are efficiently directed to the proper target sequence. They succeeded in labeling two different pairs of chromosomes using gRNAs specific to sequences on chromosomes 9 and 13. They next turned their attention to mapping pairs of intrachromosomal loci. The technique successfully resolved loci with physical map distances of 2 and 75 Mbp, with the calculated fluorescent distances



Live cell dual-color CRISPR labeling of two loci on human chromosome 9

USING CRISPR TO VISUALIZE THE GENOME (CONT'D)

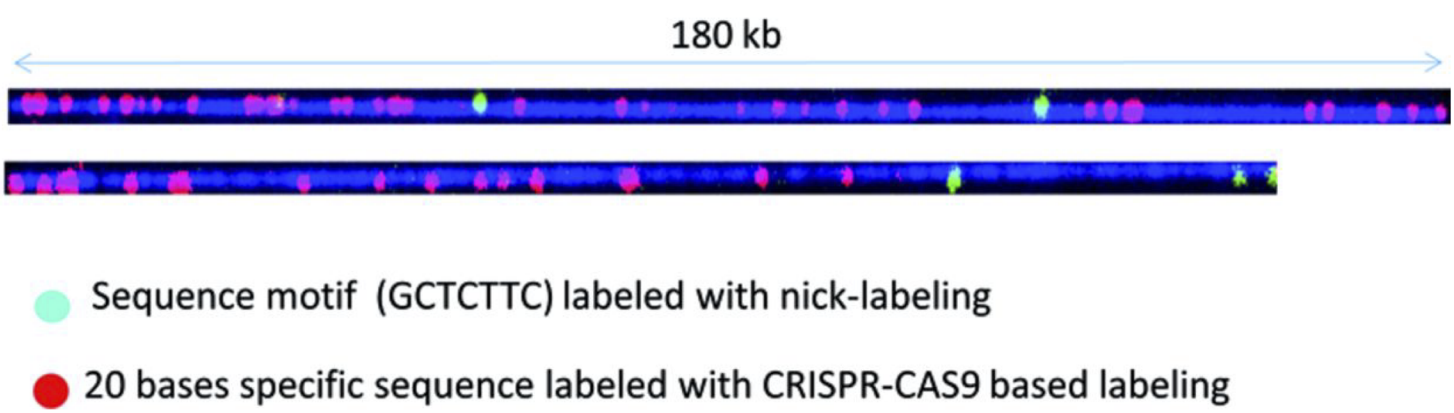
correlating with the previously established physical map. In comparing pairs of targets ~2 Mbp apart, they noticed that they could evaluate the degree of chromatin compaction even for this small distance! To the authors' knowledge, this work represents the first mapping of intrachromosomal loci, a major benchmark in characterizing the 4D nucleome.

In Vitro Application: Genome Structure

In addition to 4D nucleome applications, CRISPR visualization is also useful for studying genome structure. Shotgun sequencing breaks DNA up into small segments, which are then assembled based on overlapping sequence. However, in areas with repetitive or duplicated sequences, the sequences may be assembled incorrectly. Optical mapping can create restriction maps of the genome, better capturing structural variation, but areas that do not contain restriction sites cannot be mapped.

McCaffrey et al. adapted CRISPR for whole genome mapping by designing gRNAs to specifically target repetitive sequences. In their technique, DNA strands are incubated with Cas9 nickase, a gRNA, Taq polymerase and fluorescently labeled nucleotides. At the site of the nick, the polymerase incorporates the labeled nucleotides. Labeled DNA is subsequently imaged in a nanochannel. When testing their method using a bacterial artificial chromosome, McCaffrey et al. found a very low false positive rate of 0.6%, indicating limited labeling in the BAC region not containing the gRNA-specified sequence. The technique can also be used with traditional sequence motif labeling, which recognizes various restriction sites, to map the location of gRNA-specified repeats.

Using gRNAs binding Alu sequences, which account for up to 10% of the human genome, McCaffrey et al. created barcode-like patterns that differ among DNA sequences. This technique may become especially useful for loci of clinical significance, such as the polyglutamine tracts in Huntington's disease, where certain variants may prove pathogenic. Importantly, this advance makes genome mapping more flexible and high-throughput, for both targeted and whole genome approaches.



CRISPR labeling creates information-rich barcodes. Two pieces of human genomic DNA were labeled using an Alu sequence-specific gRNA (red) as well as traditional nick-labeling with the restriction enzyme Nt. BspQI (blue) Whereas the Nt.BspQI recognition site is found infrequently in genomic DNA, the Alu sequences are highly abundant, permitting the generation of genome-specific barcodes. Figure from McCaffrey et al., 2015. Used under a CC-BY-NC 4.0 license.

USING CRISPR TO VISUALIZE THE GENOME (CONT'D)

Further Reading

1. Chen, Baohui, et al. "Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System." Cell 155 (2013): 1379-1491. PubMed PMID: 24360272. PubMed Central PMCID: PMC3918502.
2. Ma, Hanhui, et al. "Multicolor CRISPR labeling of chromosomal loci in human cells." Proc. Natl. Acad. Sci. USA 112(10) (2015): 3002-7. PubMed PMID: 25713381. PubMed Central PMCID: PMC4364232.
3. McCaffrey, Jennifer, et al. "CRISPR-CAS9 D10A nickase target-specific fluorescent labeling of double strand DNA for whole genome mapping and structural variation analysis." Nucleic Acids Research (2015) pii: gkv878. PubMed PMID: 26481349.

OPTOGENETICS + CRISPR: USING LIGHT TO CONTROL GENOME EDITING

By Caroline LaManna | March 8, 2016

Scientists around the world have been making major improvements to CRISPR technology since it's initial applications for genome engineering in 2012. Like CRISPR, optogenetics has also been making headlines over the past decade. Optogenetics uses genetically encoded tools, such as microbial opsins, to control cellular activities using light. In 2015, scientists combined CRISPR and optogenetics 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 (Nihongaki Y, et al., Chemistry & Biology, 2015 Feb 19; 22(2):169-74) and Charles Gersbach's lab at Duke University (Polstein LR, et al., Nature Chemical Biology, 2015 Mar; 11(3): 198-200) 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 Nihongaki et al. 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 (A). 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, including making alterations to the CIB1 domain, testing various activator probes, and adding various genomic anchors to the N-terminus of both fusion constructs. 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. The authors also cycled blue light on and off and showed that ASCL1 expression followed suit - control was indeed reversible and repeatable.

With their light-activated CRISPR/Cas9 effector (LACE) system Polstein et al., utilized a similar strategy to develop an optimized photoactivatable CRISPR gene activation system, but settled on a different optimal fusion protein combination. Shown in the figure (B), 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 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.

OPTOGENETICS + CRISPR: USING LIGHT TO CONTROL GENOME EDITING (CONT'D)

Photoactivatable Genome Modifications by NHEJ and HDR

Later in 2015, the Sato lab unveiled a photoactivatable system to cleave a target DNA sequence (Nihongaki Y, et al., Nature Biotechnology, 2015 Jul; 33(7):755-60) 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 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 (C). Although the authors tried a few different photoactivatable designs (some similar to those used in the previous Nihongaki et al. system) their most successful design utilized Magnet photoswitchable proteins derived from the fungal photoreceptor, Vivid (VVD, N. crassa) (Kawano F, et al., Nature Communications, 2015 Feb 24; 6:6256). 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%).

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.

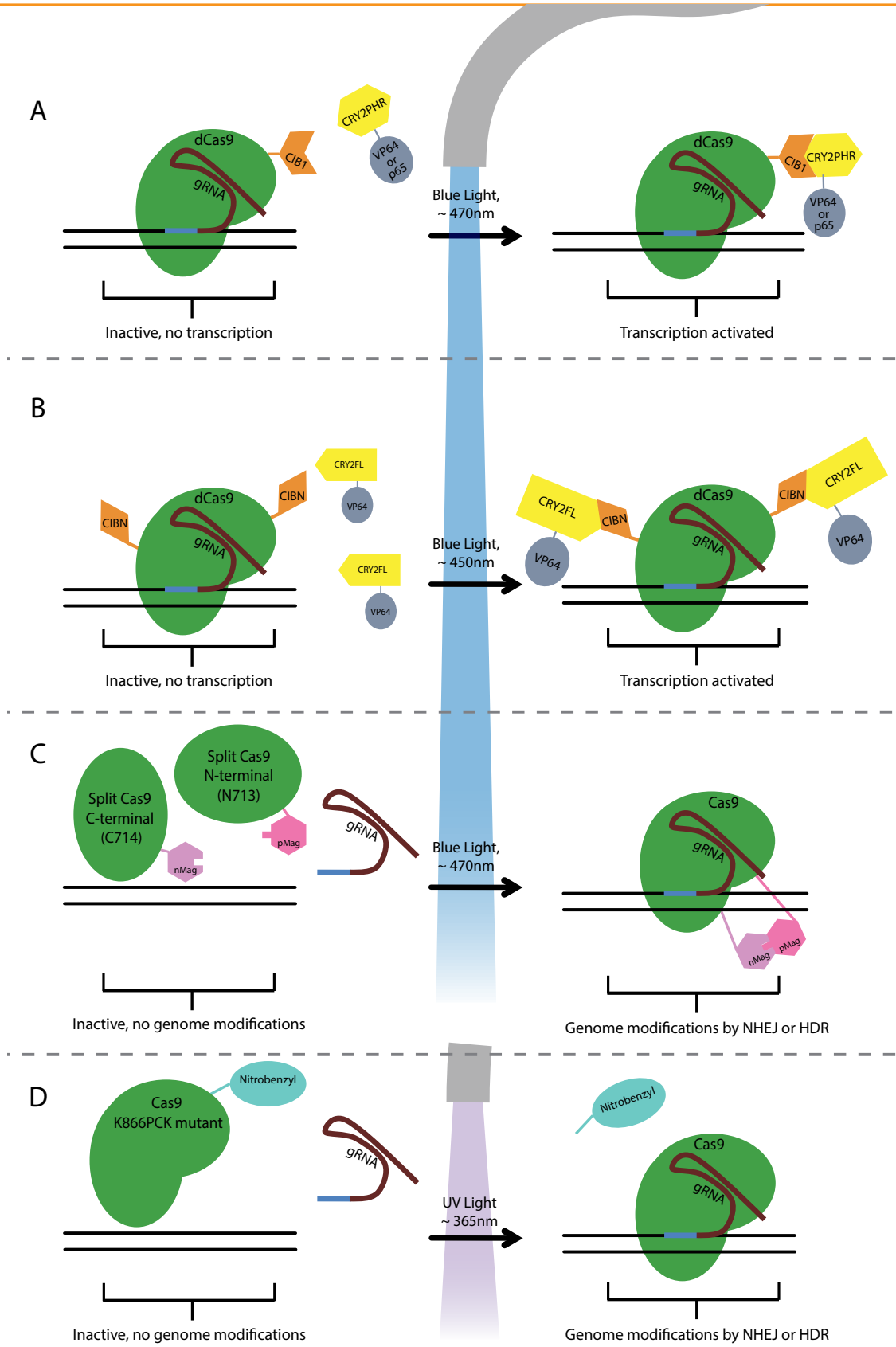
As one might expect from the modular nature of Cas9, Nihongaki et al. showed that it was possible to swap out the Cas9 domains in their split fusions and generate a photoactivatable nickase and a photoactivatable repressor (dCas9). The activity of all variants was reversible and repeatable.

Using Chemistry to Photocage CRISPR

The aforementioned techniques each employed a photoactive strategy using 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 J, et al., JACS, 2015 May 6; 137(17):5642-5). In order to insert the PCK into a specific site on 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).

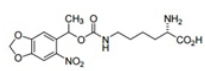
The group first tested photocaging various lysines in Cas9 to determine which best deactivated the protein's ability to cleave targeted DNA, settling on photocaging the K866 lysine, as seen in the figure below (D). Next, by using a dual reporter fluorescence assay, Hemphill et al. demonstrated that the 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.

OPTOGENETICS + CRISPR: USING LIGHT TO CONTROL GENOME EDITING (CONT'D)



OPTOGENETICS + CRISPR: USING LIGHT TO CONTROL GENOME EDITING (CONT'D)

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. For a quick run down of the applications discussed in this article, see the table below. 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!

	Publication	System Nickname	Photoactivatable Moiety	Cas9 Variant	Genome Editing Uses
A	Nihongaki Y, et al., Chem Biol, 2015 Feb 19; 22(2):169-74		CRY2 and CIB1 proteins, from A. thaliana	dCas9	Activation of gene transcription
B	Polstein LR, et al., Nat Chem Biol, 2015 Mar; 11(3): 198-200	LACE	CRY2 and CIB1 proteins, from A. thaliana	dCas9	Activation of gene transcription
C	Nihongaki Y, et al., Nat Biotech, 2015 Jul; 33(7):755-60	paCas9	Magnet proteins, from the fungal photoreceptor Vivid (N. crassa)	• Split Cas9 • Split Cas9 nickase • Split dCas9	Genome modifications by NHEJ and HDR, repression of gene expression by CRISPRi
D	Hemphill J, et al., JACS, 2015 May; 137(17):5642-5		Nitrobenzyl caged lysine 	Cas9 K866PCK mutant	Genome modifications by NHEJ and HDR

Further Reading

1. CRISPR-Cas9-based Photoactivatable Transcription System. Yuta Nihongaki, Shun Yamamoto, Fuun Kawano, Hideyuki Suzuki, Moritoshi Sato. Chemistry & Biology. 19 February 2015; 22(2):169–174. PubMed PMID: 25619936.
2. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. Lauren R Polstein & Charles A Gersbach, Nature Chemical Biology. 2015; 11:198–200. PubMed PMID: 25664691.
3. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. Yuta Nihongaki, Fuun Kawano, Takahiro Nakajima, Moritoshi Sato. Nature Biotechnology. 2015; 33:755–760. PubMed PMID: 26076431.
4. Engineered pairs of distinct photoswitches for optogenetic control of cellular proteins. Fuun Kawano, Hideyuki Suzuki, Akihiro Furuya, Moritoshi Sato. Nature Communications. 2015 Feb 24; 6:6256. PubMed PMID: 25708714.
5. Optical Control of CRISPR/Cas9 Gene Editing. Hemphill J, Borchardt EK, Brown K, Asokan A, Deiters A. Journal of the American Chemical Society. 2015 May 6; 137(17):5642-5. PubMed PMID: 25905628.
6. Genetically encoded photocontrol of protein localization in mammalian cells. Gautier A, Nguyen DP, Lusic H, An W, Deiters A, Chin JW. Journal of the American Chemical Society. 2010 Mar 31;132(12):4086-8. PubMed PMID: 20218600.

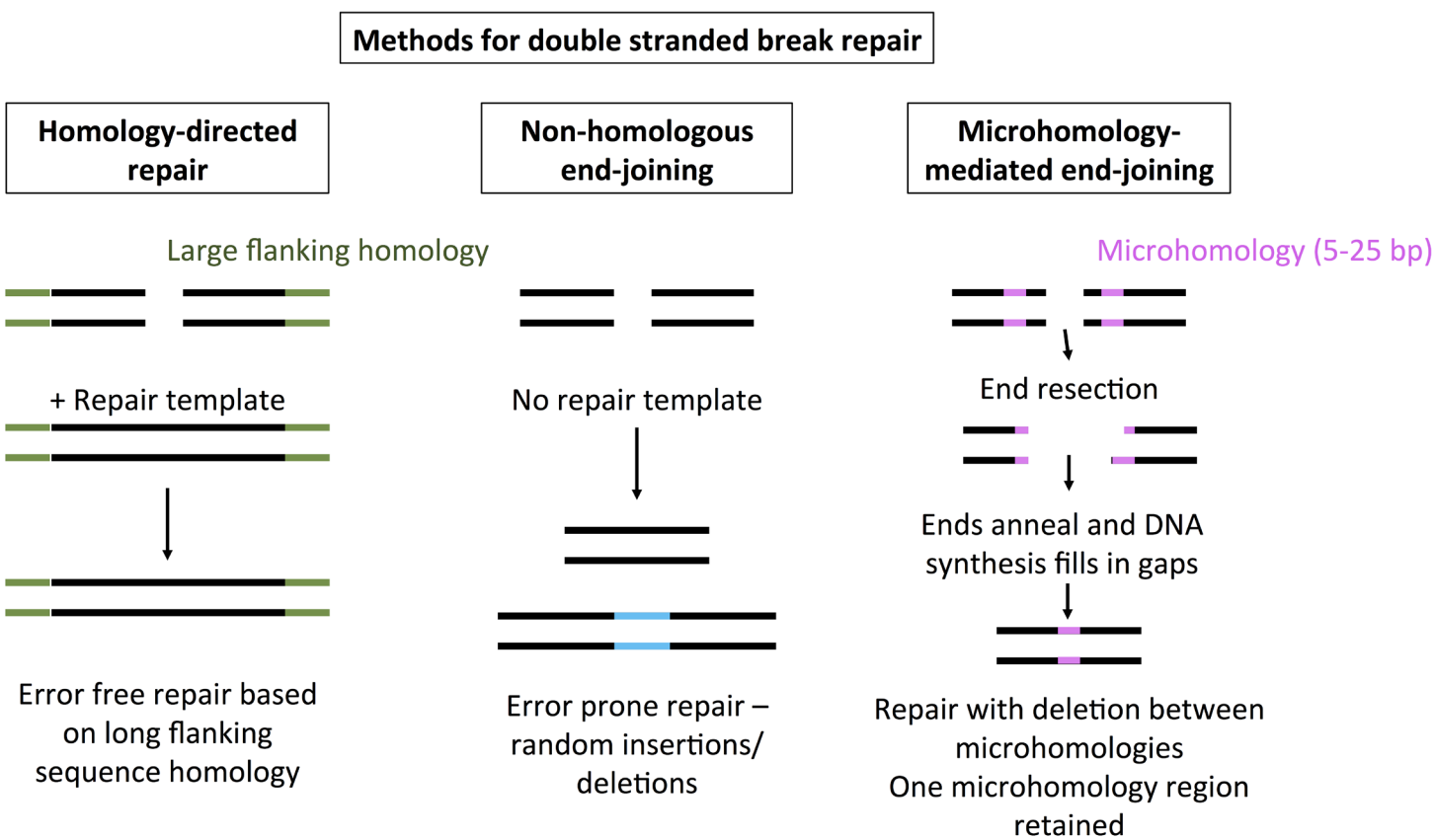
MMEJ: AN ALTERNATIVE ROUTE FOR GENE EDITING

By Mary Gearing | Februrary 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. For more mechanistic details on HDR and NHEJ, please see the linked blog posts.



MMEJ: 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 the figure above, 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 below). The double-MMEJ strategy looks very similar to HDR, but it is accomplished using much smaller regions of homology, which facilitates easier cloning.

Abbreviated PITCh Protocol

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.

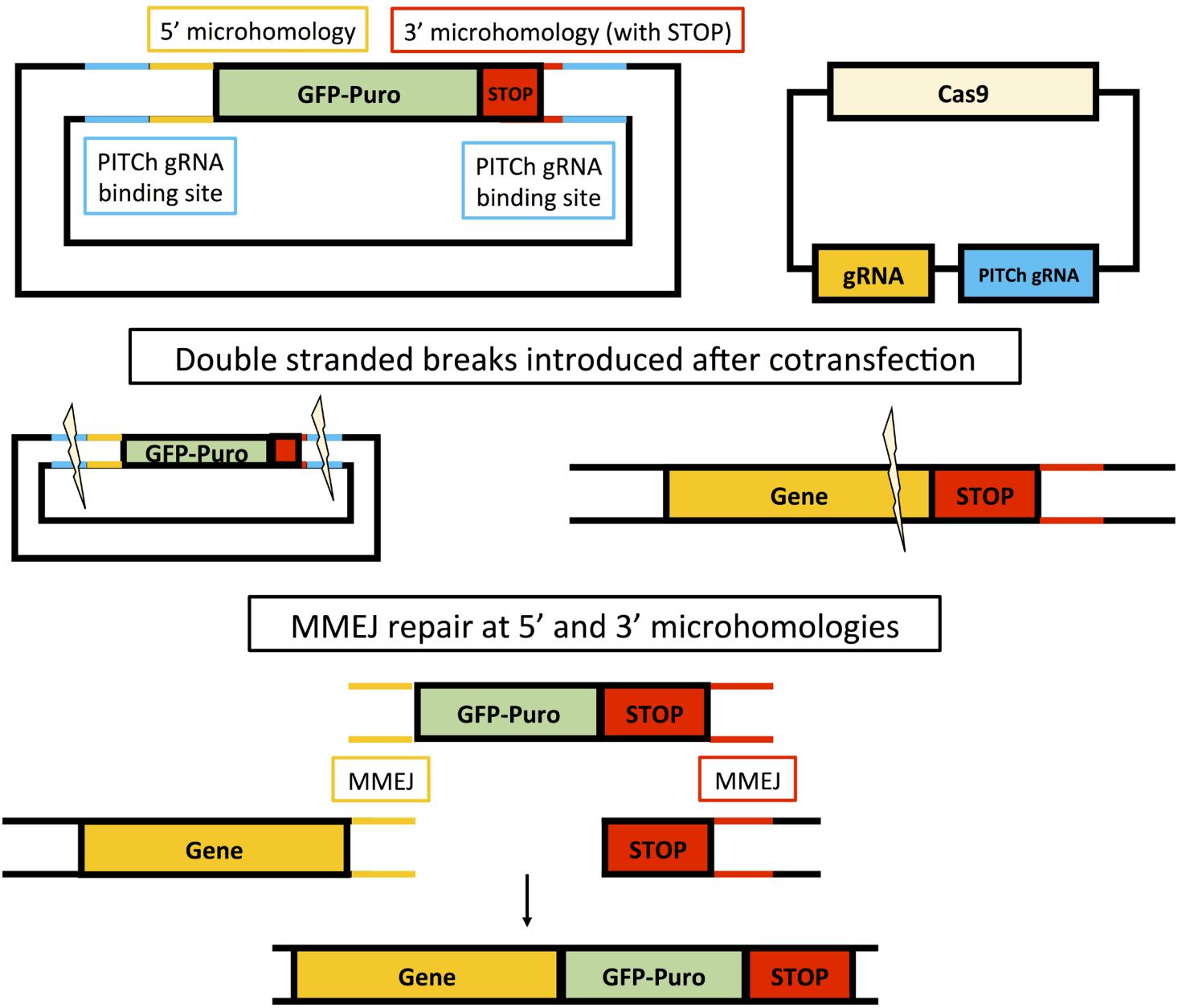
Step 3: Contransfect the PITCh vector with the vector carrying Cas9 and the PITCh- and locus-specific gRNAs

Step 4: Select for puromycin resistant cells

Step 5: PCR amplify and sequence the locus to verify correct GFP-puro insertion

MMEJ: AN ALTERNATIVE ROUTE FOR GENE EDITING (CONT'D)

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.



An overview of PITCh. The PITCh plasmid contains a GFP-Puro casette flanked by 5' and 3' microhomology and PITCh-gRNA binding sites. This PITCh plasmid is cotransfected with a plasmid carrying Cas9, the PITCh-gRNA, and the locus-specific gRNA. The PITCh-gRNA creates two DSBs in the PITCh plasmid to release the cassette, and an insertion-locus specific gRNA cleaves genomic DNA. These three double stranded breaks are repaired by two rounds of MMEJ. This double MMEJ strategy looks very similar to HDR, but it is completed using much shorter sequence homologies.

MMEJ: 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, Sakuma et al.'s 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.

Further Reading

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.

PAIRING COMBIGEM AND CRISPR FOR COMBINATORIAL GENETIC SCREENING

By Alan Wong | April 12, 2016

The complexity of biological systems can hinder our attempts to study and engineer them, but what if we had a simple tool that allowed us to rapidly decode the complexity? The CombiGEM-CRISPR technology was developed with the goal of providing an easy-to-use tool to analyze the complex combinatorial genetic networks underlying your favorite biological phenotype in a scalable way. This blog post will introduce you to this new technology, and guide you through the basics of CombiGEM-CRISPR experiments.

CombiGEM-CRISPR: The Marriage of Two Simple Platforms

CRISPR has revolutionized how we decode the genome by making it easy to create specific genetic perturbations. The ease with which one can design and synthesize CRISPR guide RNAs (gRNAs) for genome editing in large-scale has led to the rapid generation of genome-wide gRNA libraries that knock out (Doench et al., 2016; Hart et al., 2015; Koike-Yusa et al., 2014; Ma et al., 2015; Shalem et al., 2014; Wang et al., 2014), knock down (Gilbert et al., 2014), and activate (Gilbert et al., 2014; Konermann et al., 2014) individual genes for studies interrogating their functions. The continual advancements in gRNA design necessary to achieve maximal on-target and minimal off-target activities have proven themselves incredibly useful for the efficient generation of individual (and multiple) genetic perturbations in single cells. Methods to scale up multiplexed CRISPR systems for high-throughput screening are vastly useful for mapping the combinatorial genetics that underlie complex regulation in biological systems.

The CombiGEM platform provides a means to create barcoded gRNA libraries that can be used to combinatorially modify the genome, screen for a particular phenotype, and quickly profile the resultant hits (Cheng et al., 2014; Wong et al., 2015; Wong et al., 2016). CombiGEM uses iterative one-pot reactions with straightforward restriction digestion and ligation steps to build barcoded lentiviral plasmids containing one or more gRNAs. As each ligation reaction uses a pool of gRNAs as starting material, the ligation products are a diverse pool of lentiviral plasmids. Multiple rounds of digestion and ligation result in plasmid libraries containing combinations of different gRNAs in each plasmid. Each gRNA combination can be tracked and quantitatively analyzed by sequencing its set of barcodes. CombiGEM is highly flexible and can accommodate any genetic elements of interest. It can thus be tailored to address the users' specific research questions. CombiGEM has been successfully applied to functionally characterize combinatorial gene knockouts generated via multiplexed gRNA expression (Wong et al., 2016), in addition to the combinatorial expression of other genetic elements including transcription factors (Cheng et al., 2014) and microRNAs (Wong et al., 2015).

Starting Your CombiGEM-CRISPR Experiments

The first thing you will need is to get the list of effective gRNA sequences targeting your genes of interest. Thanks to the tremendous efforts made by various research teams to build and constantly improve gRNA libraries, excellent resources with effective gRNAs are publicly available, including those found at Addgene!

With a list of gRNA targeting sequences, you can then readily generate your barcoded gRNA library sequences via oligo synthesis using the format indicated below and pool-clone them into the pAWp28 storage vector (Figure 1):

Forward oligo: 5'- CACCGNNNNNNNNNNNNNNNNNNNGTTTGGGTCTTCGAGAAGACCTATTCXXXXXXXXC -3';

Reverse oligo: 5'- AATTGXXXXXXXXGAATAGGTCTTCTCGAAGACCCAAACNNNNNNNNNNNNNNNNNNNC -3',

PAIRING COMBIGEM AND CRISPR FOR COMBINATORIAL GENETIC SCREENING (CONT)

where NNN and XXX represent the 20-bp gRNA target sequence and its 8-bp barcode, respectively.

Once cloned into the storage vector, the pooled, barcoded gRNA library is now ready for CombiGEM-based assembly into the CombiGEM lentiviral vector backbone pAWp12 (Figure 2). Cloning from the pAWp28 storage vector to the pAWp12 lentiviral vector should retain the diversity in your gRNA library, but this should be verified by NGS before beginning your experiments.

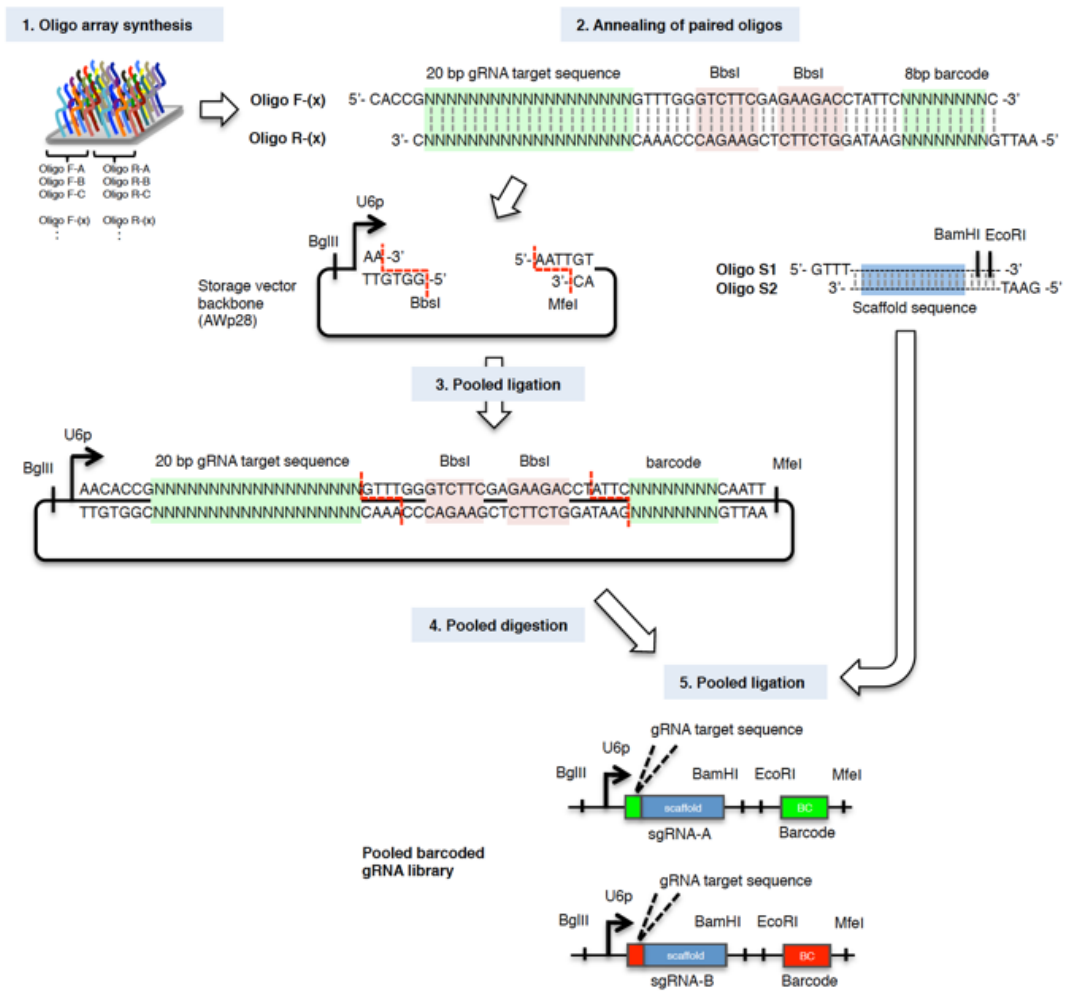


Figure 1: Strategy for assembling the barcoded gRNA library pool (Adapted from Wong et al 2016).

With CombiGEM and CRISPR platforms now being integrated, we look forward to the realization of a variety of perturbations and applications in functional genomics, cell reprogramming, and beyond. For further details, please check out Wong et al 2016 and Wong et al 2015.

PAIRING COMBIGEM AND CRISPR FOR COMBINATORIAL GENETIC SCREENING (CONT)

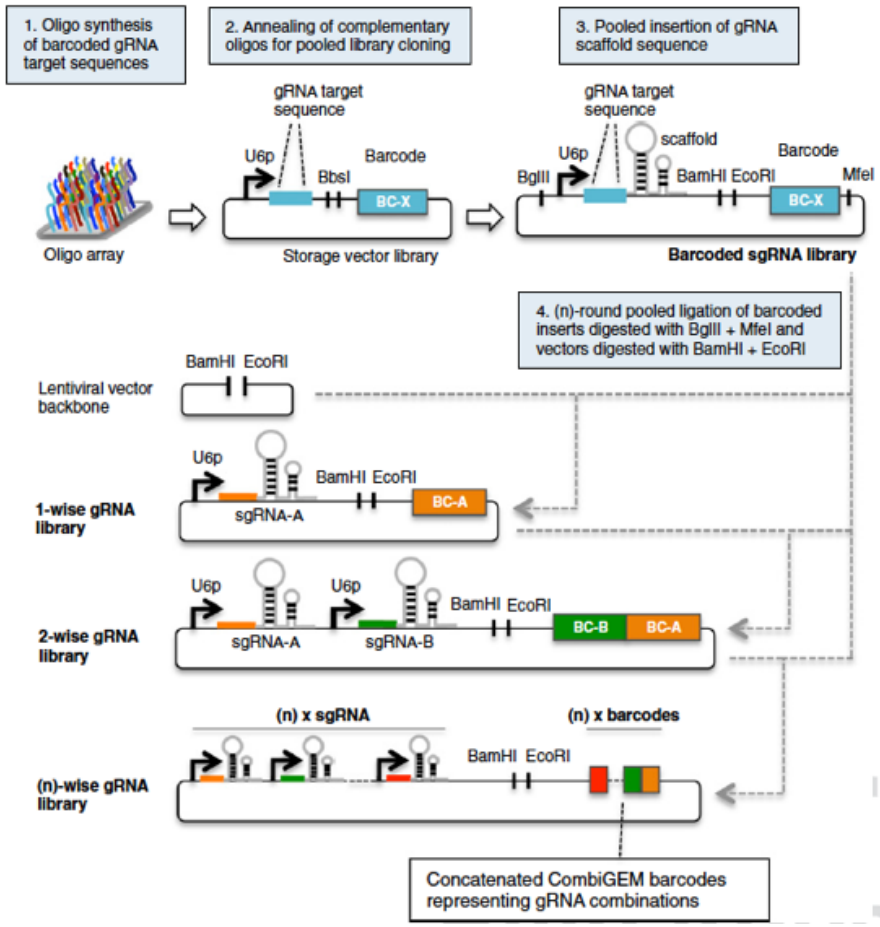


Figure 2: Strategy for assembling barcoded combinatorial gRNA libraries (adapted from Wong et al 2016).

Further Reading

1. Cheng, Allen A., Huiming Ding, and Timothy K. Lu. "Enhanced killing of antibiotic-resistant bacteria enabled by massively parallel combinatorial genetics." Proceedings of the National Academy of Sciences 111.34 (2014): 12462-12467. PubMed PMID: 25114216. PubMed Central PMCID: PMC4151723.
2. Doench, John G., et al. "Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9." Nature biotechnology (2016). PubMed PMID: 26780180. PubMed Central PMCID: PMC4744125.
3. Gilbert, Luke A., et al. "Genome-scale CRISPR-mediated control of gene repression and activation." Cell 159.3 (2014): 647-661. PubMed PMID: 25307932. PubMed Central PMCID: PMC4253859.
4. Hart, Traver, et al. "High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities." Cell 163.6 (2015): 1515-1526. PubMed PMID: 26627737.
5. Koike-Yusa, Hiroko, et al. "Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library." Nature Biotechnology 32.3 (2014): 267-273. PubMed PMID: 24535568.
6. Konermann, Silvana, et al. "Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex." Nature (2014). PubMed PMID: 25494202. PubMed Central PMCID: PMC4420636.
7. Ma, Hongming, et al. "A CRISPR-based screen identifies genes essential for West-Nile-virus-induced cell death." Cell reports 12.4 (2015): 673-683. PubMed PMID: 26190106. PubMed Central PMCID:

PAIRING COMBIGEM AND CRISPR FOR COMBINATORIAL GENETIC SCREENING (CONT)

PMC4559080.
8. Shalem, Ophir, et al. “Genome-scale CRISPR-Cas9 knockout screening in human cells.” Science 343.6166 (2014): 84-87. PubMed PMID: 24336571. PubMed Central PMCID: PMC4089965.
9. Wang, Tim, et al. “Genetic screens in human cells using the CRISPR-Cas9 system.” Science 343.6166 (2014): 80-84. PubMed PMID: 24336569. PubMed Central PMCID: PMC3972032.
10. Wong, Alan SL, et al. “Massively parallel high-order combinatorial genetics in human cells.” Nature Biotechnology (2015). PubMed PMID: 26280411. PubMed Central PMCID: PMC4785103.
11. Wong, Alan SL, et al. “Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM.” Proceedings of the National Academy of Sciences 113.9 (2016): 2544-2549. PubMed PMID: 26864203. PubMed Central PMCID: PMC4780610.

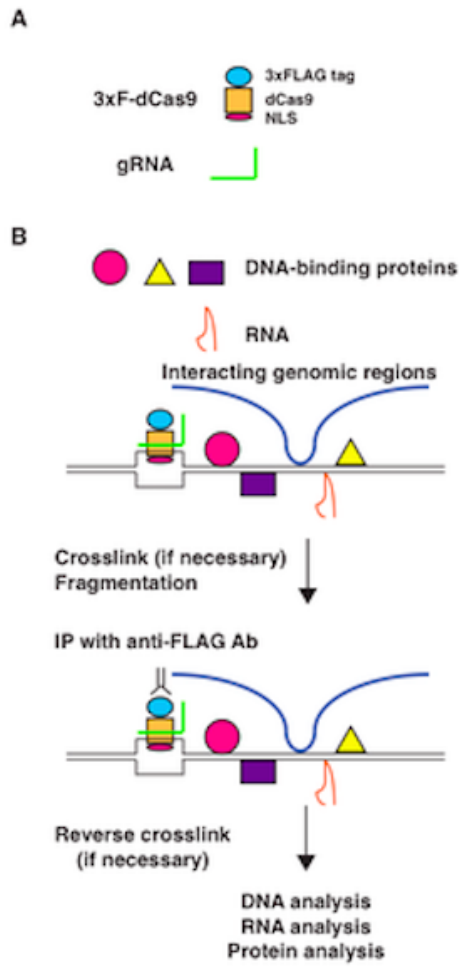
CRISPR PURIFICATION

By Mary Gearing | December, 2015

Chromatin immunoprecipitation (ChIP) has helped revolutionize our understanding of chromatin binding partners. Moving forward, researchers would like to more broadly understand which proteins and RNA bind specific genomic loci, and how these binding partners change over time or in disease states. Although methods such as iChIP and PICh have been used previously, the CRISPR-based method enChIP is advantageous due to its simplicity and multiplexing potential.

Isolating a specific genomic locus is difficult, so researchers have had to get a little creative. iChIP inserts a recognition site for a DNA-binding protein, such as LexA, next to the locus of interest, allowing one to perform ChIP using a specific antibody. iChIP works well in cell lines, but its use in model organisms requires making a transgenic organism. Another technique, PICh, uses a complementary nucleic acid probe to bind a locus, but the approach has only been used for high copy loci.

To solve these issues, Addgene Advisory Board member Hodaka Fujii's lab has created enChIP (engineered DNA-binding molecule-mediated chromatin immunoprecipitation). As you may have guessed, enChIP uses a gRNA to specify a given locus, directing FLAG-tagged dCas9 to the appropriate location. After crosslinking, the sample is immunoprecipitated using a FLAG antibody, then reverse crosslinked. DNA, RNA, and proteins bound to the locus can then be identified through a combination of next-generation sequencing, RNA-seq, and mass spectrometry (see figure to the right).



Overview of the enChIP process

To validate the specificity of gRNA targeting and the IP, you can compare the abundance of your locus of interest in the input and IP DNA using qRT-PCR. Ideally, that locus should be enriched in the IP sample, but predicted off-target (or randomly selected) loci should not. One way to minimize the effects of off-target binding is to compare two treatments (e.g. +/- insulin). Off-target binding will likely be seen in both conditions, eliminating potential false positives. Conducting enChIP with different gRNAs also minimizes the effects of off-target binding; a molecule identified across multiple experiments is likely a true positive. Fujii has found that off-target binding decreases when the mismatches are closer to the PAM site, an important consideration in gRNA design.

Application: Retroviral enChIP System

Fujita and Fujii created a retroviral enChIP system for use in difficult-to-transfect cell lines, as well as primary cells. Targeting the IRF-1 promoter region, they found that about 10% of input DNA was immunoprecipitated using the viral enChIP system. Since this promoter is stimulated by IFNγ, they wanted to see how protein binding changed with IFNγ stimulation. To compare the two conditions, they combined enChIP with SILAC

CRISPR PURIFICATION (CONT'D)

(Stable Isotope Labeling using Amino acids in Cell culture). In SILAC, cells representing one condition are grown in "Light" media; cells of the other condition are grown in "Heavy" media, which contains amino acids incorporating heavy isotopes. Heavy isotope labeling causes protein mass shifts detectable by mass spectrometry, so protein abundance in the two conditions can be directly compared.

Using SILAC and enChIP, Fujita and Fujii found many classes of proteins that bound the stimulated IRF-1 promoter, including histone deacetylases previously thought to play a role in promoter activation. To confirm these interactions, they performed ChIP using antibodies against the proteins found using enChIP. This study shows that enChIP can shed a great deal of light on transcriptional mechanisms. Fujii also sees potential applications of enChIP to tough genomic questions like epigenetic regulation, genomic imprinting, and X chromosome inactivation.

Important Considerations for enChIP gRNA Design

To help you better use enChIP, Fujita and Fujii have created a short list of points to consider when designing your gRNAs. For a more detailed guide, check out this review.

- gRNAs for use with enChIP shouldn't bind to conserved genomic regions. These regions often contain functional regulatory sequences, including recognition sites for DNA-binding molecules.

- When analyzing a promoter: The gRNA should bind several hundred base pairs upstream of the transcription start site. This location avoids interference with the recruitment of transcription factors and polymerases. Design gRNAs to bind 100-300 bases from the transciption start site and test to see if gene expression remains constant (via qRT-PCR.) If gene expression decreases, try designing gRNAs targeting other areas of the promoter sequence.

- When analyzing an enhancer/silencer: gRNAs can be designed to bind directly adjacent to the region of interest.

enChIP plasmids are available from Addgene. If you'd like to learn more about the work in Hodaka Fujii's lab, check Addgene's interview with him!

Further Reading

1. 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." Biochem Biophys Res Commun. 439(1) (2013):132-6. PubMed PMID: 23942116.

2. Fujita, Toshitsugu, and Hodaka Fujii. "Identification of Proteins Associated with an IFNγ-Responsive Promoter by a Retroviral Expression System for enChIP Using CRISPR." PLoS One 9(7) (2014) :e103084. PubMed PMID: 25051498. PubMed Central PMCID: PMC4106880.

3. Fujii, Hodaka, and Toshitsugu Fujita. "Isolation of Specific Genomic Regions and Identification of Their Associated Molecules by Engineered DNA-Binding Molecule-Mediated Chromatin Immunoprecipitation (enChIP) Using the CRISPR System and TAL Proteins." Int J Mol Sci. 16(9) (2015):21802-12. PubMed PMID: 26370991. PubMed Central PMCID: PMC4613281.

4. Fujita, Toshitsugu, and Hodaka Fujii. "Efficient isolation of specific genomic regions retaining molecular interactions by the iChIP system using recombinant exogenous DNA-binding proteins." BMC Mol Biol. 15 (2014):26. PMID: 25428274. PubMed Central PMCID: 4253623.

CRISPR PURIFICATION (CONT'D)

5. Dejardin, Jerome, and Robert E. Kingston. "Purification of proteins associated with specific genomic loci." Cell 136(1) (2009):175-186. PubMed PMID: 19135898. PubMed Central PMCID: PMC3395431.

VALIDATING YOUR GENOME EDIT

By Melina Fan | July 30, 2015

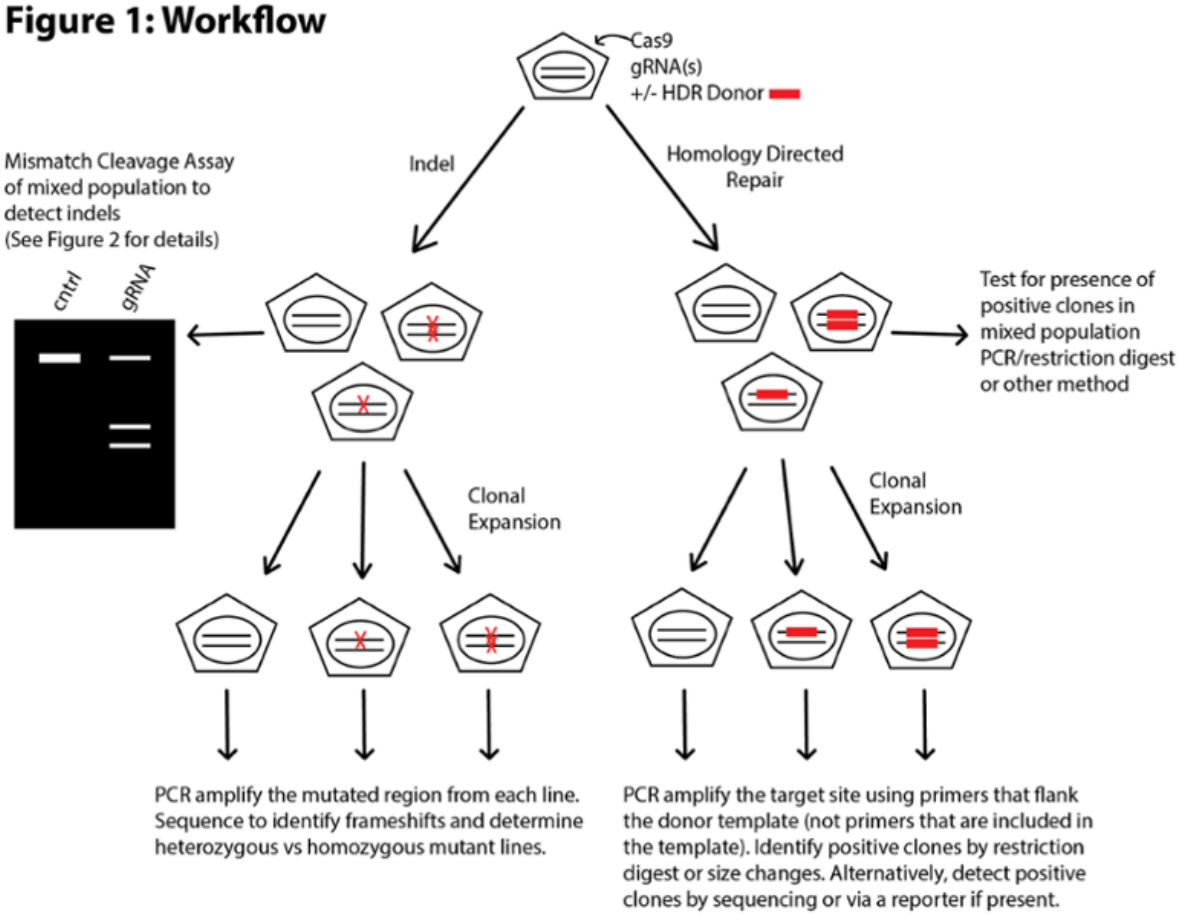
You've created your gRNA expression construct and used Cas9 to introduce it into your target cells. Hooray! You're ready to begin reading out data, right? Almost. In this section we'll explain how to verify that your cells were appropriately edited. We'll also cover the basic techniques for detecting 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 post, we will focus on diploid mammalian cells, but many of the principles will hold across different model organisms.

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) 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.

Figure 1: Workflow



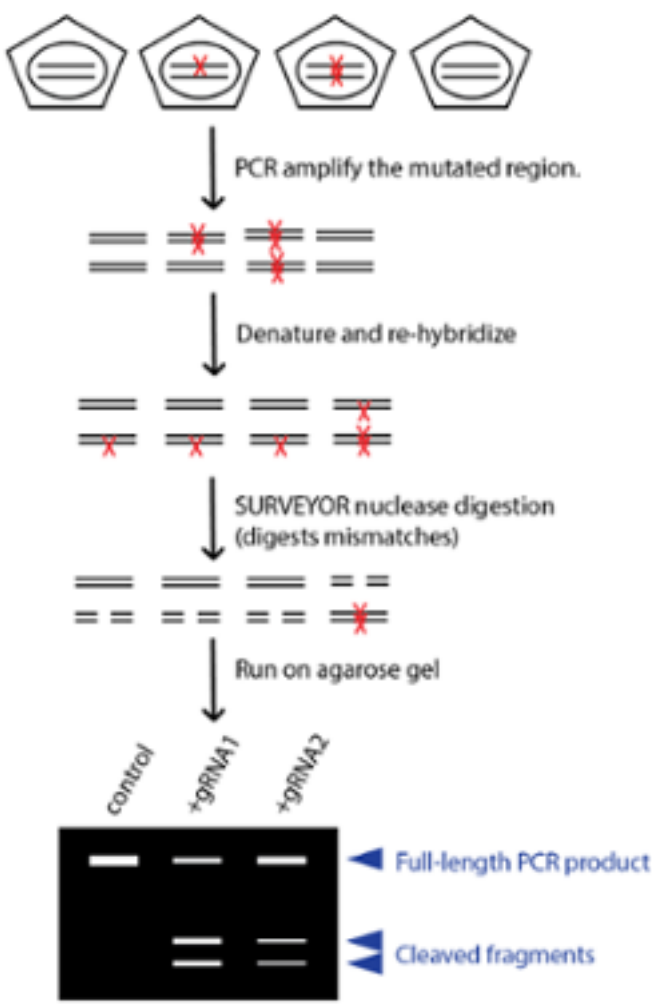
VALIDATING YOUR GENOME EDIT (CONT'D)

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. When possible, you should also assess protein expression via western blot as a further form of validation.

Mismatch Cleavage Assay to Detect Indels

Figure 2: Mismatch Cleavage Assay



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
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.

Detect Homology Directed Repair

If you want to mutate your region of interest using HDR, it is advisable to first determine whether your gRNA is efficiently cutting 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.

VALIDATING YOUR GENOME EDIT (CONT'D)

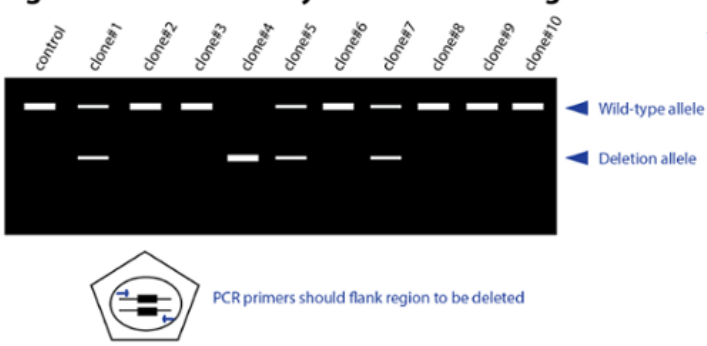
When designing your HDR donor template, plan ahead for detection of integration events. For instance, you could purposefully introduce or remove restriction sites which would alter the digestion pattern of PCR products. Alternatively, you could include a reporter element for detection of HDR at the DNA, 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.

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: PCR to Identify Clones Containing Deletions



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.

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.

Thank you to David Scott (Dr. Feng Zhang's Lab), Joel McDade (Addgene), and Marcy Patrick (Addgene) for helpful comments and edits.

VALIDATING YOUR GENOME EDIT (CONT'D)

Further Reading

1. Qiu, Peter, et al. "Mutation detection using Surveyor nuclease." Biotechniques 36.4 (2004): 702-707. PubMed PMID: 15088388.
2. Ran, F. Ann, et al. "Genome engineering using the CRISPR-Cas9 system." Nature Protocols 8.11 (2013): 2281-2308. PubMed PMID: 24157548. PubMed Central PMCID: PMC3969860.

SEQUENCING OPTIONS FOR CRISPR GENOTYPING

By Søren Hough | October 4, 2016

This post was contributed by guest blogger Søren Hough, the Head Science Writer at Desktop Genetics.

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 the Validating Your Genome Edit section, 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.

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.

SEQUENCING OPTIONS FOR CRISPR GENOTYPING (CONT'D)

Table 1: Biased Sequencing Methods

Technique	Sequence Information	Detection Limit	Advantages	Disadvantages
Mismatch Cleavage Assay	Not Provided	5%	Inexpensive, simple	Low-throughput, low sensitivity
Sanger Sequencing	Provided	20% (variable)	Sequence-level data	Low-throughput, inappropriate for heterogenous populations, low sensitivity
Targeted Amplicon Sequencing	Provided	0.01% (variable)	Sequence-level data, extremely sensitive	Do not sequence all DSBs, may miss unpredicted off-target breaks

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 0.01% (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

SEQUENCING OPTIONS FOR CRISPR GENOTYPING (CONT'D)

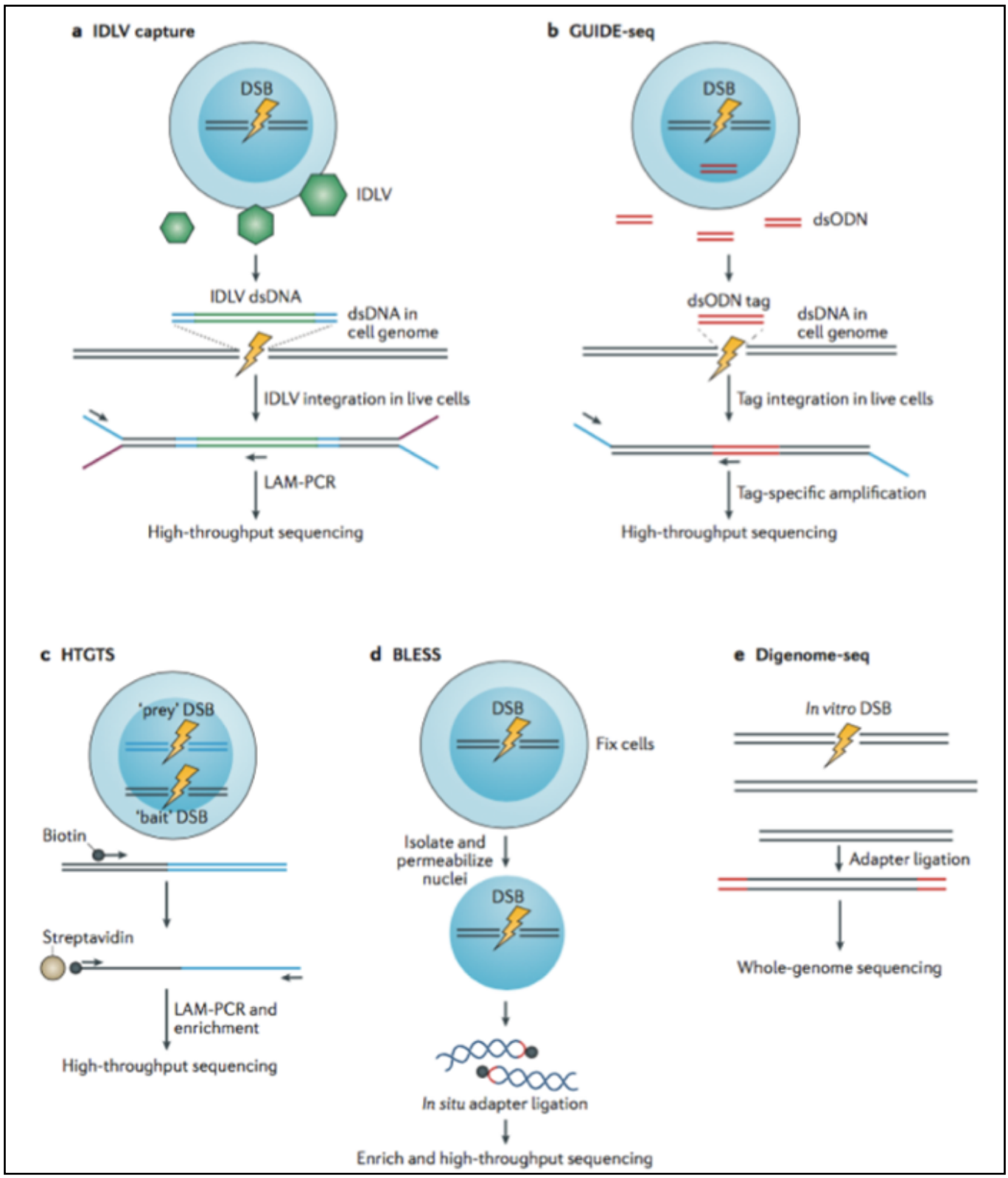


Figure 1: Schematic representations of the indicated unbiased sequencing methods. Figure reproduced from Tsai and Joung 2016 with permission.

SEQUENCING OPTIONS FOR CRISPR GENOTYPING (CONT'D)

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.

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 2: Unbiased Genotyping Options

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

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

SEQUENCING OPTIONS FOR CRISPR GENOTYPING (CONT'D)

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.

Further Reading

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. 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.
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

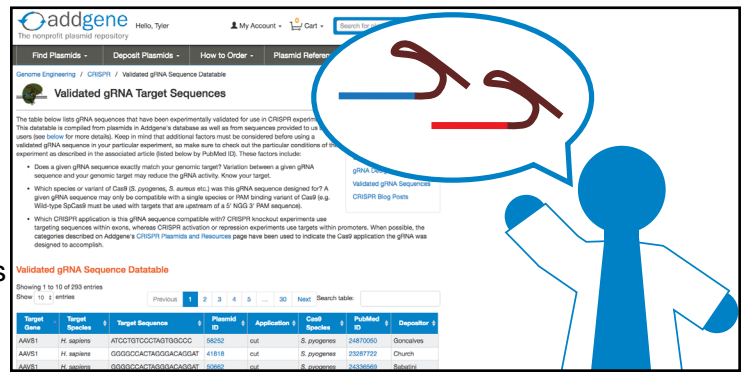
SEQUENCING OPTIONS FOR CRISPR GENOTYPING (CONT'D)

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.
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 Genetics 17.5 (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.

ADDGENE'S VALIDATED GRNA TABLE

By Nicole Waxmonsky and Brook Pyhtila | May 17, 2016

Resource sharing shortens the time needed to go from planning an experiment to performing one. At Addgene, over 120 labs have deposited CRISPR reagents, including many gRNA-containing plasmids (McDade et al, 2016). Many of the gRNAs contained within these plasmids have been used successfully in peer reviewed articles. If you're targeting your favorite gene with CRISPR, using one of these validated gRNAs can save you the time that would be spent making and testing entirely new gRNA designs. You can now easily find many validated gRNAs in our newly curated Validated gRNA Target Sequence Table.



What are Validated gRNAs?

Proper target selection and gRNA design are essential for a CRISPR experiment. The selection process is often time consuming and there is no guarantee that the selected sequence will yield a useful gRNA. Alternatively, you can select a target sequence which has already been shown to work in CRISPR experiments.

Addgene's dedication to resource sharing has led us to develop a searchable and sortable datatable that contains validated gRNA sequences. Validated in this context refers to the fact that every gRNA listed in the table has been used effectively in a published article. This blog post will serve as a guide to use our validated gRNA Target Sequence Table.

> *Disclaimer: The efficacy of gRNAs is impacted by the target location, target sequence, and model system. Even if a gRNA has been demonstrated to work in your system, it is worth spending time sequencing the genomic target to determine whether any sequence variations (e.g. SNPs) exist. Additionally, consider testing multiple gRNAs targeted to the same locus to ensure their effects are specific to that locus.*

The data in the validated gRNA table has been derived from scientists' submissions, either their deposited gRNA plasmids or from submission of a spreadsheet containing the validated gRNA sequences alone. Sequence submission is welcomed for validated gRNAs that are not generated from a plasmid. If you, for example, prefer to deliver your CRISPR components as RNPs (ribonucleoproteins) and therefore usually transcribe your gRNAs *in vitro*, you can send us the sequences of your effective gRNAs, and we'll add them to this list.

Access the spreadsheet for submitting your gRNA sequences here.

Features of the Validated gRNA Table

You can access the datatable here. Current table columns are target gene, species, sequence, Addgene plasmid ID (if we have one associated with the gRNA), application, cas9 species, PubMed ID, and depositor. The CRISPR applications we currently have data for are: cut, activate, interfere, visualize, nick, purify, tag, scaffold, CRISPR-display, and dCas9-FokI.

ADDGENE'S VALIDATED GRNA TABLE (CONT'D)

Validated gRNA Sequence Datatable

Showing 1 to 10 of 293 entries
Show 10 entries Previous 1 2 3 4 5 … 30 Next Search table:

Target Gene	Target Species	Target Sequence	Plasmid ID	Application	Cas9 Species	PubMed ID	Depositor
AAVS1	*H. sapiens*	ATCCTGTCCCTAGTGGCCC	58252	cut	*S. pyogenes*	24870050	Goncalves
AAVS1	*H. sapiens*	GGGGCCACTAGGGACAGGAT	41818	cut	*S. pyogenes*	23287722	Church
AAVS1	*H. sapiens*	GGGGCCACTAGGGACAGGAT	50662	cut	*S. pyogenes*	24336569	Sabatini
AAVS1	*H. sapiens*	GGGGCCACTAGGGACAGGAT	70661	cut	*S. pyogenes*	26472758	Sabatini
AAVS1	*H. sapiens*	GTCCCCTCCACCCCACAGTG	41817	cut	*S. pyogenes*	23287722	Church
AAVS1	*H. sapiens*	GGGGCCACTAGGGACAGGAT	72833	cut	*S. pyogenes*	27052166	Kanemaki
AAVS1	*H. sapiens*	TGTCCCTAGTGGCCCCACTG		cut	*S. pyogenes*	26789497	Corn
AAVS1	*H. sapiens*	ACAGTGGGGCCACTAGGGAC		cut	*S. pyogenes*	26789497	Corn
ACTB	*H. sapiens*	UGGAGCGAGCAUCCCCCAAA	74705	RNA targeting	*S. pyogenes*	26997482	Yeo
actII-orf4	*S. coelicolor*	ATTACCAGGGACCGGAGTTC	62552	cut	*S. pyogenes*	25739462	Jiang

This table is both sortable and searchable, which can help exclude the data that you are not interested in. You can use the carats in the datatable headers to sort alphabetically or in ascending or descending order, in the case of numeric columns.

The data in any column are searchable and you can search by any part of the data. For example, entering 'C. elegans', or 'elegans' will filter the table to show *C. elegans* gRNA sequences. You can also filter by multiple keywords that are separated by a single space to get even more specific. For example, to show the gRNAs that were used for activation in a human system, search for 'sapiens activate' to filter the table and find exactly what you're looking for.

Validated gRNA Sequence Datatable

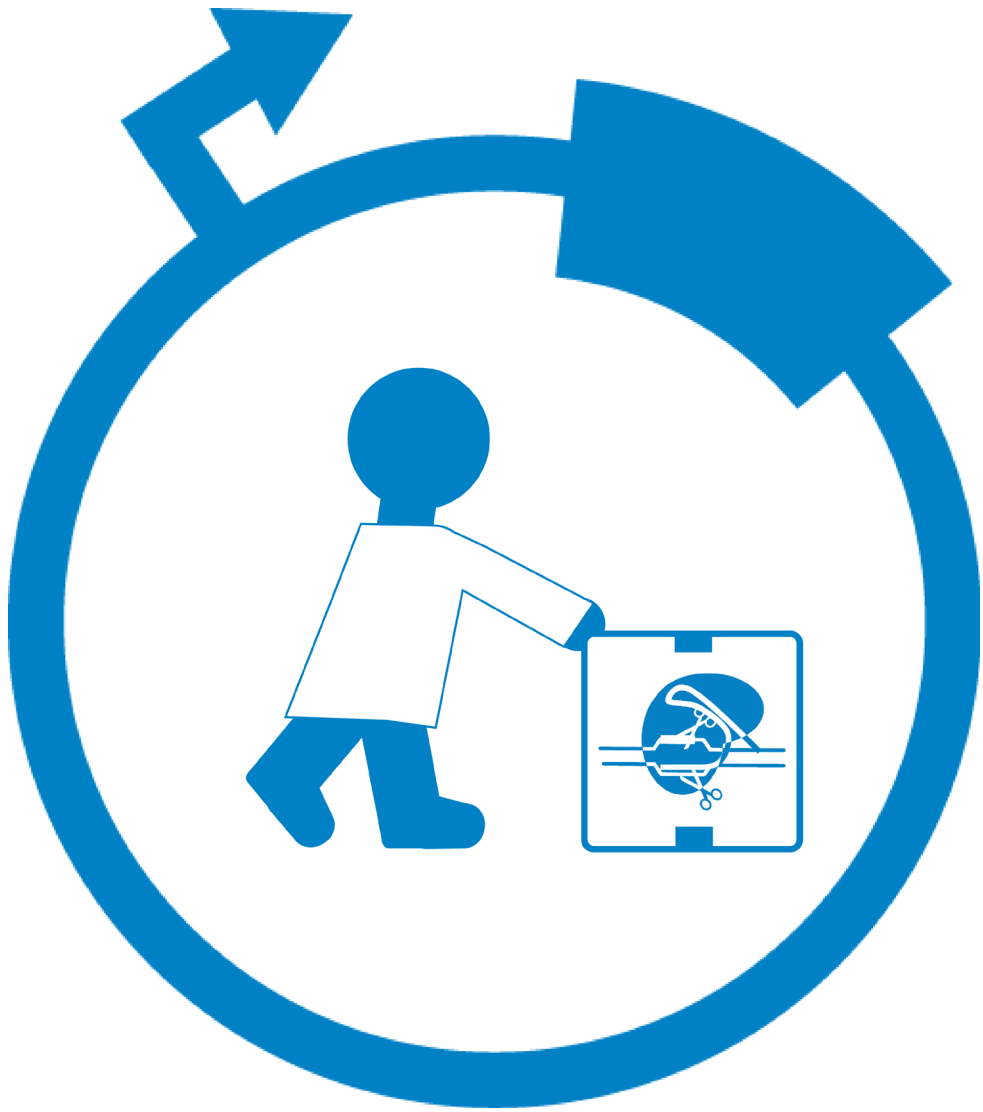
Showing 1 to 10 of 30 entries (filtered from 293 total entries)
Show 10 entries Previous 1 2 3 Next Search table: elegans

Target Gene	Target Species	Target Sequence	Plasmid ID	Application	Cas9 Species	PubMed ID	Depositor
avr-14	*C. elegans*	GATTGGAGAGTTAGACCACG	58981	cut	*S. pyogenes*	24879462	Mello
avr-15	*C. elegans*	GTTTGCAATATAAGTCACCC	58982	cut	*S. pyogenes*	24879462	Mello
dpy-10	*C. elegans*	CGCTACCATAGGCACCACG	71516	cut	*VRER Cas9 mutant*	26680661	Fire
dpy-10	*C.elegans*	GCTACCATAGGCACCACGAG	65630	cut	*S. pyogenes*	26044730	Katic
dpy-10	*C. elegans*	GCTACCATAGGCACCACGAG	59933	cut	*S. pyogenes*	25161212	Fire
dpy-10	*C. elegans*	TCCGCTACCATAGGCACCA	71479	cut	*VQR Cas9 variant*	26680661	Fire
dpy-10	*C. elegans*	GCTACCATAGGCACCACGAG	70047	cut	*S. pyogenes*	26187122	Seydoux
fbf-2	*C. elegans*	GTAGTCACGGCGATGATTA	65597	cut	*S. pyogenes*	25249454	Seydoux
fbf-2	*C. elegans*	TAATCATCGCCGTGACTAC	65591	cut	*S. pyogenes*	25249454	Seydoux
K08F4.2	*C. elegans*	AATCACTCCCTGTTTGTGT	66085	cut	*S. pyogenes*	25249454	Seydoux

Search using the box highlighted in red above. When searching, it is worth trying alternate names, particularly for genes.

If you find a sequence that will work for your experiment, it is recommended that you confirm the details and final experimental outcome in the original publication. As you develop and confirm new gRNAs, please consider submitting their sequences (and plasmids!) so that this shared resource can continue to grow.

CRISPR EXPRESSION SYSTEMS AND DELIVERY METHODS



MULTIPLEXIBLE CRISPR EXPRESSION SYSTEMS

By Mary Gearing | January, 2016

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. 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!

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...

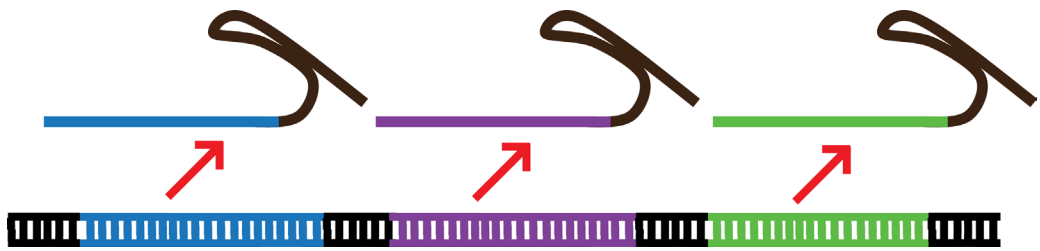


Figure 1: Multiplexing allows researchers to express multiple gRNAs from a single construct. DNA and gRNA sizes are not to scale.

Multiplexing Basics

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 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.

Golden Gate Assembly Methods

If you want to scale up to 3-7 gRNAs, Addgene has two Golden Gate Assembly-based mammalian systems from the Gersbach and Yamamoto labs. Golden Gate assembly methods are also available for plants (Qi-Jun Chen and Yao-Guang Liu labs) and zebrafish (Wenbiao Chen lab).

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.

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

MULTIPLEXIBLE CRISPR EXPRESSION SYSTEMS (CONT'D)

together to allow for ordered assembly into a destination vector that expresses Cas9. If you're confused, check out the schematic below.

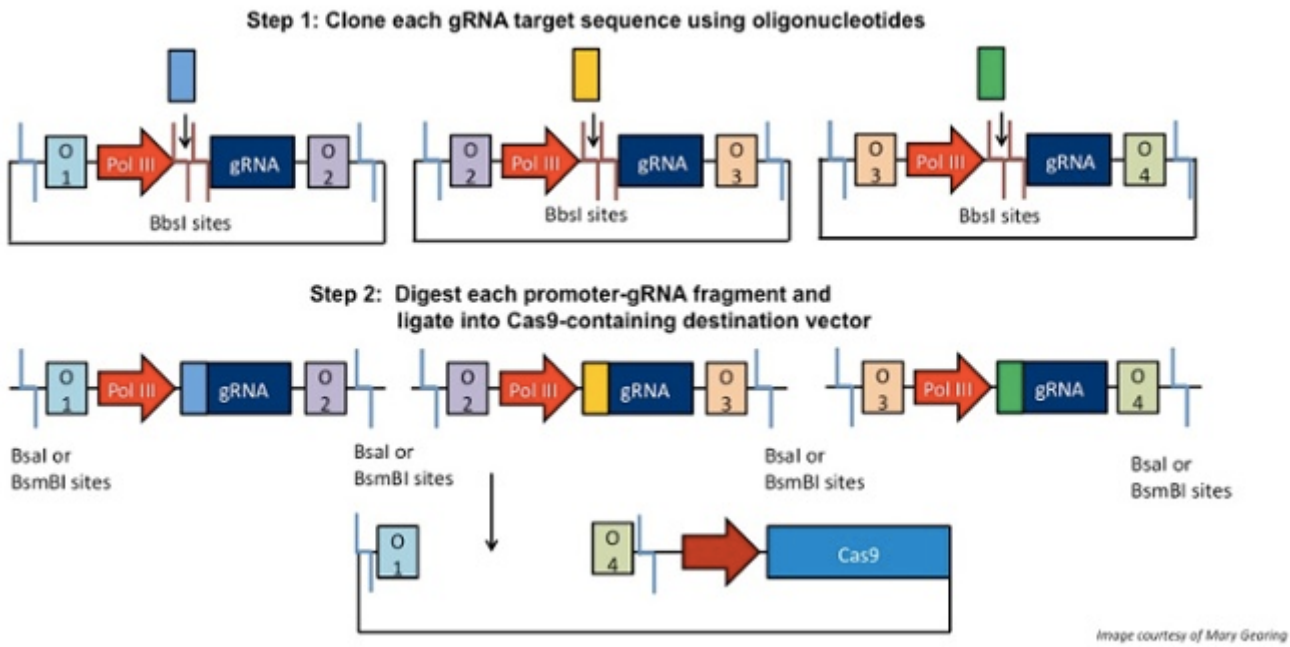


Figure 2: gRNA target sequences (colored rectangles) are cloned into various plasmids using oligonucleotides. These plasmids contain Type IIS restriction sites that flank the promoter-gRNA constructs. When these plasmids are digested, unique overhangs (here, O1-4) adjacent to the cut sites "link" fragments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

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 one of four Cas9 or dCas9-containing ampicillin-resistant destination vectors. Destination vector options include humanized wt Cas9, dCas9 (transcriptional repressor), and dCas9-VP64 (transcriptional activator)-containing plasmids. If you're interested in combining CRISPR/Cas9 with FACS-based screening, these plasmids are for you. Each destination vector contains GFP, enabling you to select cells with high GFP expression and therefore high multiplicity of infection (MOI). These high MOI 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 like with the Gersbach plasmids; the correct flanking regions are specified by the destination vector. 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

MULTIPLEXIBLE CRISPR EXPRESSION SYSTEMS (CONT'D)

cloned into the Cas9-containing, ampillicin-resistant destination vector. These constructs are digested using BsaI and assembled to produce a plasmid encoding the gRNAs, Cas9, and ampicillin resistance. As with the Gersbach lab plasmids, multiple Cas9 variants are available. The standard kit includes wt humanized Cas9 and the D10A nickase mutant (Cas9n). The accessory pack adds two more options: dCas9 (transcriptional repression) and Fok1-dCas9 (dimeric nuclease.)

Gateway Assembly Method

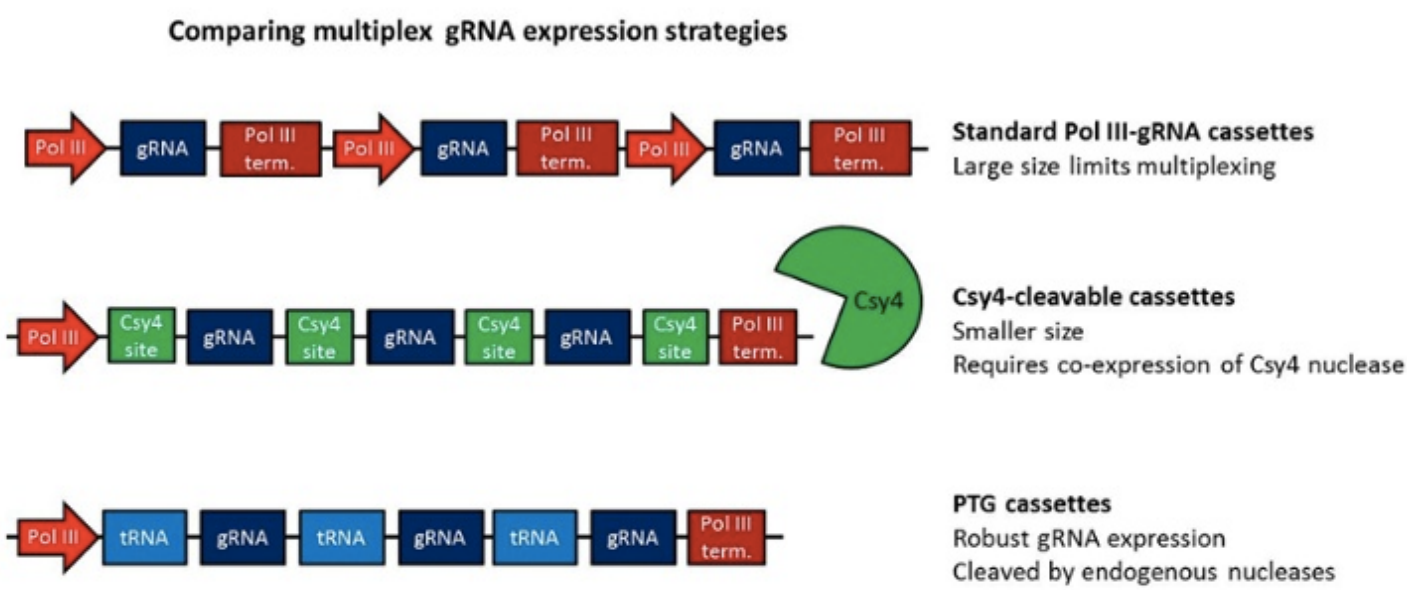
Frew Lab Multiple Lentiviral Expression Systems (MuLE) Kit

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.

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 sites. If



MULTIPLEXIBLE CRISPR EXPRESSION SYSTEMS (CONT'D)

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.

Multiplexing in Plants

© Marie-Lan Nguyen / Wikimedia Commons / CC-BY 2.5

Qi-Jun Chen Lab Golden Gate/Gibson Assembly Multiplexing Plasmids

These plasmids allow you to assemble 2-4 gRNAs through Golden Gate or Gibson Assembly. gRNAs are inserted into the pCBC vectors using BsaI, and promoter-gRNA fragments are PCR amplified for cloning into one of three *Zea mays* codon-optimized Cas9-containing binary vectors. These vectors are based off of pGreen, pSoup, and pCAMBIA, three of the most popular plant vectors that together are suitable for a wide variety of applications. The Chen lab system is compatible with both monocot and dicot plants.

Liu Lab Golden Gate/Gibson Assembly Multiplexing Plasmids

These plasmids can be used to successfully express up to 8 gRNAs after Golden Gate or Gibson Assembly! Using BsaI, gRNAs are cloned into one of 12 pYLsGRNA plasmids, which contain various promoters and reporters, and subsequently inserted into a Cas9-containing destination vector based off of pCAMBIA. wt Cas9 is plant-optimized with a high 5' GC-content, and plasmids are available for both monocot and dicot plants, with a choice of either hygromycin or Basta selection.

Yang Lab Single Transcript Multiplexing Plasmids

These plasmids are similar to the Csy4 polycistronic system described above, except that they use an endogenous nuclease system to cleave the gRNAs. gRNAs are flanked by glycine tRNAs to create polycistronic glycine tRNA-gRNA (PTG) constructs. Eukaryotic RNases P and Z recognize the tRNA sequences, cleave them, and release the gRNAs. PTGs are assembled into a wt Cas9-containing vector using Golden Gate assembly, and up to 8 gRNAs may be expressed simultaneously. Vectors for both transient expression and Agrobacterium-mediated transformation are available. For more information on this system, check out this blog post.

Multiplexing in Zebrafish

Image Attribution: Pogrebnoj-Alexandroff / Wikimedia Commons / CC-BY 3.0

Wenbiao Chen Lab Golden Gate Assembly Multiplex Plasmids

These plasmids allow expression of 2-5 gRNAs in zebrafish. As in the Yamamoto system, custom destination vectors are used depending upon the total number of gRNAs you wish to clone, so you don't have to clone any filler sequences. You also have the option of including a previously-designed tyr gRNA, which causes hypopigmentation, thus marking cells that have undergone genomic modification. In this system, Cas9 must be supplied on a separate plasmid.

MULTIPLEXIBLE CRISPR EXPRESSION SYSTEMS (CONT'D)

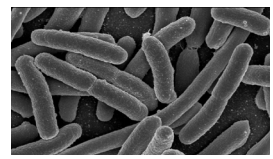
Multiplexing in *Drosophila*

Image Attribution: Aka / Wikimedia Commons / CC-BY-SA 2.5

Bullock Lab Multiplex Plasmid

Two gRNAs can be assembled using Gibson Assembly or SLIC cloning. gRNAs are expressed from two Drosophila U6 promoters. Cas9 must be supplied on a separate plasmid.

Multiplexing in *E. coli*



Koffas Lab CRISPathBrick Multiplex Plasmid

This system allows you to assemble type II-A CRISPR arrays for dCas9-based transcriptional repression. The CRISPathBrick plasmid contains a nontargeting spacer flanked by two CRISPR repeats. The spacer can be digested using BsaI, allowing a spacer-repeat "brick" to be inserted. The BsaI site remains intact, allowing subsequent "bricks" to be added one by one. This approach is especially useful for combinatorial analyses. For example, if you were to develop an array using 3 distinct spacer-repeats (more are possible), you could easily create 7 unique arrays (e.g. for spacers A, B, and C, you could obtain arrays A, B, C, AB, AC, BC, and ABC).

Check out a curated, up-to-date list of our gRNA multiplexing vectors on our gRNA page!

Further Reading

1. 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.
2. 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.
3. 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.
4. 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 PMCID: PMC4396471.
5. 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.
6. Xing, Hui-Li, et al. "A CRISPR/Cas9 toolkit for multiplex genome editing in plants." BMC Plant Biology 14 (2014): 327. PubMed PMID: 25432517. PubMed Central PMCID: PMC4262988.
7. Ma, Xingliang, et al. "A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants." Molecular Plant 8(8) (2015): 1274-1284. PubMed PMID: 25917172.
8. Xie, Kabin, Minkenberg, Bastian, & Yinong Yang. "Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system." Proceedings of the National Academy of Sciences USA 112(11) (2015): 3570-5. PubMed PMID: 25733849. PubMed Central PMCID: PMC4371917.
9. Yin, Linlin, et al. "Multiplex Conditional Mutagenesis Using Transgenic Expression of Cas9 and sgRNAs."

MULTIPLEXIBLE CRISPR EXPRESSION SYSTEMS (CONT'D)

Genetics 200(2) (2015): 431-441. PubMed PMID: 25855067.

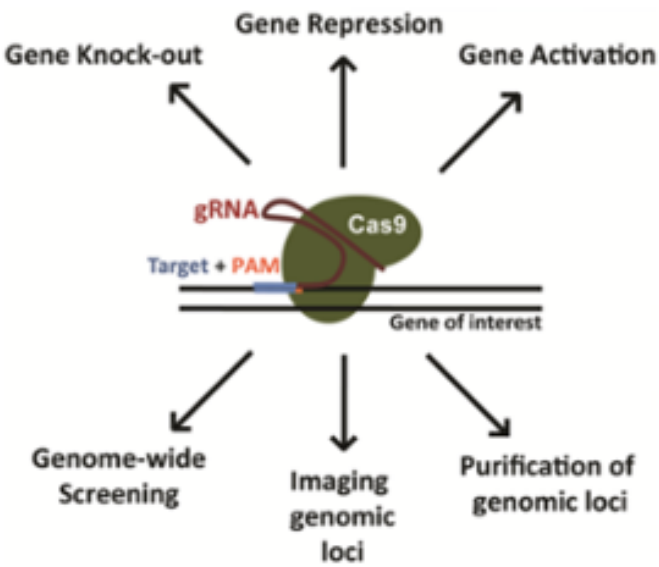
10. Port, Fillip, et al. "Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila." Proceedings of the National Academy of Sciences USA 111(29) (2014): E2967-2976. PubMed PMID: 25002478. PubMed Central PMCID: PMC4115528.

11. 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.

MAMMALIAN EXPRESSION SYSTEMS AND DELIVERY METHODS

By Nicole Waxmonsky | September 24, 2015

CRISPR technology has been widely adopted for genome editing purposes for numerous reasons including that it's cheaper, faster, and easier than prior editing techniques. With more and more CRISPR tools being published each month, you may be considering using CRISPR for your next experiment. In this section we'll provide an overview of some CRISPR mammalian expression systems, the typical applications for each, and potential delivery methods.



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 the type of outcome you are trying to achieve: Do you want to permanently knock-out or knock-in a gene? Do you want to enhance 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 the Cas9 nickase are appropriate for knocking-in, knocking-out, or introducing mutations and tags, while a "dead" or dCas9 can be used in conjunction with activator or repressor domains to control gene expression.

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 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 the specific application you have chosen. When thinking about step 3, however, you may be surprised at the number of options available--how do you choose?

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 the CRISPR components should be in based on your model system. CRISPR reagents can be delivered via transfection, nucleofection, viral infection, 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 the CRISPR components into your target cells.

MAMMALIAN EXPRESSION SYSTEMS AND DELIVERY METHODS (CONT'D)

Mammalian CRISPR Expression Systems

Each model system will have its best practices for efficient delivery of CRISPR components. If you are new to your model system, a good first step would be to consult the literature to see if anyone has published work with 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.

Expression System	Components 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 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. 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/Cas9.
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, then delivered to target cells (e.g. 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 riboprotein 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.

MAMMALIAN EXPRESSION SYSTEMS AND DELIVERY METHODS (CONT'D)

Delivery Methods for Mammalian Cell Lines

As mentioned above, the expression system you choose in many ways dictates the best method for introducing CRISPR components into your target cells. DNA delivery into mammalian cell lines is quite broad and includes several different methods, so we've further broken down this category by common cell types and recommended delivery methods below:

Method	Transformed cell lines (HeLa, HEK 293)	Stem cells (hES, iPS)	Primary cells (fibroblasts, epithelial cells)
Transfection • Lipid-Mediated • Cationic polymers • Calcium Phosphate	+		
Electroporation • Nucleofection	+	+	
Viral Delivery • Lentivirus • Retrovirus • Adenovirus • AAV	+		+

This table is not inclusive of all methods, nor are these methods limited exclusively to in vitro *cell culture. The user should review the current literature about their preferred model.*

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.

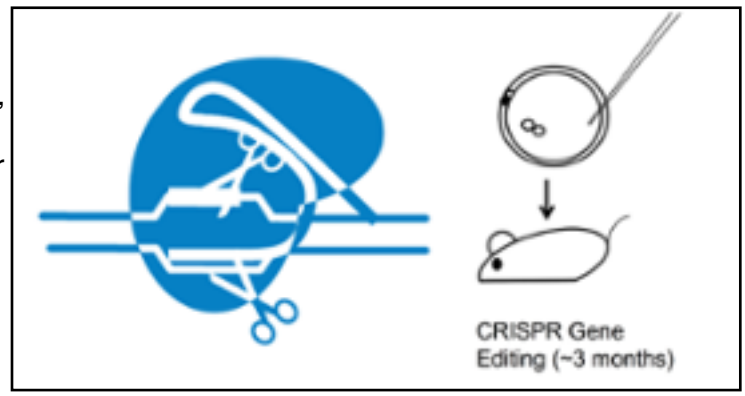
EGFP validation of sgRNAs: Can evaluate component delivery and sgRNA efficacy by cloning in genome target sequence into EGFP reporter.

Target DNA reporter system: Can evaluate component delivery with validated tools.

Special thanks to Joel McDade for creating the Expression Systems table. Marcy Patrick contributed to the writing of this section.

GENERATING MOUSE MODELS USING CRISPR/CAS9

By Wenning Qin and Haoyi Wang | July 12, 2016



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 x 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.

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, straight forward to use, and robust in performance. In its natural setting, CRISPR-Cas9 is an acquired immune system in bacteria and 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* (SP). For editing the genome, this 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).

GENERATING MOUSE MODELS USING CRISPR/CAS9 (CONT'D)

CRISPR Mouse Model Basics

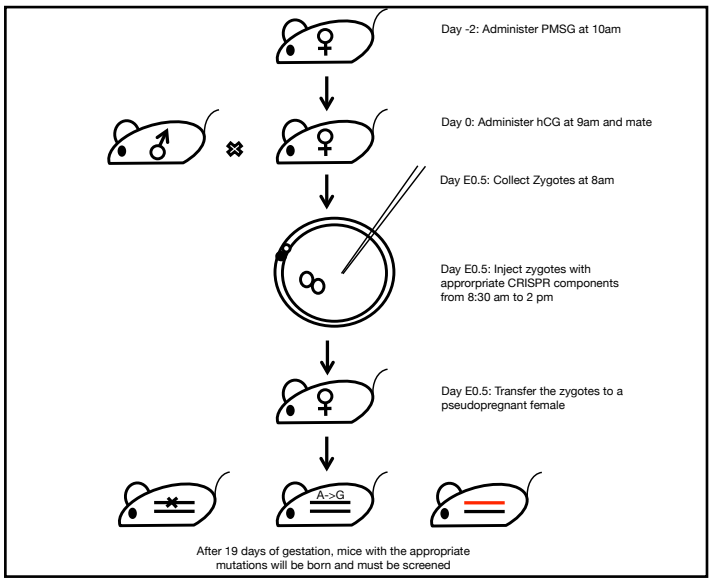


Figure 1: Process for generating genome-edited mouse lines using CRISPR/Cas9. Purifed CRISPR/Cas9 components (Cas9, gRNA, and repair template as necessary) can be directly injected into mouse zygotes and the zygotes implanted into a surrogate mother. Offspring are then screened for the desired mutations. Figure adapted from Qin et al., 2016.

To create a mouse model, the gRNA, Cas9, and donor oligonucleotide or plasmid components are brought together and injected into either the pronucleus or the cytoplasm of fertilized mouse eggs. Alternatively, the gRNA, Cas9, and donor oligonucleotide can also 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 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 CRISPR in Mice

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 the 3 major advantages of CRISPR mouse editing when compared to more conventional methods. First, one can work with almost any strains 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. Lastly,

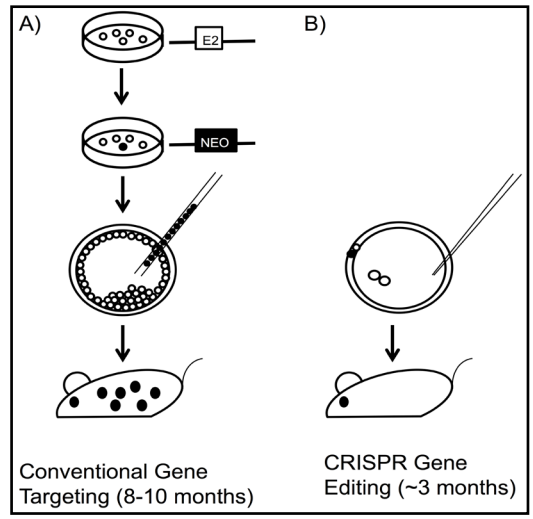


Figure 2: In conventional gene editing (A), you must first generate your desired mutation in mouse ES cells, select for the mutation, and inject the ES cells into a mouse embryo. This laborious process takes longer (~8-10 months) than CRISPR gene editing in mice (B). With CRISPR, you can inject CRISPR components (see Figure 1) directly into a zygote and get your desired mouse with the appropriate edits in much less time (~3 months). Image courtesy of Wenning Qin.

GENERATING MOUSE MODELS USING CRISPR/CAS9 (CONT'D)

it 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.

Advantages of CRISPR in Mice

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. In our hands, experiments incorporating fragments larger than 5 kb can be more challenging. Also when working with CRISPR, one must be aware that not all gRNAs are created equal. Some work better than others. There are a variety of resources online to guide you in gRNA design. We have been using Benchling, but there are many other gRNA design tools available and you can get some additional advice from John Doench's gRNA design section. 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!

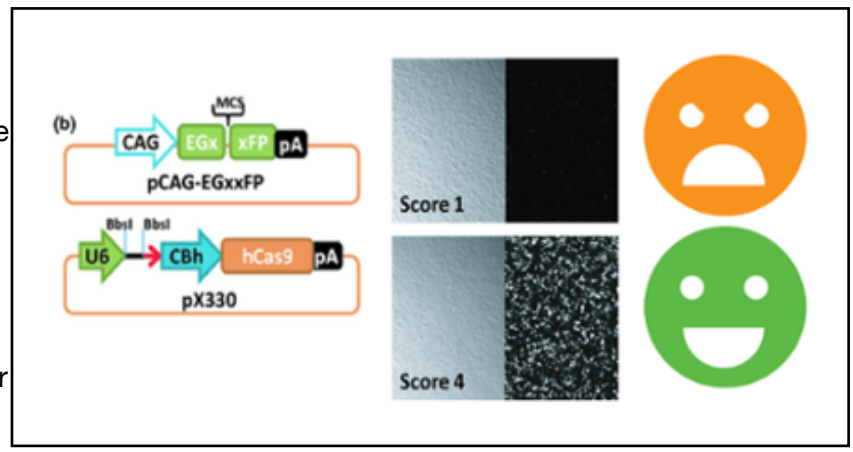
Last but not least, when using CRISPR, remember to appreciate the fact that it was first discovered in an obscure bacterium and that we still have much to learn from biology in all its forms!

Further Reading

1. Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E. 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32:1262-1267. PubMed PMID: 25184501. PubMed Central PMCID: PMC4262738.
2. Cong L., Ran F.A., Cox D., Lin S., Barretto R., Habib N., Hsu P.D., Wu X., Jiang W., Marraffini L.A., Zhang F. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013 339(6121):819-23. PubMed PMID: 23287718. PubMed Central PMCID: PMC3795411.
3. Yang, H., Wang, H., Shivalila. C.S., Cheng, A.W., Shi, L., Jaenisch, R. 2013. One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering. Cell 2013 154:1370-1379. PubMed PMID: 23992847. PubMed Central PMCID: PMC3961003.
4. Qin, W, Kutny, P. M., Maser, R. S., Dion, S. L., Lamont, J. D., Zhang, Y, Perry, G. A., Wang, H. 2016. Generating Mouse Models Using CRISPR-Cas9-Mediated Genome Editing. Curr. Protoc. Mouse Biol. 6:39-66. PubMed PMID: 26928663. PubMed Central PMCID: PMC4848752.
5. Qin, W., Dion. S.L., Kutny, P.M., Zhang. Y, Cheng, A.W., Jillette, N.L., Malhotra, A., Geurts, A.M., Chen. Y.G., and Wang. H. Efficient CRISPR/Cas9-Mediated Genome Editing in Mice by Zygote Electroporation of Nuclease. Genetics 2015 200(2): 423–430. PubMed PMID: 25819794. PubMed Central PMCID: PMC4492369.
6. Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153:910-918. PubMed PMID: 23643243. PubMed Central PMCID: PMC3969854.

TIPS FOR CRISPR GENE EDITING IN MICE

By Samantha Young | July 28, 2016



The use of CRISPR/Cas9 for gene editing has expanded since its adaptation for use in mammalian cells in 2012-2013. Researchers are now using this system in ever more creative ways, (Wang et al., 2013, Cho et al., 2014). There are several variants of the CRISPR/Cas9 system floating around, and many pre-designed plasmids containing these variants are ready for purchase. But what is the easiest and fastest way to use the system in mice? In this section, we will outline a simple method for selecting the guide RNA, validating its efficacy *in vitro*, and using it in mouse embryos to generate gene modified mouse lines. Hopefully this post will help get your *in vivo* research up and running as soon as possible!

Selecting Your Target

CRISPR/Cas9 is often used to study the biology surrounding a particular gene (although it has been quite successfully adapted for use in screening experiments as well). Therefore, if you've started down the CRISPR path, you've probably got a gene you'd like to manipulate in mind. Now it's time to consider where you would like to attack this gene. Should you cut it at the beginning of the coding sequence to delete the entire gene? Perhaps you are more interested in the c-terminus and would like to target that region. Or you may even wish to generate a point mutation and need to focus your attention to that particular region. No matter where you are looking, you can either scan the genome yourself for potential gRNA sites, or you can utilise several online tools that do it for you. One such option is the CRISPR Design tool supplied by the Zhang lab. This simple website invites you to input the target region (approximately 500 bp is best) and select the organism. It will then tell you the available gRNA targets in that region. The CRISPR Software Matchmaker can be used to determine the pros and cons of using many different types of gRNA design software.

When you are happy with your selection, ordering the sgRNA is simple and easy. Several companies supply gRNA sequences, such as GenScript, ThermoFisher Scientific and OriGene, or you can order them yourself through whatever DNA supply company you currently use. You can also find many previously validated gRNAs at Addgene.

In Vitro Screening for the Best gRNA

Not all gRNAs are effective. At the time of writing many of the specifics surrounding why some gRNAs work better than others remain unknown, but fret not! The key to saving time and energy during your *in vivo* work is to screen you gRNA *in vitro* first. *In vitro* screening is a simple and fast method for validating your selected gRNAs, and has the bonus benefit of supplying you with primers for validating your edit in mice once they have been generated (more on that later).

To begin the *in vitro* screening process, you should first develop primers to the approximately 500 base pair sequence surrounding your gRNA. Once amplified, you can insert this region into a pCAG-EGXXFP plasmid using standard cloning techniques (The XX indicates where your target is inserted, disrupting the EGFP

TIPS FOR CRISPR GENE EDITING IN MICE (CONT'D)

fluorescent signal, for details on this plasmid, see Mashiko et al., 2013). Testing the efficacy of your gRNAs is then a simple matter of transfecting HEK293T cells with your modified pCAG-EGXXFP plasmid along with the individual gRNAs in their PX330 plasmids (or similar plasmids for gRNA expression), one gRNA at a time. After 48 hours of incubation, you assess the level of fluorescence in your cells under a microscope. The higher the amount of fluorescence, the better that gRNA is at causing a double stranded break at the target site, allowing the two halves of EGFP to recombine by homology directed repair, and resulting in the expression of EGFP. Using this technique, you can assess several sgRNAs at the same time, and choose the most effective ones to increase your chances of gene modification *in vivo*.

CRISPR/Cas9 Gene Editing in Mice

Now that you've validated your gRNAs for use in mammalian cells, it's time to use them in mouse embryos. When researchers first started using CRISPR/Cas9 for genome editing in mice, they would often microinject gRNAs along with RNAs encoding Cas9 into the oocyte. This original method was developed as the use of CRISPR/Cas9 in mammalian cells was being explored and it allowed control over each individual component of the system. However, it is easier to directly inject the pX330 plasmids used in the validation assay (above) and it is just as effective (Mashiko et al., 2014). The plasmid is sturdier than the delicate RNA and therefore there has less chance of contamination with RNase leading to degradation that can set you back days. Working with plasmid DNA is much easier and doesn't require the extra precautions needed when dealing with RNA (e.g. face mask, and a dust free and specially cleaned work bench).

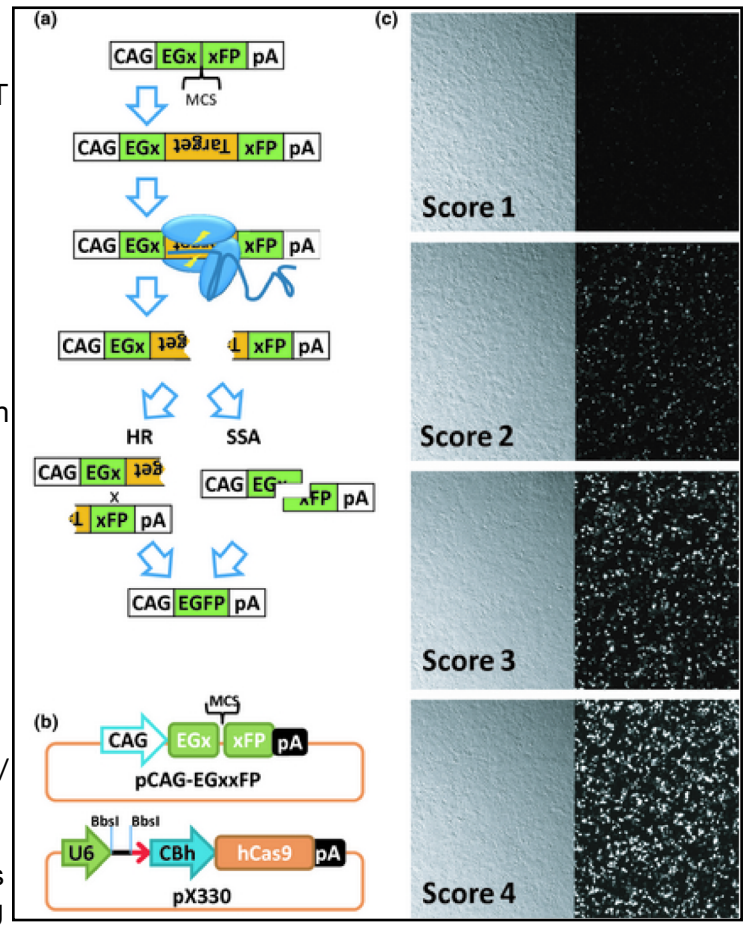


Figure 1: Validation of double strand break (DSB) mediated homology directed repair by enhanced green fluorescent protein (EGFP) reconstitution. (a) Scheme of validation for DSB mediated EGFP expression cassette reconstitution. When the target sequence was cut by the sgRNA guided Cas9 endonuclease, homology directed repair (HR, homologous recombination; SSA, single strand annealing) took place and reconstituted the EGFP expression cassette. (b) pCAG-EGxxFP target plasmid and pX330-sgRNA plasmid (Mashiko et al.). The pCAG-EGxxFP target plasmid contains overlapping 5′ and 3′ EGFP fragments under the ubiquitous CAG promoter. The approximately 500 bp genomic fragment containing the sgRNA target sequence can be placed in a multi-cloning site (MCS) between EGFP fragments. The pX330 plasmid contains a humanized Cas9 expression cassette and an sgRNA expression cassette. The target sgRNA can be cloned directionally into the BbsI site. (c) The pCAG-EGxxFP target plasmid was co-transfected with pX330 plasmids containing sgRNA sequences into HEK293T cells. The fluorescence intensity was classified into four groups (4; brighter than control, 3; same as control (Cetn1/sgRNA1), 2; darker than control, 1; very dark). Brighter fluorescence indicates a more successful gRNA design. Adapted from Development, Growth & Differentiation Volume 56, Issue 1, pages 122-129, 26 DEC 2013 DOI: 10.1111/dgd.12113 http://onlinelibrary.wiley.com/doi/10.1111/dgd.12113/full#dgd12113-fig-0001.

Once the Pups are Born

It takes roughly 19 days for the injected mouse embryos to develop. Once they are born, it's time to screen them for your desired edit. Good news! Remember the primers you designed to generate your pCAG-EGXXFP plasmid? They are the perfect primer sets to use to genotype your pups. A simple PCR with these primers will let you know if there are any major deletions, and sequencing of the PCR products will give specific details on what exactly has been inserted or deleted.

TIPS FOR CRISPR GENE EDITING IN MICE (CONT'D)

Each lab will use a slightly different method to genotype their pups. Taking a toe, a finger, a tail tip, or a combination of these will allow you to determine which pup is which once the sequencing comes back. Alternatively, if you can afford to wait, you can hold off on genotyping until the pups are 2-4 weeks old and use the ear tag method (creating different patterns of ear holes and using that material for genotyping), but bear in mind that this method will require you to house mice that may have no mutation, costing money and space. As a rule, screen all of your pups unless you have an enormous litter (over 15 pups). In this case you should be able to find your mutant from screening as little as 8-10 pups. The level of fluorescence you saw in the validation assay should give an accurate indication of how many pups will contain a mutation and many of the mice with the mutation will be homozygous for it in every cell; however, if the microinjection was done as the pronuclei were fusing, you may have several mosaic mice. To figure out what alleles are present in these mice, you can clone the PCR products into a pBluescript cloning vector and sequence the resulting plasmid.

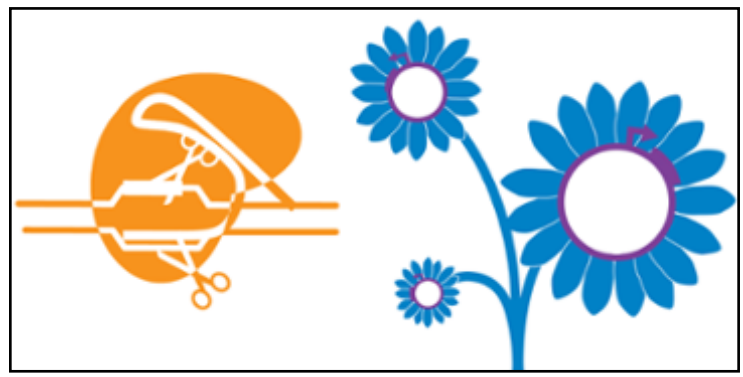
It can take roughly 3 months from the day you decide to do a gene modification experiment in mice to the day you have your F0 generation (barring life's little accidents and stuff ups). This is a huge improvement on previous mouse work, which could take 8 months or more. With such a rapid acceleration in the production of mouse lines that could be used to study a wide variety of phenomena, it's possible that we'll learn much more from our favourite mammalian model organisms in the near future.

Further Reading

1. Cho, Seung Woo, et al. "Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases." Genome research 24.1 (2014): 132-141. PubMed PMID: 24253446. PubMed Central PMCID: PMC3875854.
2. Mashiko, Daisuke, et al. "Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA." Scientific reports 3 (2013). PubMed PMID: 24284873. PubMed Central PMCID: PMC3842082.
3. Mashiko, Daisuke, et al. "Feasibility for a large scale mouse mutagenesis by injecting CRISPR/Cas plasmid into zygotes." Development, growth & differentiation 56.1 (2014): 122-129. PubMed PMID: 24372541.
4. Wang, Haoyi, et al. "One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering." Cell 153.4 (2013): 910-918. PubMed PMID: 23643243. PubMed Central PMCID: PMC3969854.

ENGINEERING THE PLANT GENOME USING CRISPR/CAS9

By Joel McDade | October 11, 2016



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 (1), 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 (2,3). 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 (4). 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.

This section will present a general overview of plant genome engineering using CRISPR, highlight the specific modifications to CRISPR machinery that allow for the use of CRISPR in plants and outline the various plant genome engineering tools that are available to academic researchers through Addgene.

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 (more information on multiplex CRISPR can be found in this section of the eBook).

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 type5. 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

ENGINEERING THE PLANT GENOME USING CRISPR/CAS9 (CONT'D)

RNAs (such as mRNAs for gene expression). 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 (5). 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.

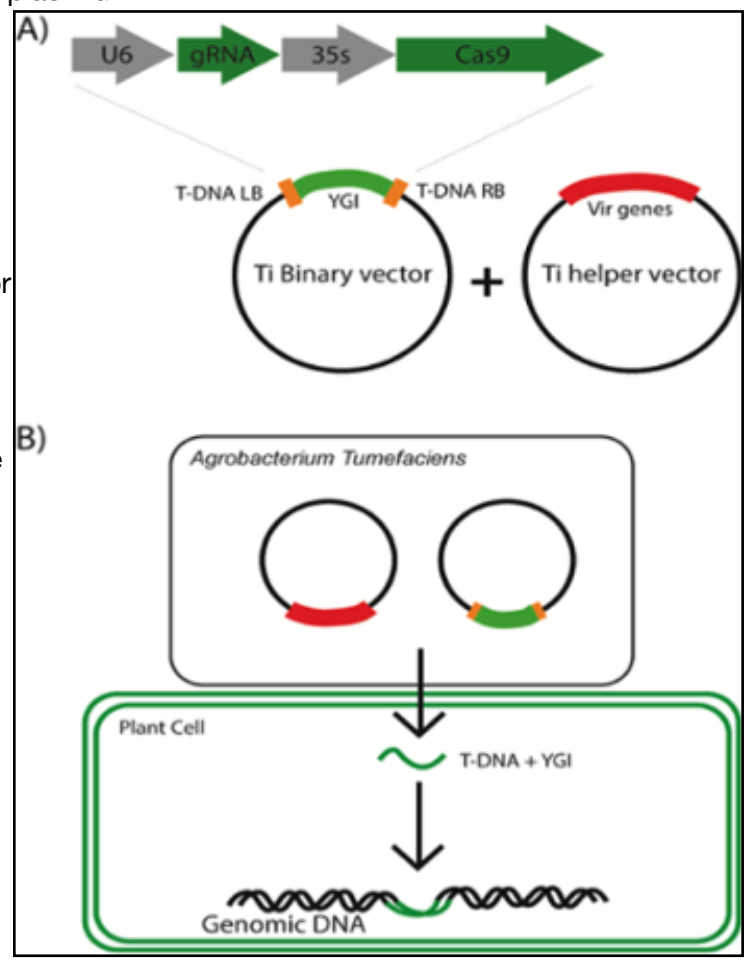


Figure 1: Simplified schematic of agrobacterium-mediated transformation in plant cells. Agrobacterium tumefaciens can be used as a vehicle to deliver your gene of interest (YGI) into plant cells. The system typically consists of a Ti plasmid into which YGI is inserted and a Ti "helper" plasmid, which contains VIR genes necesssary for T-DNA processing and insertion into the plant genome. The Ti plasmid and helper are transformed into agrobacterium tumefaciens and exposed to plant cells. The region of the Ti plasmid between the T-DNA borders (yellow boxes) is recognized by the Vir genes on the helper plasmid and used to insert the intervening region into the plant genome.

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

ENGINEERING THE PLANT GENOME USING CRISPR/CAS9 (CONT'D)

your experiment, but, if you're working with plants, don't be afraid to try your hand at CRISPR.

Further Reading

1. Waltz, Emily. "Gene-edited CRISPR mushroom escapes US regulation."Nature News 532 (2016): 293. PubMed PMID: 27111611.
2. Wang, Fujun, et al. "Enhanced Rice Blast Resistance by CRISPR/Cas9-Targeted Mutagenesis of the ERF Transcription Factor Gene OsERF922."PLoS One 11.4 (2016): e0154027. PubMed PMID: 27116122. PubMed Central PMCID: PMC4846023.
3. Wang, Yanpeng, et al. "Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew."Nature Biotechnology 32.9 (2014): 947-951. PubMed PMID: 25038773.
4. Khatodia, Surender, et al. "The CRISPR/Cas Genome-Editing Tool: Application in Improvement of Crops." Frontiers in Plant Science 7 (2016). PubMed PMID: 27148329. PubMed Central PMCID: PMC4835450.
5. Belhaj, Khaoula, et al. "Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system." Plant Methods 9.1 (2013): 1. PubMed PMID: 24112467. PubMed Central PMCID: PMC3852272.
6. Lowder, Levi G., et al. "A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation." Plant Physiology 169.2 (2015): 971-985. PubMed PMID: 26297141. PubMed Central PMCID: PMC4587453.
7. Gelvin, Stanton B. "Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool." Microbiology and Molecular Biology Reviews 67.1 (2003): 16-37. PubMed PMID: 12626681. PubMed Central PMCID: PMC150518.

AN "ELEGANS" APPROACH TO BETTER CRISPR/CAS9 EDITING EFFICIENCY

By Jordan Ward | January 27, 2015

This section was contributed by Jordan Ward who is a postdoctoral fellow at UCSF.

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 knock out sequences and introduce mutations and epitopes with unprecedented ease. In the last year, several advances in *C. elegans* genome editing using CRISPR/Cas9 have emerged, which I will describe below. These new *C. elegans* approaches rapidly enrich for editing events without the need for any selective marker to remain in the edited animal. To my knowledge these approaches have not yet been extended to other organisms/cell lines, though it is likely that many aspects will broadly improve editing efficiency.

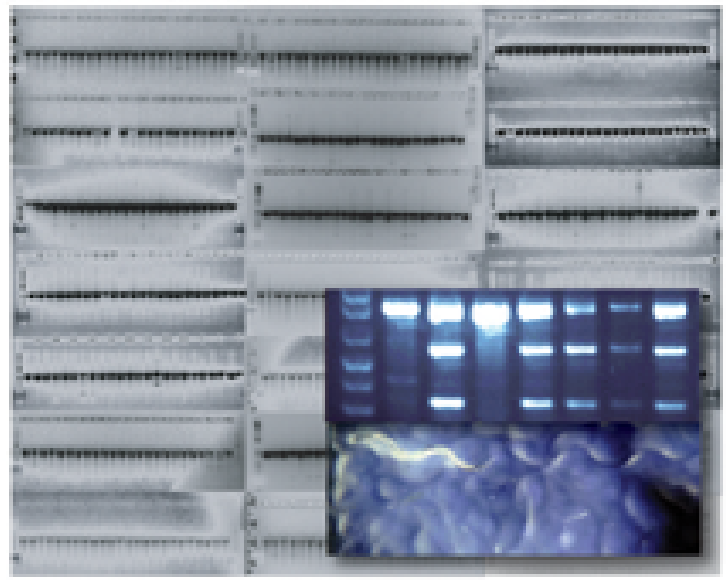


Image courtesy of Jordan Ward, UCSF.

Development of CRISPR/Cas9 Editing Strategies in *C. elegans*

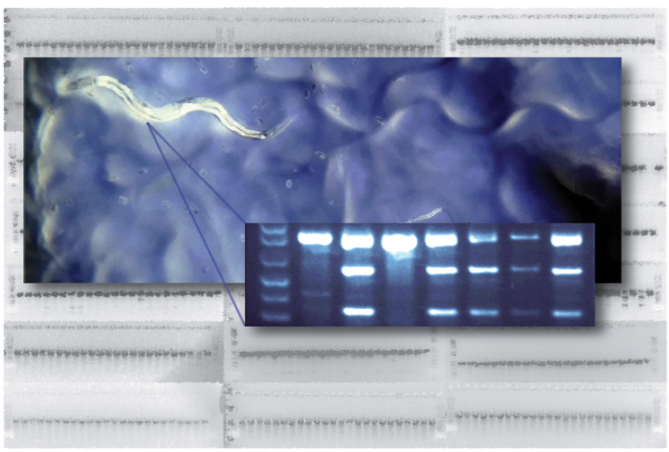
As brilliantly and irreverently illustrated in the parable of the Geneticist vs the Biochemist, we geneticists are a lazy bunch who love to rely on the awesome power of genetic selection. Knock-in events tend to be rare, and the challenge with any experiment is inefficient recovery of these edits. Initial 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. Including Cre-mediated marker excision, it takes approximately four weeks to obtain a homozygous, outcrossed knock-in ready for experimentation.

In the last year, several manuscripts published in *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 (i.e. PCR-derived dsDNA) with 30-60 basepair homology arms to knock in large epitopes, such as GFP (Paix et al., 2014). 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. Additionally, they take from 8-14 days to recover knock-in homozygotes ready for experimentation, depending on the screening strategy.

AN "ELEGANS" APPROACH TO BETTER CRISPR/CAS9 EDITING EFFICIENCY (CONT'D)

To Find a Needle, Remove the Haystack

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.



In the background are 504 restriction digests of PCR products used to identify a single 2xFLAG knock-in by direct screening. The center images depict the power of lethal mutation co-selection. Only animals rescued for the conditional lethal mutation survive the selection; from a representative co-selection experiment, five knock-ins were recovered from seven rescued F1 animals (gel image under worm). Image courtesy of Jordan Ward (UCSF).

In my recent *Genetics* paper (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 ºC, yet exhibit complete embryonic lethality at 25 ºC. 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 range 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 non-homologous end-joining 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.

Future Perspectives

It is interesting to note that recipients of the 2015 Breakthrough Prize included Jennifer Doudna and Emmanuelle Charpentier for their pioneering CRISPR/Cas9 work, 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.

AN "ELEGANS" APPROACH TO BETTER CRISPR/CAS9 EDITING EFFICIENCY (CONT'D)

Further Reading

1. Chen, Changchun, Lorenz A. Fenk, and Mario de Bono. "Efficient genome editing in Caenorhabditis elegans by CRISPR-targeted homologous recombination." Nucleic Acids Research 41.20 (2013): e193-e193. PubMed PMID: 24013562. PubMed Central PMCID: PMC3814388.
2. Tzur, Yonatan B., et al. "Heritable custom genomic modifications in Caenorhabditis elegans via a CRISPR–Cas9 system." Genetics 195.3 (2013): 1181-1185. PubMed PMID: 23979579. PubMed Central PMCID: PMC3813848.
3. Dickinson, Daniel J., et al. "Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination." Nature Methods 10.10 (2013): 1028-1034. PubMed PMID: 23995389. PubMed Central PMCID: PMC3905680.
4. Kim, Heesun, et al. "A Co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans." Genetics 197.4 (2014): 1069-1080. PubMed PMID: 24879462. PubMed Central PMCID: PMC4125384.
5. Arribere, Joshua A., et al. "Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans." Genetics 198.3 (2014): 837-846. PubMed PMID: 25161212. PubMed Central PMCID: PMC4224173.
6. Paix, Alexandre, et al. "Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 sites in Caenorhabditis elegans." Genetics 198.4 (2014): 1347-1356. PubMed PMID: 25249454. PubMed Central PMCID: PMC4256755.
7. Ward, Jordan D. "Rapid and precise engineering of the Caenorhabditis elegans genome with lethal mutation co-conversion and inactivation of NHEJ repair." Genetics 199.2 (2015): 363-377. PubMed PMID: 25491644. PubMed Central PMCID: PMC4317648.
8. Böttcher, Romy, et al. "Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells." Nucleic Acids Research 42.11 (2014): e89-e89. PubMed PMID: 24748663. PubMed Central PMCID: PMC4066747.

EVEN MORE ELEGANT: SINGLE INJECTION CRISPR/CAS9 IN C. ELEGANS

By Mary Gearing | July 7, 2015

In the summer of 2013, a remarkable nine papers describing CRISPR/Cas9 genome engineering methods for *C. elegans* were released, signaling a new era in *C. elegans* research. Homology directed repair (HDR), which enables insertion of custom genomic modifications, is very robust in *C. elegans*, and the methods for HDR-mediated modification continue to be improved. New work from Bob Goldstein's lab at the University of North Carolina has made using CRISPR in *C. elegans* even easier - now, one can generate a fluorescent protein fusion, transcriptional reporter, and loss-of-function allele in just one injection step! The entire protocol takes about 2-3 weeks but requires less than eight hours worth of hands-on time.

Creating a Modular, Selection-based Gene Tagging Approach

Multiple features of *C. elegans* have made CRISPR/Cas9 so successful in this species, namely, the robust HDR mentioned above, an accessible germline, and multiple selectable markers. Bob Goldstein's lab previously used a transgene selection strategy for their *C. elegans*/CRISPR experiments, but this strategy required the use of unhealthy, difficult-to-inject *unc-119(ed3)* animals, as well as a second injection step to remove the selection marker. Other labs have used PCR-based screening to detect mutants, but this method limits screening to a few hundred worms, making the detection of rare events difficult.

Dickinson et al. sought to create a new and efficient selection-based CRISPR/Cas9 method to generate knock-ins on any genetic background. Their new method uses a selection cassette containing genes for hygromycin resistance and the Roller movement phenotype (access the video below to see the Roller phenotype in action!). To make these markers easy to remove after injection, they added flanking loxP sites and a heat-shock inducible Cre to generate a self-excising cassette (SEC).

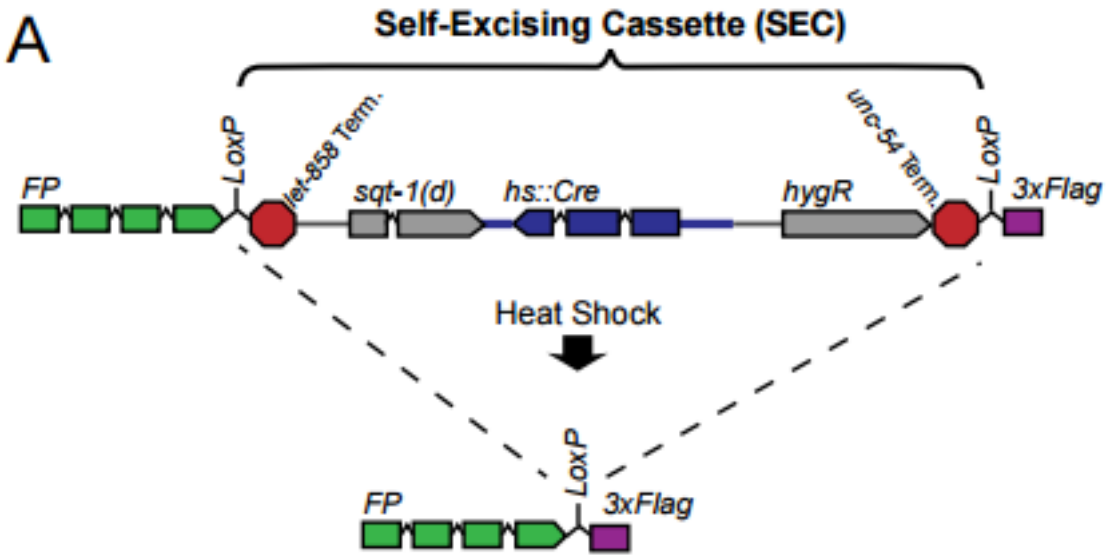
Screen shot of the movie from Dickinson et al. displaying the Rol phenotype. Fewer animals on the right display the Rol phenotype after the SEC casette has been removed. Used with permission from the Genetics Society of America. View the movie here.

To make it possible to fluorescently label proteins in *C. elegans*, they inserted the SEC into a synthetic intron between a fluorescent protein (FP) and a 3xFLAG tag (see construct schematic below). After self-excision, the resulting loxP scar is located in this synthetic intron, and only the fluorescent protein and 3xFLAG-tag sequences are inserted into the genome.

To use this SEC-based system, first design and clone 500-700 bp homology arms for your gene of interest into the FP-SEC vector using Gibson assembly. This construct will serve as the repair template plasmid to create fluorescent and 3xFLAG-tag fusion proteins.

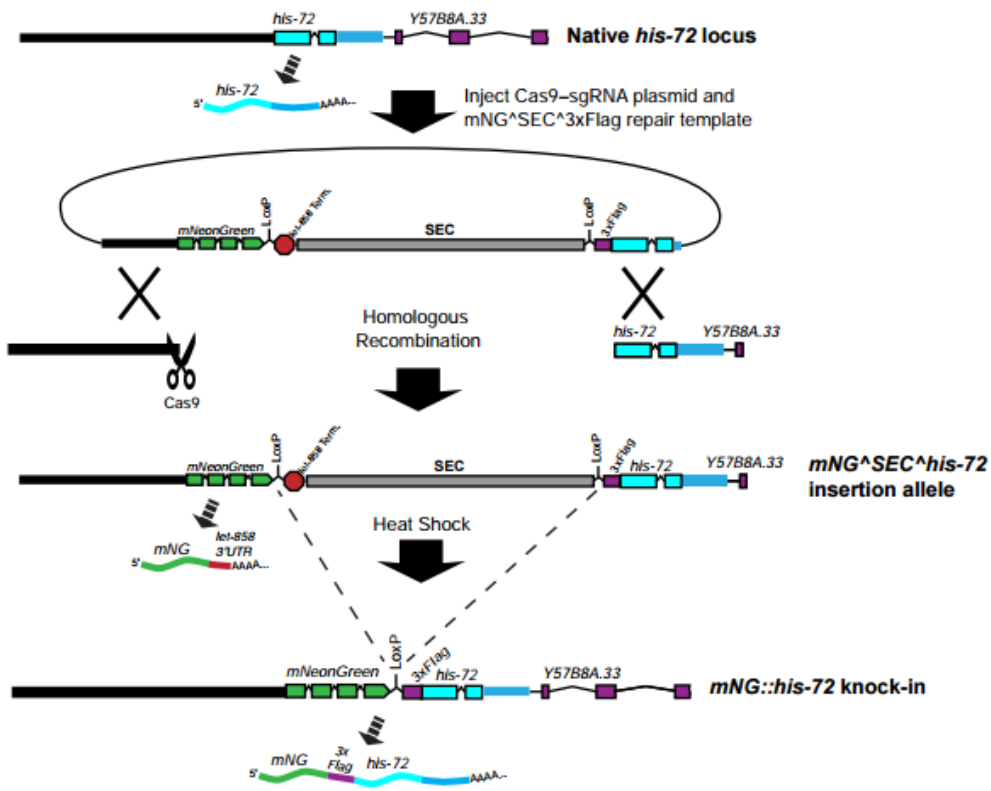
The repair template and Cas9-sgRNA plasmid are then injected into the germline of young adult worms. Once the worms have laid eggs, hygromycin is added to the plates to select for candidate knock-ins. These candidate animals can then be further screened based on their Roller phenotype. If the SEC is inserted at the 5' end of a gene, the SEC separates the targeted gene from its promoter, creating a loss-of-function allele as well as a promoter-FP fusion for the gene of interest. L1/L2 larvae from these strains are then heat shocked to remove the SEC and after 5-6 days of growth, wild-type animals that have undergone SEC excision are screened based on the loss of the Roller phenotype. These animals contain a fluorescent protein fusion on the targeted gene.

EVEN MORE ELEGANT: SINGLE INJECTION CRISPR/CAS9 IN C. ELEGANS (CONT'D)



Self-excising cassette schematic from Dickinson et al. Used with permission from the Genetics Society of America.

After initial testing, Dickinson et al. tested this system across 7 genes in parallel; when they injected 60-80 animals per gene, they obtained the desired insertions on the first attempt for 6/7 genes. The data are clear - the SEC-based system is both simple to use and highly efficient for *C. elegans* genome modification.



Schematic from Dickinson et al. illustrating the steps to fluorescently tag a C. elegans *protein. The his-72 locus is shown as an example. Used with permission from the Genetics Society of America.*

EVEN MORE ELEGANT: SINGLE INJECTION CRISPR/CAS9 IN C. ELEGANS (CONT'D)

Applications of this System

Dickinson et al. have developed a modular system that requires very little hands-on labor and could facilitate the tagging of every protein in the *C. elegans* genome. Although they tested fluorescent protein fusions, this system could be used to make many other genome modifications, including targeted mutations. SEC-based systems may prove effective in other species or in cell culture. In addition to CRISPR/Cas9 applications, SEC selection could replace other, more restrictive *C. elegans* selectable marker systems. This novel, rapid, and user-friendly system represents another leap forward for the CRISPR/Cas9 revolution!

Further Reading

1. Dickinson, Daniel J., et al. "Streamlined genome engineering with a self-excising drug selection cassette." Genetics 200.4 (2015): 1035-1049. PubMed PMID: 26044593. PubMed Central PMCID: PMC4574250.
2. Dickinson, Daniel J., et al. "Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination." Nature Methods 10.10 (2013): 1028-1034. PubMed PMID: 23995389. PubMed Central PMCID: PMC3905680.
3. Paix, Alexandre, et al. "Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 sites in Caenorhabditis elegans." Genetics 198.4 (2014): 1347-1356. PubMed PMID: 25249454. PubMed Central PMCID: PMC4256755.

CRISPR METHODS FOR BACTERIAL GENOME ENGINEERING

By Mary Gearing | March 3, 2016

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 (HDR) 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 do CRISPR editing in *E. coli* and other commonly-used bacterial species. Read on to learn about the tools available for bacteria and some of the applications for which they've been used.

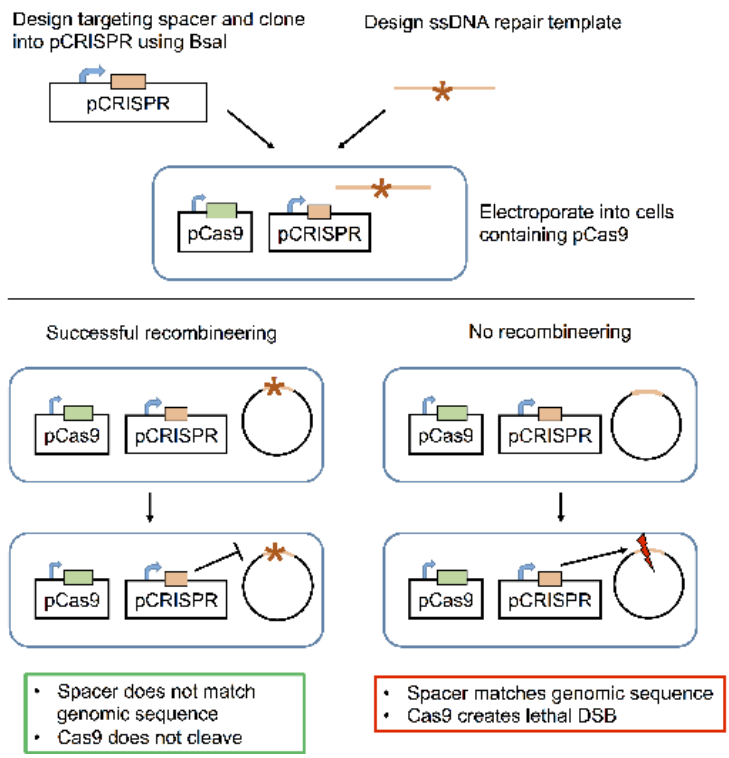
The Beginnings of Bacterial CRISPR Engineering

Much 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:

1. pCas9: carries Cas9 and chloramphenicol resistance
2. 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. 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 designed by Jiang et al. also functions in *S. pneumoniae* and can be used to generate multiple mutations simultaneously.

CRISPR METHODS FOR BACTERIAL GENOME ENGINEERING (CONT'D)

What's New: CRISPR Multiplexing in *E. coli* and Other Bacteria

CRISPR is clearly a powerful tool for bacterial engineering, and the work needed to adapt CRISPR systems to cover most bacterial species is ongoing. The good news is that CRISPR multiplexing is now available for multiple bacterial species/genera. See below for a list of plasmid-based systems available from Addgene - we hope your favorite bacteria are included!

Genome Editing

Huimin Zhao Lab pCRISPomyces Plasmids

Streptomyces bacteria produce a wide variety of bioactive natural products. To easily explore and engineer pathways within this genus, Cobb et al. created two "pCRISPomyces" systems for use in Streptomyces. 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 also be 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.

Sheng Yang Lab *E. coli* and *T. citrea* Scarless Editing Plasmids

Jiang et al.'s two-plasmid system combines the power of recombineering with that of CRISPR to create a system for scarless, iterative genome engineering. 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.

Prather Lab noSCAR *E. coli* Plasmids

Like Jiang et al.'s system, Reisch and Prather's noSCAR system incorporates phage recombination machinery into CRISPR editing to create scarless modifications. This two-plasmid system uses ssDNA or dsDNA repair templates to produce point mutations, insertions, or deletions. Tet-inducible Cas9 is located on the pCas9cr4 plasmid, and the targeting gRNA and recombination machinery are carried by pKDsg-xxx. pKDsg-xxx is easily cured after the desired modification has been made, allowing for multiple sequential rounds of editing. For most of the experiments conducted, 100% of colonies tested via colony PCR or sequencing displayed the desired mutation, indicating that the noSCAR method is highly efficient.

Tao Chen Lab *E. coli* Iterative Editing Plasmids

Li et al.'s two-plasmid system enables easy metabolic engineering in *E. coli* carrying phage recombination machinery. pGRB supplies one or more gRNAs inserted using Golden Gate Assembly. Cas9cur contains Cas9

CRISPR METHODS FOR BACTERIAL GENOME ENGINEERING (CONT'D)

and an inducible gRNA targeting the bla resistance gene, which can be used to rapidly cure the bla-containing pGRB. Either ssDNA or dsDNA may be used as a repair template. Each cycle of genome modification takes two days, and the system displays ~100% modification efficiency for deletions as large as 12 kb and insertions as large as 2 kb. Multiple mutations can be introduced simultaneously, albeit at a lower efficiency (83% for 2 mutations and 23% for 3 mutations).

Transcriptional Repression

Bacterial CRISPR methods are also available for transcriptional activation and repression. As RNA interference does not function in bacteria, most previous efforts to regulate gene expression were limited to the use of inducible promoters or direct gene knockout. In contrast, CRISPR offers a much more user-friendly way to modulate gene expression. Both Bikard et al. and Qi et al. developed early systems for use in E. coli; while Bikard et al. used a native minimal CRISPR array, Qi et al. employed a gRNA-based design more familiar to those using CRISPR in eukaryotes. As in other systems, catalytically dead (dCas9) targeted to a promoter or gene body can repress transcription by physically blocking the elongation complex, and gRNAs targeting the noncoding strand repress transcription more efficiently than those targeting the coding strand. Bikard et al. also successfully activated transcription in *E. coli* and *S. pneumoniae* using a Cas9-RNA polymerase omega subunit fusion guided to bind 80-100 bases upstream of the transcription start site.

Koffas Lab CRISPathBrick Multiplex Plasmid

This system allows you to assemble type II-A CRISPR arrays for dCas9-based transcriptional repression in *E. coli*. The pCRISPathBrick plasmid contains dCas9 and a nontargeting spacer flanked by two CRISPR repeats. The spacer can be digested using BsaI, allowing a spacer-repeat "brick" to be inserted. The BsaI site remains intact, allowing subsequent "bricks" to be added one by one. This approach is especially useful for combinatorial analyses. For example, if you were to develop an array using 3 distinct spacer-repeats (more are possible), you could easily create 7 unique arrays (e.g. for spacers A, B, and C, you could obtain arrays A, B, C, AB, AC, BC, and ABC).

Beisel Lab Type I CRISPR Plasmids

Luo et al. took a different approach to transcriptional repression: instead of adding a Type II system for transcriptional repression, they co-opted a native Type I system in *E. coli*. Whereas Type II systems require a single protein for DNA cleavage, Type I systems employ a multi-Cas protein complex. Luo et al. deleted one component of the Type I complex, the nuclease cas3, from the *E. coli* genome in order to create an inactive complex. Without Cas3 present, the other Type I Cas proteins can tightly bind a given locus and block transcription. pcrRNA.con and pcrRNA.ind are constitutive and arabinose-inducible empty array plasmids into which desired spacers can be cloned. Using an endogenous Type I system instead of the common Type II system decreases the amount of genetic material that must be transformed into the bacteria, and opens up more potential PAM sites (in this case, CTT and CCT).

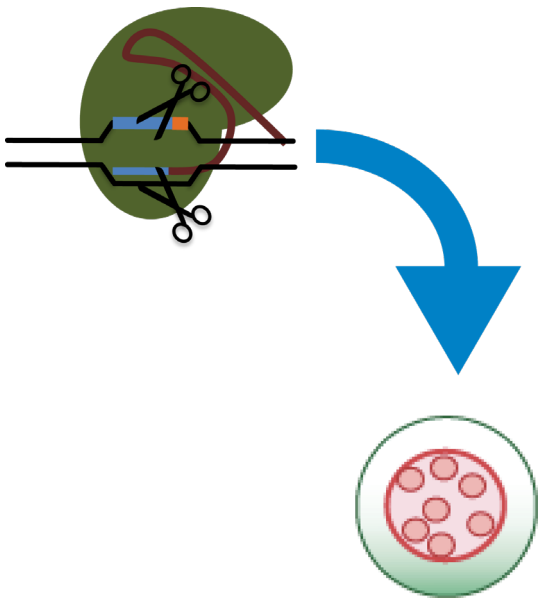
CRISPR METHODS FOR BACTERIAL GENOME ENGINEERING (CONT'D)

Further Reading

1. Jiang, Wenyan, et al. (2013). "RNA-guided editing of bacterial genomes using CRISPR-Cas systems." Nat Biotechnol. 31(3): 233-9. PubMed PMID: 23360965. PubMed Central PMCID: PMC3748948.
2. Bikard, David, et al. (2013). "Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system." Nucleic Acids Res. 41(15): 4729-37. PubMed PMID: 23761437. PubMed Central PMCID: PMC3753641.
3. Qi, Lei. S., et al. (2013). "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." Cell 152(5): 1173-83. PubMed PMID: 23452860. PubMed Central PMCID: PMC3664290.
4. Cobb, Ryan E., Wang, Yajie, and Huimin Zhao. (2015). "High-Efficiency Multiplex Genome Editing of Streptomyces Species Using an Engineered CRISPR/Cas System." ACS Synth Biol. 4(6): 723-8. PubMed PMID: 25458909. PubMed Central PMCID: PMC4459934.
5. Jiang, Yu, et al. (2015). "Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. Appl Environ Microbiol. 81(7): 2506-14. PubMed PMID: 25636838. PubMed Central PMCID: PMC4357945.
6. Reisch, Chris R. and Kristala L. J. Prather. "The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in Escherichia coli." Sci Rep. 5: 15096. PubMed PMID: 26463009. PubMed Central PMCID: PMC4604488.
7. Li, Yifan, et al. (2015). "Metabolic engineering of Escherichia coli using CRISPR-Cas mediated genome editing." Metab Eng. 31:13-21. PubMed PMID: 26141150.
8. Cress, Brady F., et al. (2015). "CRISPathBrick: Modular Combinatorial Assembly of Type II-A CRISPR Arrays for dCas9-Mediated Multiplex Transcriptional Repression in E. coli." ACS Synth Biol. 4(9): 987-1000. PubMed PMID: 25822415.
9. Luo, Michelle L., Mullis, Adam S., Leenay, Ryan T. and Chase L. Beisel. (2015). "Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression." Nucleic Acids Res. 43(1): 674-681. PubMed PMID: 25326321. PubMed Central PMCID: PMC4288209.

USING SaCAS9 FOR EASIER AAV-BASED CRISPR DELIVERY

By Mary Gearing | July 14, 2015



CRISPR/Cas9 genome editing has quickly become the most 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 chimeric sgRNA together (~4.2 kb) into an AAV vector is challenging due to the low packaging capacity of AAV (~4.5 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.

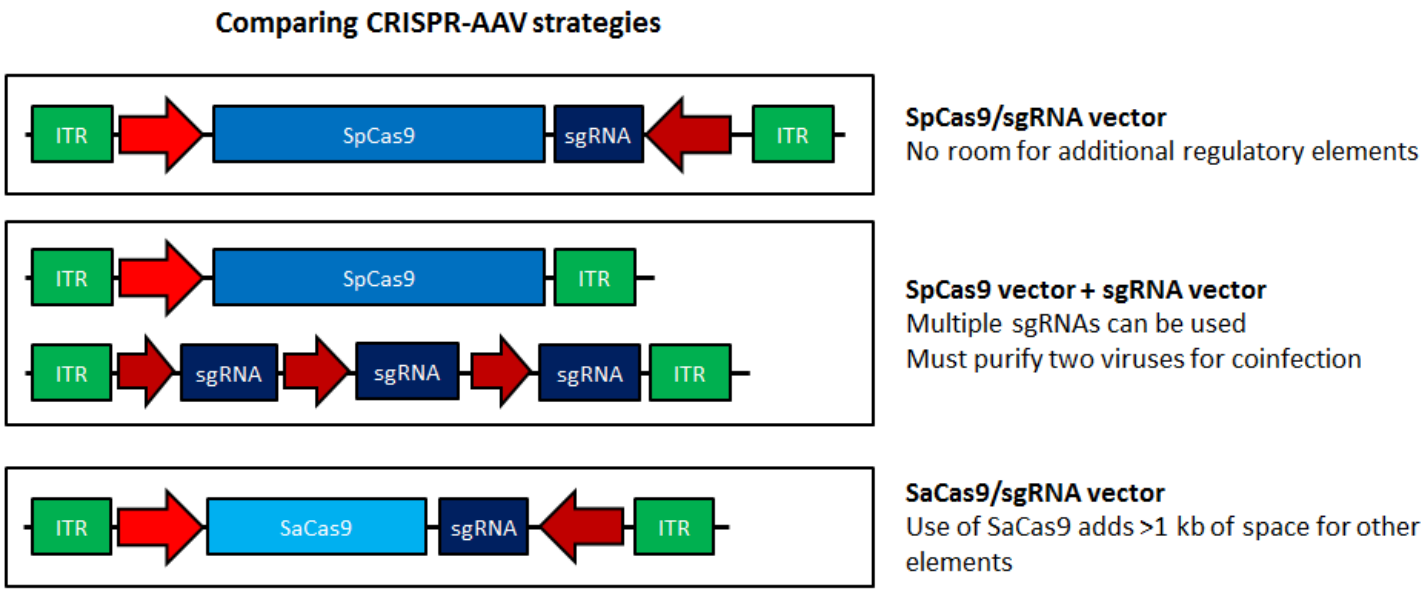
Cas9 Orthologs: Shorter, but Just as Potent and Specific?

The previous two AAV strategies described above showed successful target modification, indicating that AAV is a good delivery vehicle for Cas9. To maximize the genetic capacity of AAV, Gang Bao's group has developed a split-intein Cas9 that can be separated into two AAV cassettes, providing even more room for regulatory sequences and additional gRNAs in each cassette. However, to fit Cas9 and gRNAs into one AAV construct, the construct must be made even 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.

Ran et al. recently developed a new strategy to overcome these drawbacks. To discover a shorter, but equally potent Cas9 enzyme, they analyzed over 600 Cas9 orthologs and found that they could be divided into two groups: one with orthologs of ~1350 amino acids, which includes SpCas9, and one with orthologs of ~1000 amino acids. From the pool of shorter orthologs, only *Staphylococcus aureus* Cas9 (SaCas9) 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, Ran et al. used an approach called BLESS (direct in situ breaks labelling, enrichment on streptavidin and next-generation sequencing). Using this sensitive method, Ran et. al found that SaCas9 did not display higher levels of off-target activity than SpCas9, confirming its suitability for *in vivo* studies.

USING SaCAS9 FOR EASIER AAV-BASED CRISPR DELIVERY (CONT'D)



Testing AAV-SaCas9 In Vivo

To test the efficiency of AAV-SaCas9 *in vivo*, Ran et al. created an all-in-one SaCas9 and sgRNA construct using the liver-specific serotype AAV8. Since the efficiency of CRISPR/Cas9 genome editing varies across targets, they tested two genes in mice. For both genes, they observed 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 AAV-SaCas9-sgRNA constructs mediate genome modification, but they did so without a substantial immune response or toxicity.

The work of Zhang's group illustrates the potential of combining an advantageous vector delivery system (AAV) with a potent genome modification technique (CRISPR/Cas9). In this "best of both worlds" scenario, *in vivo* genome editing without substantial toxicity or off-target effects will likely become much easier than we could have imagined.

If you're interested in using SaCas9 in your research, the AAV targeting constructs are available from Addgene.

Further Reading

1. Ran, F. Ann, et al. "*In vivo* genome editing using *Staphylococcus aureus* Cas9." Nature 520.7546 (2015): 186-191. PubMed PMID: 25830891. PubMed Central PMCID: PMC4393360.
2. Swiech, Lukasz, et al. "*In vivo* interrogation of gene function in the mammalian brain using CRISPR-Cas9." Nature Biotechnology 33.1 (2015): 102-106. PubMed PMID: 25326897. PubMed Central PMCID: PMC4492112.
3. Cong, Le, et al. "Multiplex genome engineering using CRISPR/Cas systems." Science 339.6121 (2013): 819-823. PubMed PMID: 23287718. PubMed Central PMCID: PMC3795411.
4. Fine, Eli J., et al. "Trans-spliced Cas9 allows cleavage of HBB and CCR5 genes in human cells using compact expression cassettes." Scientific Reports 5 (2015). PubMed PMID: 26126518. PubMed Central PMCID: PMC4486982.

USING SaCAS9 FOR EASIER AAV-BASED CRISPR DELIVERY (CONT'D)

5. Senís, Elena, et al. "CRISPR/Cas9-mediated genome engineering: An adeno-associated viral (AAV) vector toolbox." Biotechnology Journal 9.11 (2014): 1402-1412. PubMed PMID: 25186301.
6. Garneau, Josiane E., et al. "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA." Nature 468.7320 (2010): 67-71. PubMed PMID: 21048762.
7. Nishimasu, Hiroshi, et al. "Crystal structure of Cas9 in complex with guide RNA and target DNA." Cell 156.5 (2014): 935-949. PubMed PMID: 24529477. PubMed Central PMCID: PMC4139937.

ADENOVIRAL DELIVERY OF CRISPR/CAS9 AIMS TO EXPAND GENOME EDITING TO PRIMARY CELLS

By Kendall Morgan | September 30, 2014

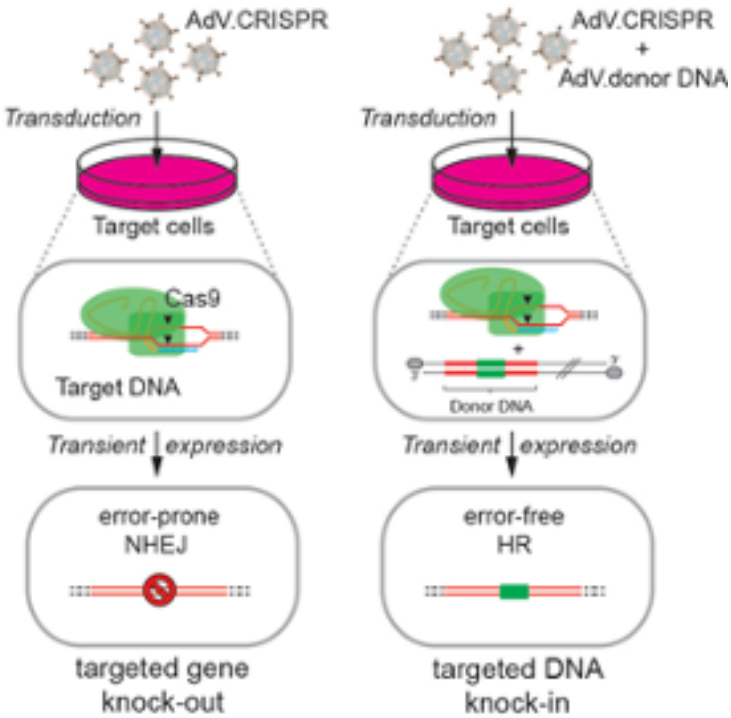


Image courtesy of Manuel Gonçalves.

Researchers have shown that it is possible to deliver RNA-guided CRISPR/Cas9 nuclease complexes using adenoviral vectors (AdVs), to a wide range of human cells, including mesenchymal stem cells, and in a rather straightforward manner. 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. "Although AdVs are being deployed for delivering zinc-finger nucleases into human cells, we think they are still underused in the emerging field of genome editing," Gonçalves said. "In contrast, AdVs are extensively being explored for genetic vaccination and oncolytic approaches. In genome editing, they are not used much, but we do think they have a very bright future."

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 mammallian cells.

Delivering gRNA and Cas9 Using Adenoviral Vectors

In a *Scientific Reports* paper introducing the delivery method in May 2014, the researchers report 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%.

A second paper published online in *Nature Methods* in August 2014 found 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. 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."

"We think this [most recent] work gives additional rationale for investigating the usefulness of adenoviral vector technology in the context of genome editing," he said, adding that he hopes others will now begin to make use of the new AdV delivery tools for a variety of applications.

"It would be rewarding if these reagents and protocol are picked up and people start to explore and test this

ADENOVIRAL DELIVERY OF CRISPR/CAS9 AIMS TO EXPAND GENOME EDITING TO PRIMARY CELLS (CONT'D)

method of introducing the CRISPR system into a broader range of cells – primary cells and cells that are not transformed – and eventually also consider *in vivo* applications."

Start Your CRISPR/Cas9 Research!

To find more information about the adenoviral delivery of CRISPRS/Cas9 using the Gonçalves lab's 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.

Further Reading

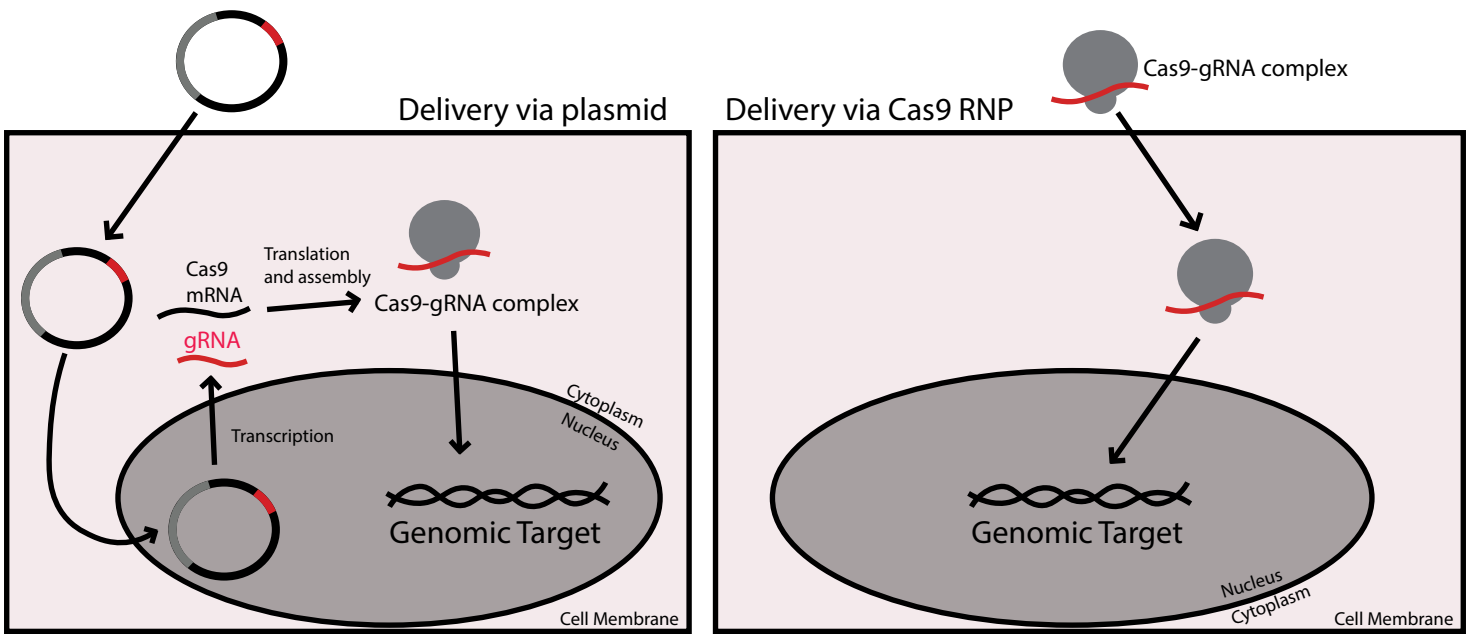
1. Maggio, Ignazio, et al. "Adenoviral vector delivery of RNA-guided CRISPR/Cas9 nuclease complexes induces targeted mutagenesis in a diverse array of human cells." Scientific Reports 4 (2014). PubMed PMID: 24870050. PubMed Central PMCID: PMC4037712.
2. Holkers, Maarten, et al. "Adenoviral vector DNA for accurate genome editing with engineered nucleases." Nature Methods 11.10 (2014): 1051-1057. PubMed PMID: 25152084.

GENOME ENGINEERING USING CAS9/GRNA RIBONUCLEOPROTEINS (RNPS)

By Joel McDade | April 21, 2016

CRISPR has quickly become the preferred system for genome engineering due to its simplicity, as it requires only Cas9 and a guide RNA (gRNA). Choosing the correct method to deliver both Cas9 and gRNAs to your target cells is absolutely critical as failure to adequately express either component will result in a failed experiment. In our previous section on Mammalian Expression Systems and Delivery Methods we provided a general overview of the most common ways in which you can deliver Cas9 and gRNAs to your target cells and discussed a few key advantages and disadvantages of each method. In this section, we will go into greater detail about why and how Cas9/gRNA Ribonucleoprotein complexes (Cas9 RNPs) are being used for genome engineering experiments and provide a general framework for getting started with Cas9 RNPs in your research.

What Are Cas9 RNPs and Why Are They So Useful?



Delivery of CRISPR components via plasmid vs. Cas9 RNP. Plasmid-based delivery of CRISPR components requires transcription, translation, and assembly of the Cas9-gRNA complex before target DNA can be modified. Cas9 RNPs are delivered to cells as pre-assembled Cas9-gRNA complexes and are free to interact with target DNA.

Cas9 RNPs consist of purified Cas9 protein in complex with a gRNA. They are assembled *in vitro* and can be delivered directly to cells using standard electroporation or transfection techniques. Cas9 RNPs are capable of cleaving genomic targets with similar efficiency as compared to plasmid-based expression of Cas9/gRNA and can be used for most of the current genome engineering applications of CRISPR, including: generating single or multi-gene knockouts in a wide variety of cell types, gene editing using homology directed repair (HDR), and generating large genomic deletions.

Cas9 RNPs differ from plasmid or viral-based delivery of CRISPR components with regards to how quickly the components are expressed and how long they are present within the cell. Plasmid or viral delivery of Cas9 and gRNA(s) requires the use of cellular transcription/translation machinery to generate functional Cas9-gRNA complexes, which results in a significant lag in peak Cas9 protein expression (>12 hours). Expression of each component continues indefinitely (for lentiviral-mediated delivery) or until the DNA is lost through cell division (for plasmid or AAV-based delivery). By contrast, Cas9 RNPs are delivered as intact complexes, are detectable

GENOME ENGINEERING USING CAS9/GRNA RIBONUCLEOPROTEINS (RNPS) (CONT'D)

at high levels shortly after transfection, and are quickly cleared from the cell via protein degradation pathways. There are two major consequences of the distinct kinetic profile of Cas9 RNPs. First, using Cas9 RNPs may increase the rate at which mutations form in target genes compared to plasmid-mediated delivery of Cas9 and gRNAs; Cas9 RNPs are delivered as functional complexes capable of cleaving target DNA and don't need to be transcribed and translated. Second, rapid clearance of Cas9 RNPs from the cell may increase CRISPR specificity by reducing the amount of time that Cas9 is available for off-target cleavage. The aforementioned characteristics of Cas9 RNPs make them useful for CRISPR applications where limited expression of Cas9 is required and specificity is a concern, such as knockout generation or homologous recombination. Experiments that require long-term expression of Cas9, such as visualizing genomic loci using fluorophore tagged dCas9 may require the use of plasmid or viral-mediated delivery.

You've Decided to Use RNPs, Now What?

The following is a general workflow for using Cas9-RNPs for genome engineering. First, you will need to obtain purified Cas9 and purified gRNAs targeting your specific locus of interest. gRNAs should be designed based on standard gRNA design principles, making sure to pick targeting sequences that are upstream of a PAM sequence and unique to the target compared to the rest of the genome. Purified gRNAs can be generated by PCR amplification of annealed gRNA oligos or *in vitro* transcription of a linearized gRNA containing plasmid (such as Addgene plasmid 42250 from Keith Joung's lab). Cas9 (or a variant of Cas9) can be purified from bacteria through the use of bacterial Cas9 expression plasmids, including these plasmids from Jacob Corn's group at the Innovative Genomics Initiative. In most cases, His-tagged Cas9 is expressed in bacterial cells and then purified using nickel affinity chromatography. Alternatively, purified Cas9 can be purchased from a variety commercial sources including NEB and Thermo Fisher.

Cas9 RNP delivery to target cells is typically carried out via lipid-mediated transfection or electroporation. Liang et al. 2015 compared electroporation to lipid based transfection of Cas9 RNPs for two DNA targets across 11 cell lines. For several of the cell lines, electroporation yielded high cleavage efficiency when lipid-based delivery completely failed, suggesting that electroporation may be more suitable for difficult to transfect cell types. Interestingly, the cell types that were resistant to lipid-mediated Cas9 RNP delivery were also resistant to lipid-mediated plasmid delivery. So, if it is difficult to deliver plasmids to your specific cell type using lipid reagents, it may be difficult to deliver RNPs, as well. However, Zuris et al. 2014 demonstrated that lipid-mediated delivery can be used to modify genomic targets in human cells in culture and mouse outer hair cells *in vivo* and recent advances in lipid chemistry may increase the efficiency of lipid-mediated Cas9 RNP delivery. Ultimately, selecting a delivery method for Cas9 RNPs will require some experimentation and optimization for your specific cell type. Once you have treated your cells with Cas9 RNPs, you should validate your edit, either by isolating individual clones and screening your target locus with Sanger sequencing or analyzing cleavage efficiency using a restriction digest-based assay (see section on validation).

Wrapping It All Up

In summary, there are several advantages to using Cas9 RNPs in your CRISPR experiment. Cas9 RNPs can be generated quickly and delivered directly to cells as fully functional Cas9-gRNA complexes. Cas9 RNPs remove the necessity of cloning targeting oligos into a plasmid backbone, which enables researchers to go from designing gRNA(s) to validating a genome edit in as little as 3-4 days. Cas9 RNPs are active immediately following transfection and are quickly degraded within the cell. These fast degradation kinetics enable Cas9 RNPs to modify target genes with reduced off-target effects. Of course, Cas9 RNPs are not without limitations. A major drawback of using Cas9 RNPs is that expression is transient. Therefore, it may be best to use plasmid-

GENOME ENGINEERING USING CAS9/GRNA RIBONUCLEOPROTEINS (RNPS) (CONT'D)

based delivery or lentiviral-mediated delivery of CRISPR components in cases where stable or elevated expression of CRISPR components is necessary.

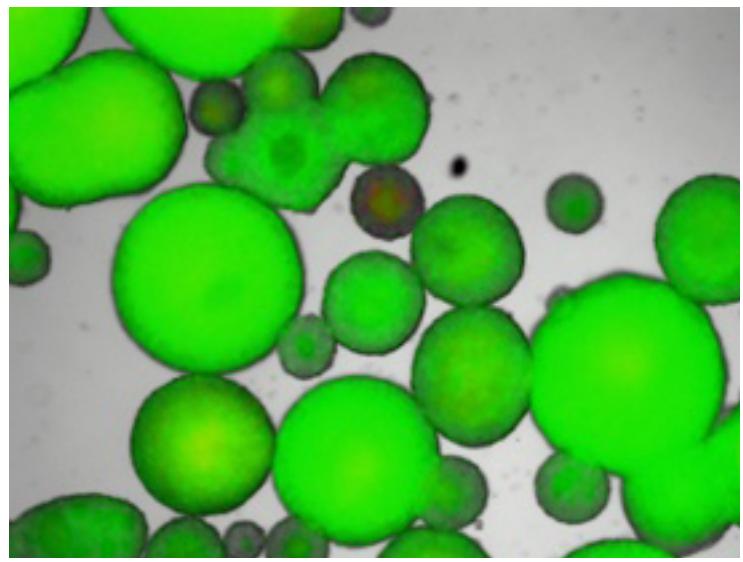
Further Reading

1. Liang, Xiquan, et al. "Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection." Journal of biotechnology 208 (2015): 44-53. PubMed PMID: 26003884.
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): 73-80. PubMed PMID: 25357182. PubMed Central PMCID: PMC4289409.
3. Kim, Sojung, et al. "Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins." Genome research 24.6 (2014): 1012-1019. PubMed PMID: 24696461. PubMed Central PMCID: PMC4032847.
4. Wang, Ming, et al. "Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles." Proceedings of the National Academy of Sciences 113.11 (2016): 2868-2873. PubMed PMID: 26929348.

ANOTHER PATHWAY INTO CELLS: ITOP PROTEIN ONLY TRANSDUCTION

By Mary Gearing | June 23, 2015

Primary cells recapitulate the natural biology of a cell type of interest better than immortalized lines derived from the same cell type; however, their use has been limited by technical problems. For instance, it's much more difficult to introduce a gene of interest into primary cells, so most primary cell lines require viral infection. A paper from Niels Geijsen's lab suggests that primary cells may be better transduced using only protein. Read on for a description of the lab's iTOP protein-only transduction method and its potential applications to CRISPR/Cas9 genome editing.



Neuronal stem cells transduced with iTOP resulting in GFP expression. Image courtesy of Diego D'Astolfo.

Protein Only Transduction?

Niels Geijsen's group in Utrecht wanted to pursue a nonviral strategy for transducing primary cells. What if, instead of supplying a DNA construct, they could simply induce cells to take up protein directly? It's not a new idea, but it comes with a few caveats. The most common protein-only transduction strategy is to fuse a cell penetrating peptide (CPP) to a protein of interest thereby giving it the capacity to enter cells. One such CPP was discovered in the HIV protein Tat. Although CPP fusions can translocate across a cell membrane, the presence of the CPP may alter protein function or localization. Thus, the usefulness of CPPs must be evaluated on a case-by-case basis.

While setting up their CPP system to study protein transduction, D'Astolfo et al. made the surprising discovery that both purified CPP-tagged protein and an untagged control could enter cultured cells and activate expression of a luciferase reporter. CPP was not required for protein entry into cells! After verifying the specificity of their luciferase assay, they concluded that components in the buffer used for protein purification must be helping the untagged protein enter cells. Subsequent work found that both NaCl and non-detergent sulfobetaine 201 (NDSB-201) were required for protein delivery. D'Astolfo et al. coined the term iTOP (induced transduction by osmocytosis and propanebetaine) to describe this method of protein transduction. Later studies showed that iTOP functions through macropinocytosis.

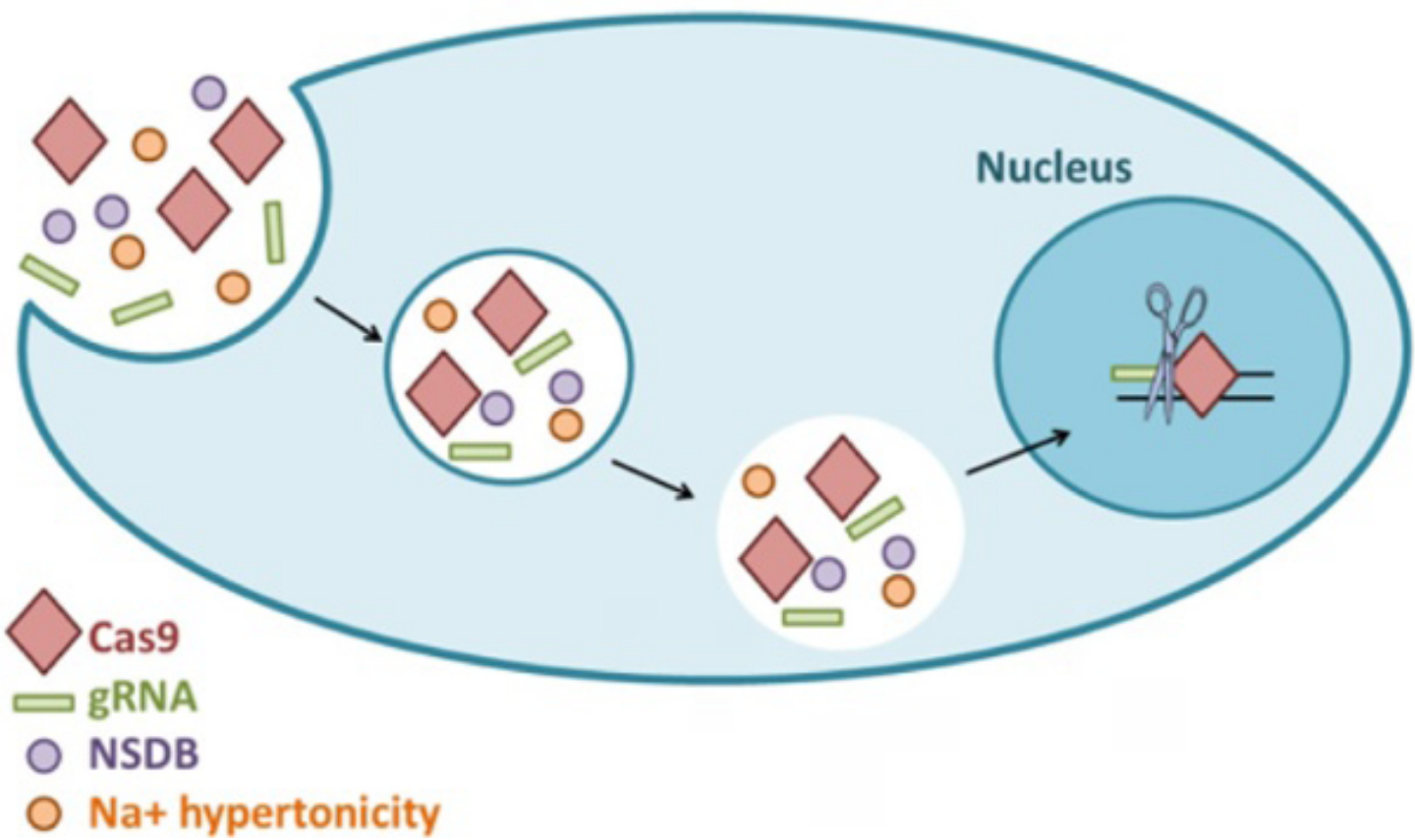
D'Astolfo et al. tested iTOP in a series of primary cell types using Cre and various loxP reporters. In both mouse and human embryonic stem cells, they found very high recombination percentages (78-79% after two rounds of transduction). Compared to CCP transduction, iTOP is at least four times more efficient in primary fibroblasts. iTOP also works in multiple types of primary mouse cells, including neuronal and gut stem cells, dendritic cells, embryonic fibroblasts, glia cells and neurons, and cell death post-transduction is low.

CRISPR/Cas9: A Potential Application of iTOP

Having tested iTOP using Cre-lox recombination, D'Astolfo et al. were eager to see if iTOP would be compatible with CRISPR/Cas9 genome editing. Although CRISPR/Cas9 is all the rage, they had another reason to test this system. It's difficult to determine how much protein iTOP causes a cell to take up, and continued protein expression would require multiple transduction events. For these reasons, iTOP is best suited to binary systems in which transient protein expression has a long-lasting effect, such as Cre- or Cas9-mediated recombination.

ANOTHER PATHWAY INTO CELLS: ITOP PROTEIN ONLY TRANSDUCTION (CONT'D)

Both primary hESCs and immortalized cells treated with iTOP efficiently took up both Cas9 protein and a gRNA (yes, iTOP can also be used for RNA transduction) targeting a gene associated with diphtheria toxin lethality and were sorted into single wells to generate clonal populations. After iTOP treatment, both primary (hESCs) and cultured cells showed up to 70% resistance to diphtheria toxin, indicating a 70% frequency of biallelic mutations. Importantly, hESCs retained their stemness, indicating that iTOP may be valuable for CRISPR/Cas9 stem cell editing.



The small molecule, NSDB, and Na+ hypertonicity are key elements of iTOP. iTOP can be used to transduce Cas9 and a gRNA, which subsequently enter the nucleus and cleave the specified target. Figure adapted from D'Astolfo et al 2015.

In certain instances, the potential benefits of nonviral transduction are numerous. First, you remove the risk of viral integration causing mutations and gene expression changes elsewhere in the genome. Second, expression is by nature transient, which may lower nuclease toxicity or the potential for off-target effects. Third, protein purification (once a protocol has been established) may be less laborious than viral purification, and it does not raise the safety concerns associated with virus work. Fourth, working directly with proteins removes the limitations of packaging size associated with viral vectors. For these reasons, iTOP, especially in tandem with genome editing, may make primary cell work much more efficient and feasible for many labs.

Further Reading

1. D'Astolfo, Diego S., *et al.* "Efficient Intracellular Delivery of Native Proteins." Cell 161.3 (2015): 674-690. PubMed PMID: 25910214.

OPTIMIZING DONOR DNA FOR ENHANCED CRISPR GENOME EDITING

By Chris Richardson | Mar 24, 2016

This post was contributed by guest blogger Chris Richardson, a Postdoctoral Researcher in Jacob Corn's lab.

CRISPR-Cas9 (Cas9) is an RNA-guided nuclease that targets and cuts genomic DNA. The interplay between Cas9 (which causes the breaks) and host cell DNA repair factors (which repair those breaks) makes Cas9 extremely effective as a genome editing reagent. This interplay falls into two broad categories and thus, causes two types of editing outcomes: Cas9 breaks repaired by the non-homologous end-joining (NHEJ) pathway disrupt target gene sequences (thus inactivating genes), while breaks repaired by homology directed repair (HDR) pathways can modify the sequence of a gene (thus altering its function). HDR is crucial for certain applications, for example, correcting the allele that causes sickle cell anemia. However, HDR occurs much less frequently than NHEJ and the efficiency of these editing reactions is low. Understanding the biological cause of this repair bias is a fascinating (and yet unanswered) question. Our recent paper (Richardson et al 2016) revealed some of the biophysical parameters that can influence the HDR/NHEJ decision.

Our conclusions reinforce a key scientific principle: that understanding a biological process in detail often suggests new strategies to manipulate that process. We began our research with a simple question: how do Cas9 and sgRNA interact with and dissociate from a DNA target? The surprising answer to this question led us to develop a new, rational approach to improve the efficiency of sequence replacement by HDR.

Understanding Cas9 Biophysics Enhances HDR

We first measured Cas9 dissociation rates and discovered that the catalytically active form of the nuclease dissociates very slowly after introducing a cut, with residence times on DNA of approximately 6 hours. To our surprise, we measured identical dissociation rates for catalytically inactive dCas9 molecules. We followed these experiments with more detailed investigations into the release of DNA post-cut. Using substrate DNA labeled on both sides of the break, we determined that Cas9 held tightly to three of the four strands of the cleaved duplex, while the fourth strand (the released strand in Figure 1 below) was free to anneal to complementary ssDNA molecules in vitro.

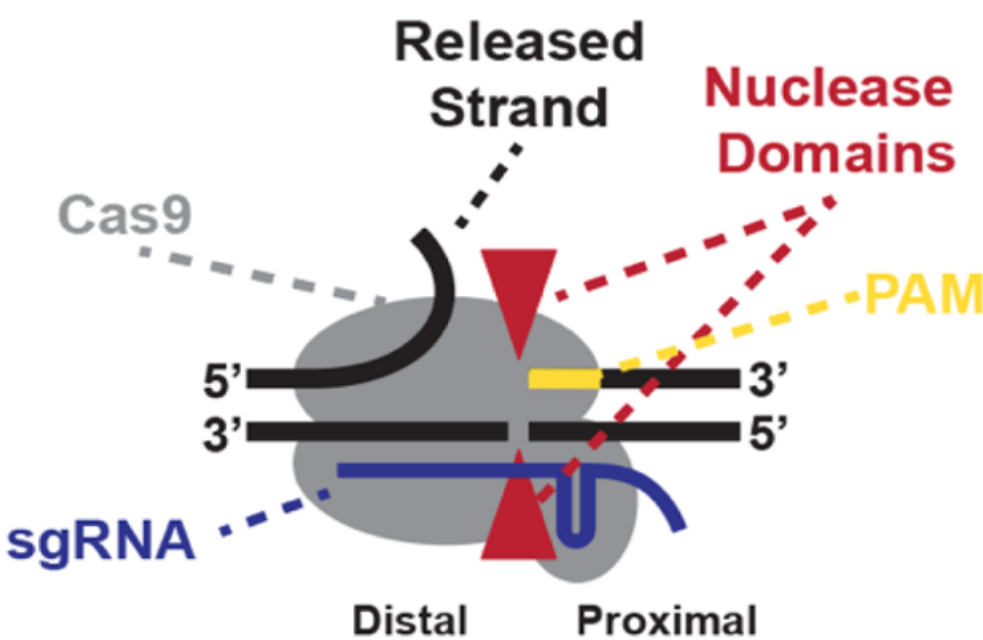


Figure 1: Cas9 Binding to the cleaved DNA Duplex

OPTIMIZING DONOR DNA FOR ENHANCED CRISPR GENOME EDITING (CONT'D)

We next wondered if we could leverage this model (built from *in vitro* studies) to boost the efficiency of sequence replacement *in vivo*. We discovered that single stranded donor DNA with the following 3 properties:

1. Complementarity to the released strand,
2. A legnth of 127 bp,
3. With 36 bp on the PAM-distal side of the break and 91 bp on the PAM-proximal side,

consistently supported higher frequencies (up to 60%!) of gene editing than single or double stranded donor DNA with other design parameters. Example single stranded donor DNA and a presumptive mechanism of action are presented in Figure 2.

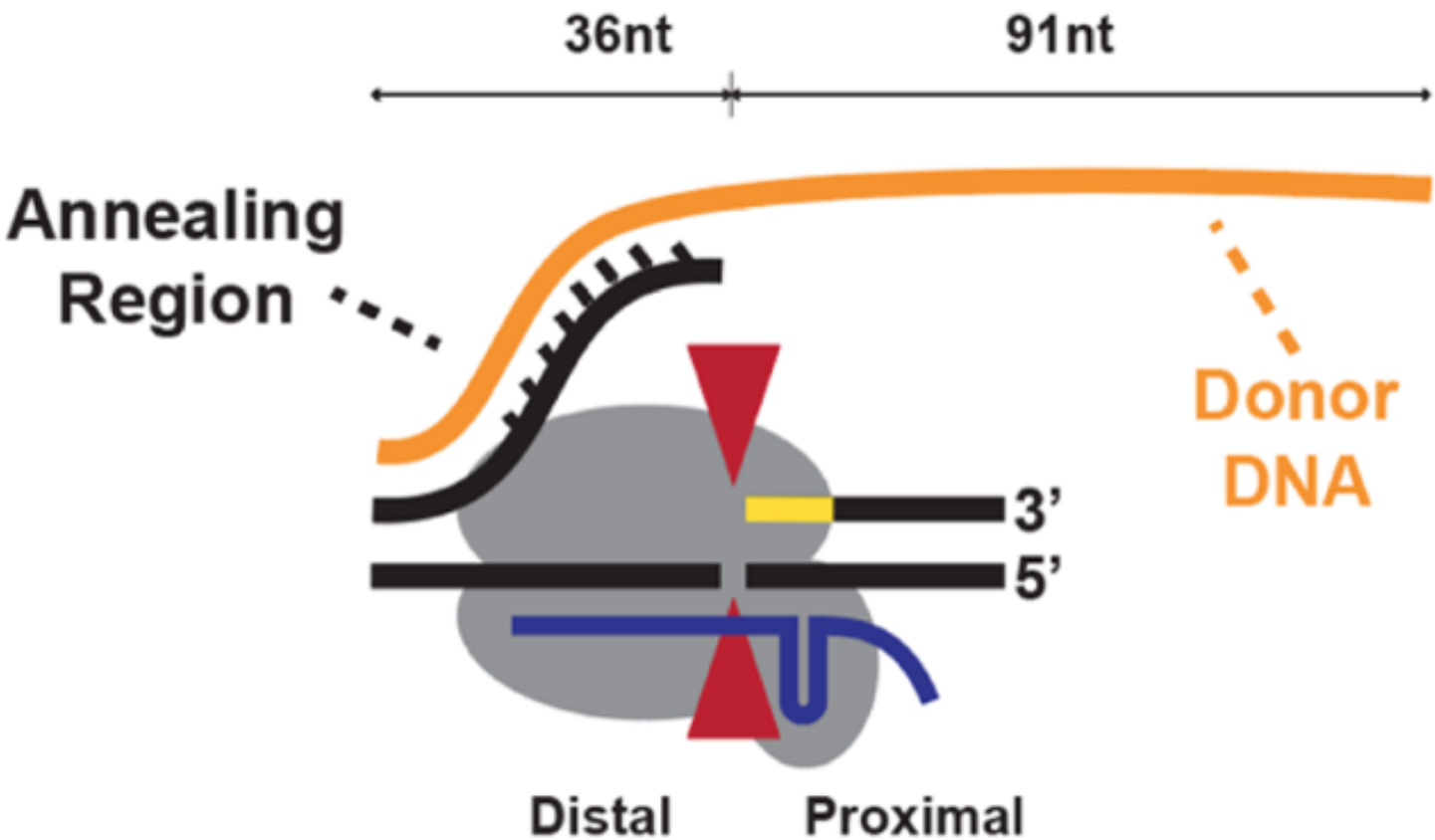


Figure 2: Single stranded donor DNA interaction with the Cas9-DNA duplex

dCas9 Enables Sequence Replacement Without Cutting

Our *in vitro* results also demonstrated that catalytically inactive dCas9 produced a "bubble" structure in uncut target DNA that was accessible to anneal with complementary single stranded DNA. We thus wondered if such a structure could drive sequence replacement in cells. Targeting three dCas9 molecules to a precisely spaced region allowed sequence replacement rates of approximately 1%. This is by no means a large number, but it was attained without any of the error-prone repair that normally accompanies Cas9 cutting. We still don't know the mechanism underlying dCas9 editing, but it could be very useful to tackle genetic diseases in which sequence replacement holds some kind of fitness advantage but error-prone repair of the gene would be disastrous (i.e. if breaking the gene is worse than leaving the mutation alone).

OPTIMIZING DONOR DNA FOR ENHANCED CRISPR GENOME EDITING (CONT'D)

CRISPR Delivery by Nucleofection

The *in vivo* editing experiments in this publication were performed using a technique called nucleofection, which introduces functional Cas9 protein and the targeting guide RNA into cells by electroporation. Results from our lab and others (Lin et al 2014, Richardson et al 2016, and Kim et al 2014) have shown that this technique supports extremely high frequencies of genome cutting and, when donor DNA is included in the nucleofection reaction, extremely high frequencies of HDR. The detailed protocol for nucleofection as well as the sequences of our validated editing reagents can be found on our plasmid pages. We encourage any labs performing genome editing experiments to try nucleofection as (in our hands) this technique supports gene disruption or gene replacement frequencies that are an order of magnitude greater than transfection approaches. Moreover, the cost of reagents is relatively modest, requiring expressed Cas9 protein, transcribed sgRNA, and optional donor DNA.

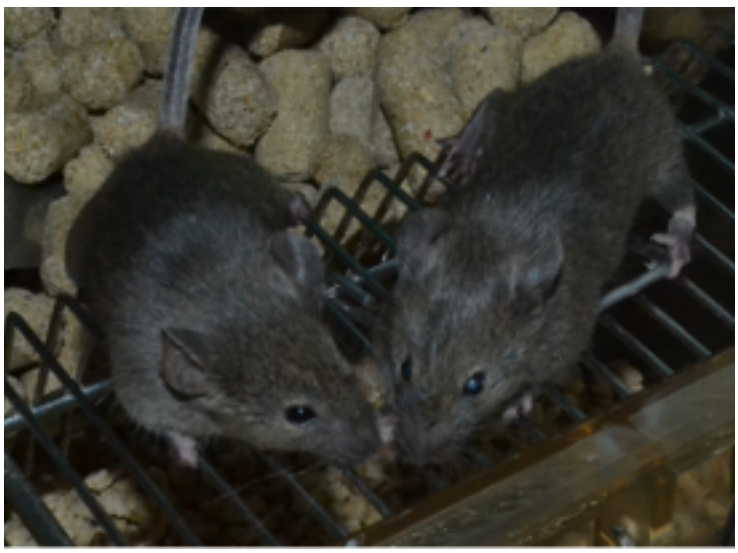
Further Reading

1. Richardson, Christopher D., et al. "Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA." Nature biotechnology (2016). PubMed PMID: 26789497.
2. Lin, Steven, et al. "Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery." Elife 3 (2015): e04766. PubMed PMID: 25497837. PubMed Central PMCID: PMC4383097.
3. Kim, Sojung, et al. "Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins." Genome research 24.6 (2014): 1012-1019. PubMed PMID: 24696461. PubMed Central PMCID: PMC4032847.

THERAPEUTIC APPLICATIONS OF CRISPR

USING CRISPR/CAS9 TO EDIT DISEASE OUT OF THE GENOME

By Kendall Morgan | January 7, 2014



CRISPR cured the mouse on the left of its cataracts. On the right, a control mouse with cataracts. Image attribution: Jinsong Li.

There can be no doubt that CRISPR/Cas9 technology has been a breakthrough for the genome-editing field. Two studies reported in *Cell Stem Cell* in December 2013 additionally show that this tool - already so useful in the laboratory - might also find its way to the clinic.

A team led by Jinsong Li from the Chinese Academy of Sciences found that mice with a dominant mutation in a gene that causes cataracts could be rescued by coinjection of Cas9 mRNA and a gRNA targeting the mutant allele into zygotes. An independent team led by Hans Clevers at the Hubrecht Institute in The Netherlands used the CRISPR/Cas9 genome editing system to correct the cystic fibrosis transmembrane conductor receptor (CFTR) by homologous recombination in cultured intestinal stem cells of patients with cystic fibrosis.

CRISPR to the Rescue

"I was really excited when I saw these two papers," said the University of California Berkeley's Jennifer Doudna, an Addgene depositor and CRISPR expert who was not involved in either study. "They are the first direct demonstration of use of genome editing to correct point mutations in two diseases with real phenotypic effect."

The difference with CRISPR/Cas9 in comparison to other genome editing techniques is its ease of use. "It's a very simple technology," Doudna explained, giving credit to Addgene for their role in making the technology so readily available. "That's why we're seeing adoption of the technology in different labs for different applications."

Li's team noted that CRISPR/Cas9 has been used to generate mutations in a wide range of organisms, but its potential to efficiently correct disease hadn't yet been realized. The researchers chose to explore its potential in mice with a dominant cataract disorder caused by a single copy of a gene known as Crygc. When all was said and done, the researchers had cured 24 mice of their disease.

Meanwhile, Hans Clevers and colleagues applied CRISPR to disease correction in adult stem cells isolated from two patients with cystic fibrosis. They were able to demonstrate functional correction of the gene in clonally expanded organoids.

"CRISPR in adult stem cells has huge potential, because the genomes of adult stem cells in organoid cultures are stable," said Gerald Schwank, first author of the cystic fibrosis report. "This is a big advantage over induced pluripotent stem cells, which need to go through a phase of crisis where they 'collect' a number of mutations."

Cystic fibrosis may not be the most likely candidate for gene therapy, however, because the disease affects multiple organ systems. Nevertheless, the findings are a great proof-of-principle for the use of CRISPR to correct single-gene conditions.

USING CRISPR/CAS9 TO EDIT DISEASE OUT OF THE GENOME (CONT'D)

Remaining Hurdles

"I think it is absolutely possible to use CRISPR to cure genetic disease in the near future," Li says, though there are some things to sort out first.

The potential for off-target effects is a major limitation. The procedure will also need to be perfected until the efficiency of repair approaches 100 percent.

"We need to be sure it's accurate enough and efficient enough for repair," Doudna said. "And it can't be too hard to apply. These papers suggest it might work well enough to consider as a therapeutic."

There are many fundamental questions about CRISPR/Cas biology still left to answer as well, she says, and those details will continue to be a primary preoccupation of her California lab and many others. "Those off-target effects are important," she said. "We also really need to get a handle on how cleavage sites are repaired." Those insights will be critical for scientists' ability to influence editing efficiency.

Further Reading

1. Wu, Yuxuan, et al. "Correction of a genetic disease in mouse via use of CRISPR-Cas9." Cell Stem Cell 13.6 (2013): 659-662. PubMed PMID: 24315440.
2. Schwank, Gerald, et al. "Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients." Cell Stem Cell 13.6 (2013): 653-658. PubMed PMID: 24315439.

TREATING DMD MICE WITH CRISPR GENE EDITING

By Mary Gearing | January, 2016

Having seen CRISPR's success in basic research, researchers are eager to apply it in a clinical setting. CRISPR is often used for animal germline modification, to repair or add in disease-causing mutations, but, until recently it hadn't been used to treat disease postnatally. Now, three papers published concurrently in Science have shown CRISPR can treat a genetic disease in a postnatal mouse model, an important proof of concept for future preclinical and clinical work.

Duchenne Muscular Dystrophy: The Most Common Inherited Disease

The genetic disease targeted is Duchenne muscular dystrophy (DMD), an X-linked recessive disorder affecting approximately 1 in 5000 males. DMD is caused primarily by frameshift mutations in dystrophin, a protein essential for proper muscle function. Without functioning dystrophin, an individual experiences progressive muscle wasting leading to death at around 30 years of age. Despite the amount of research conducted on DMD, there is still no good treatment.

The dystrophin gene is very large (79 exons), but much of the sequence appears to be nonessential. Within the mutational hotspot for dystrophin, exons 45-55, there are multiple common deletions that maintain the protein's reading frame, leading to the production of a smaller, but at least partially functional protein. Individuals with these mutations are often asymptomatic, or have only mild symptoms, a condition known as Becker muscular dystrophy (BMD).

Dystrophin's size makes it difficult to deliver via gene therapy, so researchers have set their sights on another approach. Since shorter forms of dystrophin can still be functional, exon skipping is a good option for DMD treatment. Clinical trials have used oligonucleotide exon skipping (OEN) to remove mutated exons from the dystrophin transcript. Unfortunately, the oligonucleotides only modestly improve muscle function, and they must be injected regularly.

Dystrophin and Genome Editing

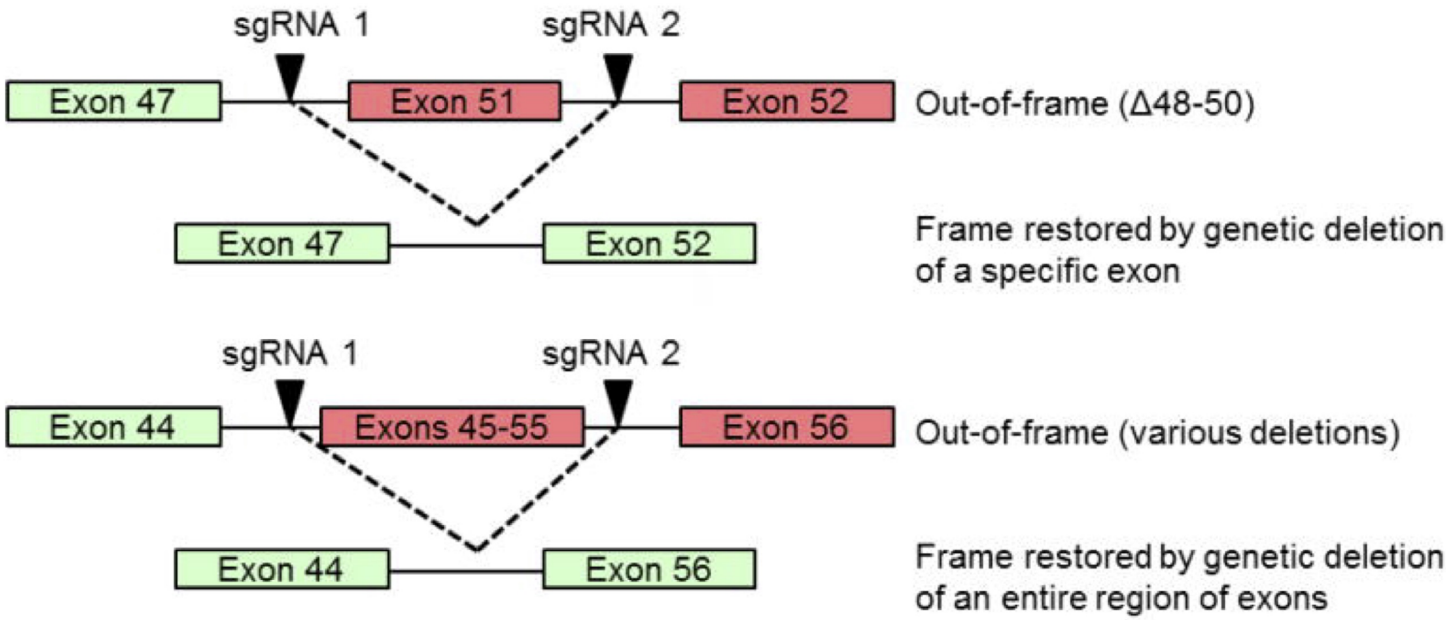
Since complex oligonucleotide treatment comes with many challenges, researchers have begun to explore genome editing approaches for exon skipping. Addgene depositor Charles Gersbach's lab used paired zinc finger nucleases to remove exon 51 in DMD patient myoblasts. They observed a 13% removal rate of exon 51, which resulted in appropriately localized dystrophin. In a subsequent study, they used CRISPR with two gRNAs to delete exon 51 or exons 45-55 in patient myoblasts; when injected into DMD mice, these cells expressed functional dystrophin.

Editing genes *in vivo* is of course much more difficult than in cell culture, but Ran et al. have shown that CRISPR and AAV can be used together for postnatal genome editing in mice. Could this approach work with dystrophin exon skipping? For such a therapy to be successful, multiple requirements must be met. First, CRISPR would need to be delivered to both cardiac and skeletal muscle cells, where precision editing of the dystrophin gene would take place, with minimal risk of off-target editing. In order for the therapy to persist over time, stem cell editing would be highly desirable. Should stem cell editing occur, the CRISPR components would only need to be expressed for a short period of time, which would prevent the accumulation of unwanted mutations over time.

DMD is a good choice for a CRISPR proof of concept treatment, as the disease is especially well suited for genome editing. The homology-directed repair (HDR) pathway is downregulated in mature tissues - no problem,

TREATING DMD MICE WITH CRISPR GENE EDITING (CONT'D)

as exon skipping proceeds through the non-homologous end joining (NHEJ) pathway. It's estimated that very little dystrophin correction (about 4%) is needed to see muscle improvement, with only 30% correction needed for normal function. Even low-frequency editing could make a huge difference in DMD, and it's estimated that exon skipping therapies would be applicable to 80% of DMD patients.



Exon skipping strategies for Duchenne muscular dystrophy. gRNAs can be designed to flank a single or multiple exons that, when deleted through NHEJ, will restore the dystrophin reading frame. gRNAs bind in intronic regions, so low-frequency indels will not cause frameshifts. Figure adapted from Osterout et al. under a CC-BY 4.0 license.

Long et al., Nelson et al., and Tabebordbar et al. each chose to try exon skipping in DMD mice, which have a mutation in dystrophin exon 23. Long et al. used SpCas9 in conjunction with AAV9, whereas Nelson et al. and Tabebordbar et al. used the shorter SaCas9 with AAV8 and AAV9, respectively. Although each study used slightly different methods, they each observed the same positive phenotypes in cardiac and skeletal muscle. Even when the genome editing frequency is low, the relative abundance of exon-skipped mRNA is high, likely because this mRNA is not subject to nonsense-mediated decay. Although Nelson et al. observed only 2% genome editing in one experiment, they found the exon-skipped transcript constituted 59% of total dystrophin mRNA, similar to the 39% observed by Tabebordbar et al. Long et al. show that the percentage of dystrophin-expressing muscle cells increases over time, and all three groups confirm dystrophin expression via Western blot, at <10% of the wild-type level. Muscle histology is improved, with vastly reduced inflammation and necrosis compared to unedited animals. In grip strength, specific force, and other muscle tests, muscle function is clearly improved, although not to wild-type levels. With regard to off-target effects, each group finds very low to no off-target activity at the ten highest-ranking predicted off-target sites.

Each paper characterizes additional, unique facets of CRISPR DMD therapy. Long et al. show that AAV-CRISPR does not cause obvious germline modification, an important finding given the controversy associated with such editing. Tabebordbar et al. show that muscle stem cells are modified by this approach, increasing the possibility that edits will persist long-term. In accordance with that result, Nelson et al. find that dystrophin restoration is maintained for at least six months.

TREATING DMD MICE WITH CRISPR GENE EDITING (CONT'D)

Future Directions and Obstacles to Clinical CRISPR Editing

Given the success of the CRISPR-mediated exon skipping approach in mice, DMD researchers are very excited. This approach may also be applicable to a number of rare genetic diseases caused by splicing defects, including ataxia telangiectasia, congenital disorder of glycosylation, and Niemann-Pick disease type C. Although the three DMD studies referenced here represent a great step forward for CRISPR gene therapy, it's important to realize that DMD is a simpler case than other genetic diseases we'd like to treat with CRISPR. As seen above, DMD can be treated with "one-size-fits-most" NHEJ-mediated editing, but most other diseases will require HDR-mediated precision editing tailored to smaller patient populations.

To bring DMD CRISPR therapy closer to the clinic, much work is still needed. First, CRISPR delivery must be optimized to:

1. Reach a high percentage of muscle cells throughout the body, especially stem cells
2. Remove any immunogenicity of the AAV vector

Once delivery has been optimized, it's important to ascertain how long the rescue phenotype will last, and more importantly, if it does extend lifespan. This work should be done in mouse and large animal models with mutations in human-relevant exons 45-55, rather than in the traditional exon 23-mutated mouse model. Potential off-target effects in muscle, as well as unwanted germline editing, must be rigorously examined over a long period of time. High fidelity Cas9s such as eSpCas9 and SpCas9-HF should be explored to reduce off-target editing. Short-term CRISPR expression would be especially desirable, as it would reduce the potential of off-target editing over time, but this approach would require robust stem cell editing to maintain the desired phenotype.

Precision genome editing faces the challenges above and more. One chief challenge is upregulating HDR in mature tissues, as this process is necessary to precisely edit point mutations. In addition to upregulating HDR, NHEJ must be downregulated to prevent the introduction of new mutations by CRISPR; editing exonic regions comes with much more risk than the intronic editing used in the exon skipping approaches here. In most genetic diseases, the mutational landscape is broad and varied, necessitating the development of many distinct CRISPR therapies to match these different mutations. For each therapy, effectiveness and off-target risk must be evaluated separately, increasing the time to clinical approval.

Even with the challenges precision editing faces, it's encouraging to see such progress in the more feasible case of DMD. If the safety and efficacy of this approach can be optimized, DMD could become one of the first diseases to be treated clinically with CRISPR. Many obstacles remain for CRISPR gene editing, but given the speed at which the technology is advancing, precision editing treatments may also be closer than we might expect.

Further Reading

1. Kaiser J. CRISPR helps heal mice with muscular dystrophy. Science. 2015 Dec 31. doi:10.1126/science.aae0169

2. Long C, Amoasii L, Mireault AA, McAnally JR, Li H, Sanchez-Ortiz E, Bhattacharyya S, Shelton JM, Bassel-Duby R, Olson EN. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science. 2015 Dec 31. pii: aad5725. PubMed PMID: 26721683.

3. Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Rivera RM, Madhavan S, Pan X, Ran

TREATING DMD MICE WITH CRISPR GENE EDITING (CONT'D)

FA, Yan WX, Asokan A, Zhang F, Duan D, Gersbach CA. *In vivo* genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science. 2015 Dec 31. pii: aad5143. PubMed PMID: 26721684.

4. Reardon S. Gene-editing summit supports some research in human embryos. Nature. 2015 Dec 3. doi:10.1038/nature.2015.18947

5. Siva K, Covello G, Denti MA. Exon-skipping antisense oligonucleotides to correct missplicing in neurogenetic diseases. Nucleic Acid Ther. 2014 Feb;24(1):69-86. PubMed PMID: 24506781. PubMed Central PMCID: PMC3922311.

6. Tabebordbar M, Zhu K, Cheng JK, Chew WL, Widrick JJ, Yan WX, Maesner C, Wu EY, Xiao R, Ran FA, Cong L, Zhang F, Vandenberghe LH, Church GM, Wagers AJ. *In vivo* gene editing in dystrophic mouse muscle and muscle stem cells. Science. 2015 Dec 31. pii: aad5177. PubMed PMID: 26721686.

7. Ousterout DG, Kabadi AM, Thakore PI, Perez-Pinera P, Brown MT, Majoros WH, Reddy TE, Gersbach CA. Correction of dystrophin expression in cells from Duchenne muscular dystrophy patients through genomic excision of exon 51 by zinc finger nucleases. Mol Ther. 2015 Mar;23(3):523-32. PubMed PMID: 25492562. PubMed Central PMCID: PMC4351462.

8. Ousterout DG, Kabadi AM, Thakore PI, Majoros WH, Reddy TE, Gersbach CA. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. Nat Commun. 2015 Feb 18;6:6244. PubMed PMID: 25692716. PubMed Central PMCID: PMC4335351.

9. 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. *In vivo* genome editing using Staphylococcus aureus Cas9. Nature. 2015 Apr 9;520(7546):186-91. doi: 10.1038/nature14299. Epub 2015 Apr 1. PubMed PMID: 25830891. PubMed Central PMCID: PMC4393360.

TARGETING HIV-1 WITH CRISPR

By Mary Gearing | August 26, 2015

Over 25 million people worldwide are currently infected with the lentivirus HIV-1. Today, HIV-1 can be controlled with antiviral therapies such that the virus is undetectable in the blood. But the virus doesn't completely disappear; it just hides in latently infected cells. To truly cure HIV-1, researchers need to vanquish these hidden viral reservoirs, and CRISPR may be the way to accomplish this tough job! Kamel Khalili's lab at Temple University has demonstrated two potential strategies for CRISPR-HIV therapeutics - one using dCas9-SAM to activate HIV-1 transcription and destroy infected cells, the other using wild-type Cas9 to remove the HIV-1 genome from infected cells. Read on to learn how CRISPR can take on HIV-1 *in vitro*, and what obstacles must be overcome for clinical success.

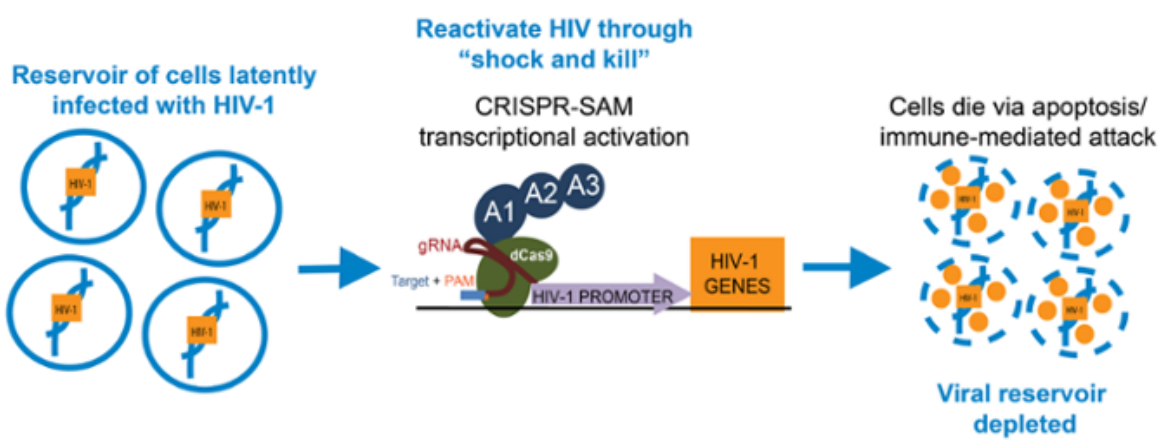
ART and HIV-1 Reservoirs

HIV-1 infects cells in the immune system, notably CD4+ T-cells, and eventually leads to acquired immunodeficiency syndrome (AIDS) in untreated individuals. Symptoms of AIDS include rapid weight loss and an increased risk of infection, including both common infections and opportunistic infections not usually seen in healthy individuals. Antiretroviral therapy (ART) can virtually eliminate plasma HIV-1, improving life expectancy and quality for HIV-1 patients. However, ART is not an HIV-1 cure. In patients who stop ART, viral levels will soon skyrocket back to pre-treatment levels due to the viral reservoirs located in latently infected cells. Although ART-treated patients lack plasma HIV-1, they are at increased risk for other chronic diseases, including dementia, gut disorders, nerve damage, and cardiac disease. This increased disease risk is attributed to latent HIV reservoirs, chronic inflammation, and negative metabolic effects of ART.

Why can't the body mount an immune response to destroy these reservoirs? Essentially, the immune system doesn't see any threat to which it should respond. Latently infected cells evade immune detection because they produce very little or no viral protein. To solve this problem, two strategies have been proposed.

The first, "shock and kill", aims to reactivate the latent HIV so that infected cells will produce viral proteins and die, either through cytotoxicity or the immune response. The second strategy is the simple removal of the HIV-1 genome from infected cells, an idea that became much more realistic with the advent of CRISPR.

Using CRISPR/Cas9 SAM to "Shock and Kill"

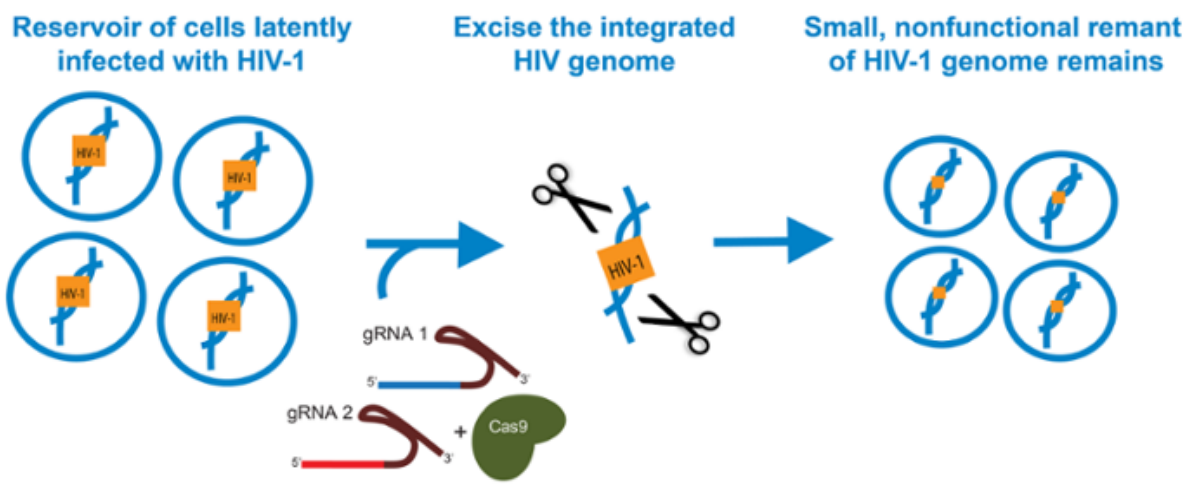


TARGETING HIV-1 WITH CRISPR (CONT'D)

To "shock and kill" the HIV-infected reservoir, researchers have previously used histone deacetylase (HDAC) inhibitors to increase transcription of the integrated HIV-1 genome. For a targeted approach, Zhang et al. turned to the CRISPR/Cas9 Synergistic Activation Mediators (SAM) system available from Addgene. This system uses the dCas9-VP64 fusion commonly used to activate transcription, but with two additional activation domains (MS2 and p65) to enhance transcriptional activation.

Zhang et al. designed multiple gRNAs to target the 5' long terminal repeat (LTR) of the HIV-1 genome, which acts as a promoter, and found that SAM targeting of the enhancer region near the NF-KB binding sites increased activation of HIV promoter-luciferase constructs. In multiple HIV-1 latent T cell lines, as well as a latent microglial line, treatment with CRISPR/Cas9 SAM increased the percentage of cells with activated HIV-1, measured via HIV-1 promoter-driven GFP expression. In cell lines that can produce HIV-1 toxic proteins, CRISPR/Cas9 SAM caused apoptosis, indicating a true "shock and kill" response.

Obstacles for CRISPR HIV-1 Therapy



Before the introduction of CRISPR, HIV-1 genome editing with ZFNs and TALENs focused primarily on disrupting CCR5, the receptor that HIV-1 uses to enter a cell. ZFNs and CRISPR/Cas9 have also been used to create mutations in genes essential to HIV-1 replication. However, this approach could permit some viral protein production, so researchers are also interested in precisely excising the HIV-1 genome from infected cells. ZFNs were used for proof-of-concept studies, but researchers believe that using CRISPR/Cas9 could improve efficiency and reduce the probability of off-target effects.

In such a therapy, CRISPR/Cas9 would need to 1.) excise the HIV-1 genome from every infected cell and 2.) prevent reinfection of those cells. Kaminski et al. designed gRNAs to target the HIV-1 5' and 3' LTRs and expressed them along with Cas9 in the T-cell line 2D10. PCR amplification and Sanger sequencing of pooled samples showed that most of the HIV-1 genome had been excised, leaving only a small portion of the LTRs joined together. Clonal Cas9/gRNA-expressing cell populations were also immune to HIV-1 reinfection.

Importantly, no off-target effects were detected. Kaminski et al. observed no negative effects on cell viability, cell cycle progression, or apoptosis with Cas9/gRNA expression. A pooled analysis found no evidence of Cas9 cleavage at predicted off-target sites with up to 7 mismatches to the gRNA target sequence.

TARGETING HIV-1 WITH CRISPR (CONT'D)

Kaminski et al. next tested their procedure in HIV-1 infected T-cells to see if they could be rescued by CRISPR. CD4+ T-cells were isolated from healthy individuals, expanded, and infected with HIV-1. For two HIV-1 strains, Cas9/gRNA lentiviral expression significantly reduced the HIV-1 copy number, although the efficiency differed from 48-100% between strains. In a similar experiment, CRISPR reduced the HIV-1 copy number by over 50% in isolated CD4+ T-cells from two HIV-1 infected patients. Although this is an exciting finding, it's important to note that these HIV-1 patients were ART-naive, so this experiment models the active HIV-1 infection period rather than the ART-induced viral control found in most HIV-1 patients.

Cutting HIV-1 Out of the Genome with CRISPR/Cas9

Both of these methods represent exciting advances in preclinical HIV-1 research that should be pursued in animal models, and it's difficult to tell if one approach might be more successful. "Shock and kill" has the advantage of killing cells infected with HIV-1 to deplete the viral reservoir, and this approach doesn't require a functional Cas9 nuclease, so there's no potential for aberrant DNA cleavage. Direct HIV-1 cleavage allows T-cells to survive, but preventing re-infection would require sustained expression of Cas9/gRNAs, which could result in higher rates of off-target cleavage.

In both cases, translating the *in vitro* work into an animal model comes with two key hurdles. The first is the delivery of the CRISPR machinery to all of the target cells. This goal may prove especially difficult given that the viral reservoir spans multiple organ systems. Both approaches may prove more successful early in the course of infection, when HIV-1 is limited to a subset of T-cells, but it is unclear if enough T-cells could be reached to ablate a well-established viral reservoir.

A second challenge for CRISPR HIV-1 therapies is sequence specificity. Whereas antiviral therapies target HIV-1 at the level of protein structure, CRISPR gRNAs require DNA sequence-specific binding. Patients' HIV-1 genomes will need to be sequenced to determine the optimal gRNAs for either "shock and kill" or viral excision approaches, and each of these gRNAs will need to be validated for high on-target binding and low off-target binding. It's also possible that HIV-1 may evolve resistance to CRISPR therapies through PAM or seed sequence mutations. Multiple gRNAs could be used to combat this problem, just as ART includes multiple drugs to lower the odds of developing resistance.

Shortly after Kaminski et al. published their results, Wang et al. showed that HIV-1 could escape from CRISPR/Cas9-induced modification targeting either the LTRs or essential genes. Many of these escape mutations were located near the Cas9 cut site, leading Wang et al. to conclude that some Cas9-derived indels may not ablate viral function, but rather promote resistance to CRISPR/Cas9. While these results are discouraging, it's important to note that these experiments were done in cell culture, and the cells infected with HIV-1 only after stable expression of Cas9/gRNA. Viral production is higher in this model than in latent HIV-1 infection, which would make it easier for an "escape" virus to propagate and infect adjacent cells. It will be important to determine if CRISPR/Cas9 escape occurs in an animal model. If these escape mutations are frequent, the dCas9-mediated "shock and kill" strategy may be a better option than direct cleavage.

Despite the potential difficulties in translating these findings into a therapy, these papers present tantalizing evidence that an HIV-1 cure may be within our reach. Similar studies have shown that CRISPR can be used to combat other viral infections, notably a mouse model of Hepatitis B, a disease that infects over 250 million people worldwide. As we've previously seen in the CRISPR field, the lessons learned in pursuing one technological application can benefit many others and accelerate the pace of research. We at Addgene hope to see the problem of *in vivo* CRISPR delivery become more tractable in the next few years, as it would open up

TARGETING HIV-1 WITH CRISPR (CONT'D)

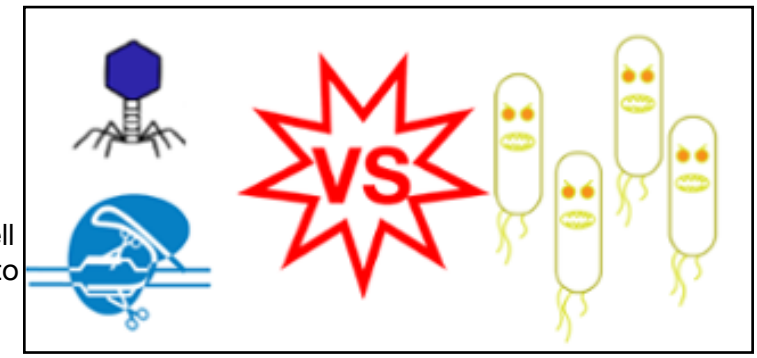
many new therapeutic possibilities for some of the world's most common diseases.

Further Reading

1. Zhang Y, Yin C, Zhang T, Li F, Yang W, Kaminski R, Fagan PR, Putatunda R, Young WB, Khalili K, Hu W. CRISPR/gRNA-directed synergistic activation mediator (SAM) induces specific, persistent and robust reactivation of the HIV-1 latent reservoirs. Sci Rep. 2015 Nov 5;5:16277. PubMed PMID: 26538064. PubMed Central PMCID: PMC4633726.
2. Kaminski R, Chen Y, Fischer T, Tedaldi E, Napoli A, Zhang Y, Karn J, Hu W, Khalili K. Elimination of HIV-1 Genomes from Human T-lymphoid Cells by CRISPR/Cas9 Gene Editing. Sci Rep. 2016 Mar 4;6:22555. PubMed PMID: 26939770. PubMed Central PMCID: PMC4778041.
3. Wang Z, Pan Q, Gendron P, Zhu W, Guo F, Cen S, Wainberg MA, Liang C. CRISPR/Cas9-Derived Mutations Both Inhibit HIV-1 Replication and Accelerate Viral Escape. Cell Rep. 2016 Apr 5. PubMed PMID: 27068471.
4. Ramanan V, Shlomai A, Cox DB, Schwartz RE, Michailidis E, Bhatta A, Scott DA, Zhang F, Rice CM, Bhatia SN. CRISPR/Cas9 cleavage of viral DNA efficiently suppresses hepatitis B virus. Sci Rep. 2015 Jun 2;5:10833. PubMed PMID: 26035283. PubMed Central PMCID: PMC4649911.
5. Khalili K, Kaminski R, Gordon J, Cosentino L, Hu W. Genome editing strategies: potential tools for eradicating HIV-1/AIDS. J Neurovirol. 2015 Jun;21(3):310-21. PubMed PMID: 25716921.

CRISPR ANTIMICROBIALS

By Mary Gearing | May 3, 2016



The crisis of antibiotic resistance is upon us, and the world is unprepared. Each year in the United States, two million people will be infected by antibiotic resistant bacteria. Even when researchers develop new antibiotics, the onset of resistance is swift, as few as five years after introduction. Current antibiotic strategies are nonspecific - they harm any bacterial cell without a resistance gene, allowing resistant bacteria to multiply, spreading their resistance genes throughout the bacterial population. But what if we could specifically target only virulent or antibiotic resistant bacteria with a weapon that they'll have less potential to become resistant to? CRISPR may provide a method for doing just that. While challenges remain in the delivery of these agents, CRISPR antimicrobials could become our newest line of defense against bacteria.

The Problem of Antibiotic Resistance

Most antibiotics target one of three essential processes: cell wall synthesis, DNA replication, or protein synthesis. Because they target such generic processes, antibiotics don't just affect the bacteria causing an infection, but also alter an individual's native, beneficial flora. Currently, there is no antibiotic strategy available to target only virulent or antibiotic resistant bacteria, and inappropriate antibiotic usage selects for resistant bacteria, allowing them to multiply and propagate their resistance genes.

Our society is hoping to combat antibiotic resistance using multiple approaches. Major restaurant chains including Chipotle and Panera Bread have pledged to eliminate the use of antibiotics by their suppliers, and public health and medical organizations are working to help educate the public about proper antibiotic usage. Although these steps are laudable, they likely won't solve the growing problem. Above all, we need new antimicrobials, ideally ones for which the development of resistance will be slow. CRISPR may be just the method needed to jumpstart the fight against antibiotic resistance.

Advantages of CRISPR Over Traditional Antibiotics

The potential advantages of CRISPR over traditional antibiotics are easy to see. As discussed in a previous section, most bacteria have poor non-homologous end joining machinery, so a CRISPR-induced double-stranded break (DSB) in the genome is lethal. If this DSB occurs on a plasmid, the plasmid will be eliminated from the bacterium, which may also induce cell death. CRISPR also enables increased specificity - rather than targeting a process essential to most bacteria, custom CRISPR antimicrobials can target specific sequences in a single virulent bacterial species, or even an antibiotic resistance gene. Importantly, native, non-pathogenic bacteria remain to help recolonize the niche, reducing the chance of an opportunistic infection like for organisms like *C. difficile*.

In 2014, Citorik et al. and Bikard et al. developed sequence-specific, antimicrobial, plasmid-based, CRISPR systems (Find the Citorik Plasmids here). In both cases, they used 1-2 CRISPR crRNAs targeting sequences found only in certain bacteria (Figure 1). Citorik et al. targeted beta-lactam and quinolone resistance genes in *E. coli* using plasmid and phagemid delivery systems. Phagemids are plasmids containing a phage origin of replication that can be packaged into replication-incompetent phage particles using a helper system. These CRISPR antimicrobials successfully resensitized a beta-lactam resistant population to the antibiotic, despite

CRISPR ANTIMICROBIALS (CONT'D)

this resistance gene being encoded by a high-copy plasmid. In the case of the genomically encoded quinolone resistance, mediated by a single base pair change in DNA gyrase, CRISPR phagemids were cytotoxic only for resistant bacteria.

Bikard et al. used a phagemid system to target virulent *S. aureus*, a common hospital-borne infection. Two rounds of phagemid treatment were sufficient to wipe out the virulent strain. Bacteria not killed by the first round of treatment either did not receive a phagemid, lost the phagemid or received a defective phagemid; Bikard et al. did not observe any mutations at the target site. In a mouse skin colonization model, phagemid treatment decreased the proportion of virulent S. aureus from 50% to 11% in just 24 hours.

Advantages of CRISPR Over Traditional Antibiotics

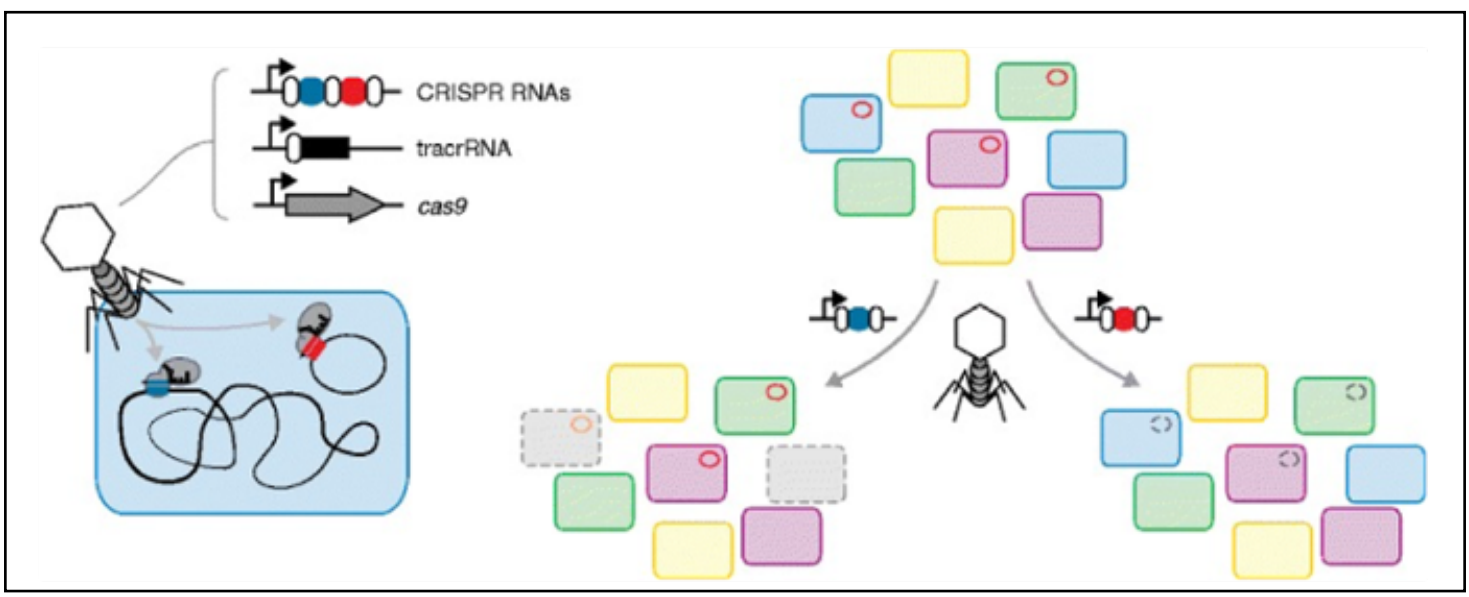


Figure 1. Single-phage strategies for bacterial CRISPR targeting. Phage-derived vectors carry CRISPR RNAs (crRNAs) targeting either genomic (blue) or plasmid-based sequences (red). Treating a mixed population, represented here by cells of different colors, specifically targets bacteria with the given genomic or plasmid sequence from a mixed population without affecting the other bacteria. Targeting a genomic sequence results in cell death. Plasmid targeting may result in plasmid loss and antibiotic resensitization, or may also cause cell death. Image from Beisel et al., licensed under a Creative Commons Attribution 4.0 International License.

Improving CRISPR Delivery and Potency with Phages

Neither group observed substantial resistance to the CRISPR antimicrobials, and they suggest that targeting multiple resistance/virulence factors may further reduce this likelihood. The key problem with these techniques is efficient delivery, as the antimicrobial must reach essentially every cell or else small pools of antibiotic resistant cells could repopulate the environment after treatment. The phagemid systems used in both papers are analogous to mammalian viral vectors in that the vectors are replication-deficient. Thus, a very large number of phagemids would be needed for an *in vivo* therapy.

Another strategy for CRISPR antimicrobials combines them with both replication-competent phages and antibiotics. Phage therapy fell out of fashion with the introduction of antibiotics, but the field is currently undergoing a renaissance. Yosef et al. designed a two-phage CRISPR system to resensitize antibiotic-resistant bacteria (Figure 2). The first phage introduced is lysogenic, integrating into the bacterial genome, and it carries CRISPR machinery targeting 1) the given resistance gene and 2) a second lytic phage. Bacteria that receive this

CRISPR ANTIMICROBIALS (CONT'D)

phage are now antibiotic sensitive, but lytic phage-resistant. When the lytic phage is introduced, it targets only the bacteria that are antibiotic resistant. Once these bacteria have been killed, antibiotics can be used to target the sensitive population. Yosef et al.'s proof-of-concept work targeted two distinct beta-lactam resistance genes, again with very high specificity. This double phage system both favors antibiotic sensitive bacteria and prevents sensitive bacteria from acquiring resistance genes. Yosef et al. suggest that it could be valuable in medical settings where antibiotic resistant bacteria have previously flourished.

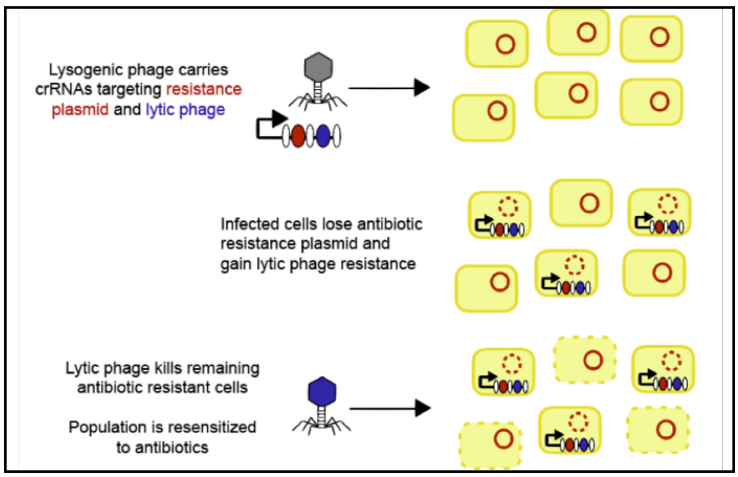


Figure 2. Two-phage strategies to resensitize bacteria to antibiotics. A lysogenic phage infects antibiotic resistant bacteria, causing them to lose their antibiotic resistance. These bacteria also gain resistance to a lytic phage through the introduction of a crRNA (shown in blue.) Infection with a lytic phage kills bacteria not previously infected by the lysogenic phage, rendering the population sensitive to antibiotics.

Phages could be used to deliver any of the three systems discussed above, with the caveat that phages are very diverse, and each phage has a narrow bacterial host range. To further potential clinical applications, Ando et al. recently engineered synthetic phages, each based on the well-studied T7 phage, to target various types of bacteria. In mixed populations, their phages were highly specific for certain bacteria, just like the CRISPR targeting methods. Phage cocktails could be used to target multiple types of pathogenic or drug-resistant bacteria.

It's a bit ironic that CRISPR, the bacterial immune system, may someday be used to target the hardest-to-kill bacteria. In using phage-derived vectors to carry CRISPR, we're turning the tables on the bacteria. Future research will show if this new strategy can help us win back ground in the fight against bacterial superbugs.

Further Reading

1. Citorik, Robert J., Mimee, Mark, and Timothy K. Lu. "Sequence-specific antimicrobials using efficiently-delivered RNA-guided nucleases." Nature Biotechnol. 32(11): 1141-5. (2014). PubMed: PMID 252409288. PubMed Central: PMCID PMC4237163.

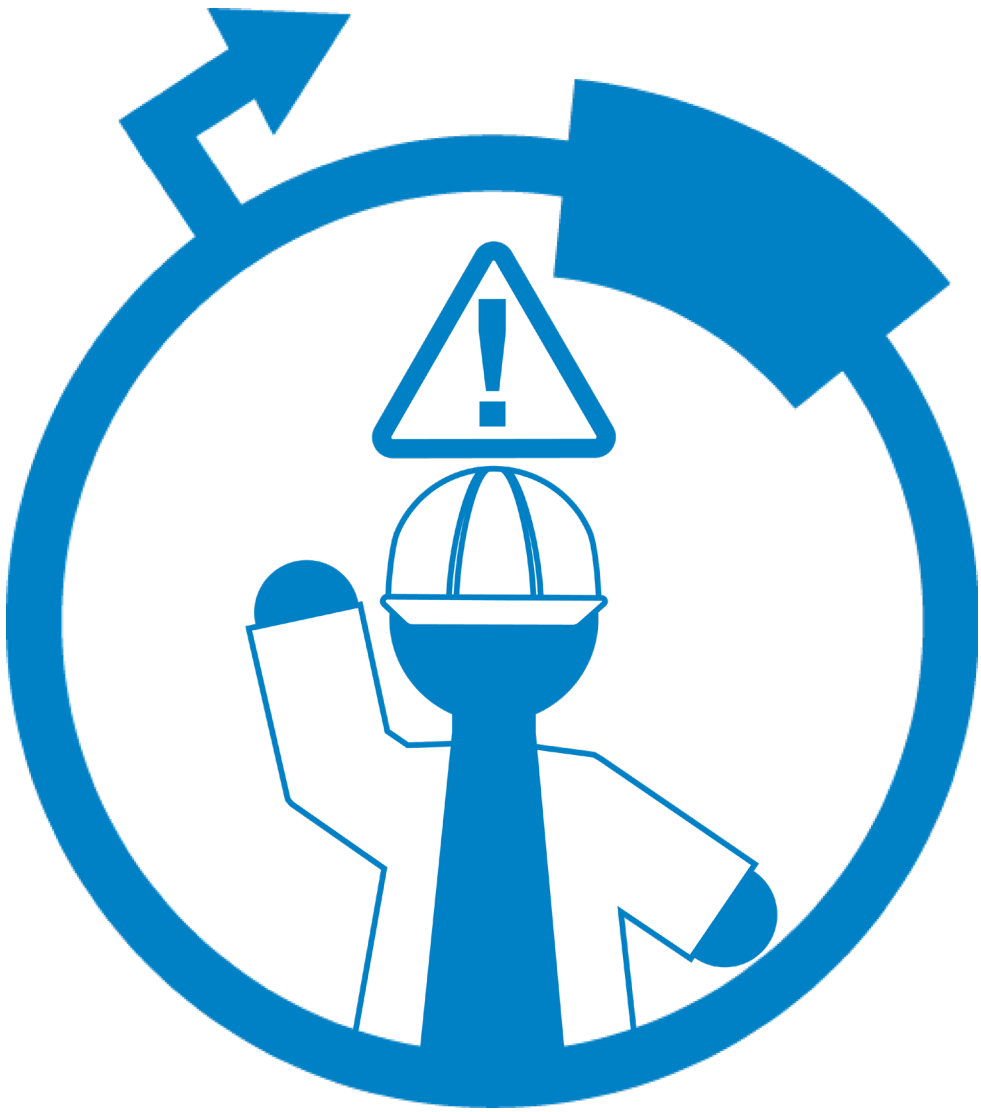
2. Bikard, David, et al. "Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials." Nature Biotechnol. 32(11): 1146-50. (2014). PubMed: PMID 25282355. PubMed Central: PMCID PMC4317352.

3. Beisel, Chase L., Ahmed A. Gomaa, and Rodolphe Barrangou. "A CRISPR design for next-generation antimicrobials." Genome Biol. 15(11):516. (2014). PubMed: PMID 25417800. PubMed Central: PMCID PMC4282009.

4. Yosef, Ido, Miriam Manor, Ruth Kiro, and Udi Qimron. "Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria." Proc Natl Acad Sci U S A. 9;112(23):7267-72. (2015). PubMed: PMID 26060300. PubMed Central: PMCID PMC4466736.

5. Ando, Hiroki, Sebastien Lemire, Diana P. Pires, and Timothy K. Lu. "Engineering Modular Viral Scaffolds for Targeted Bacterial Population Editing." Cell Syst. 1(3):187-196. (2015). PubMed: PMID 26973885.

CRISPR BIOSAFETY



GENERAL CRISPR SAFETY CONSIDERATIONS

By Mary Gearing | January, 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 at 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.

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.

GENERAL CRISPR SAFETY CONSIDERATIONS (CONT'D)

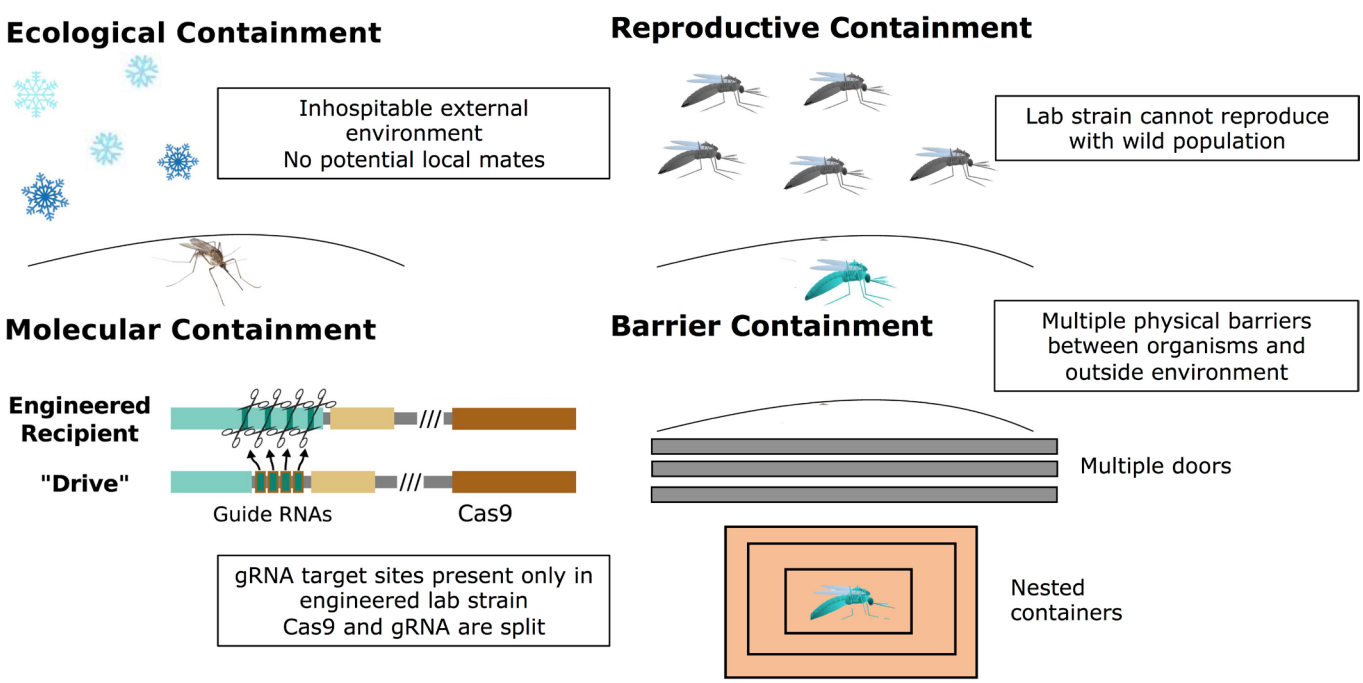
Gene Drive Containment

If you've kept up with CRISPR, chances are you've heard of gene drives. If not, check out our section 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.



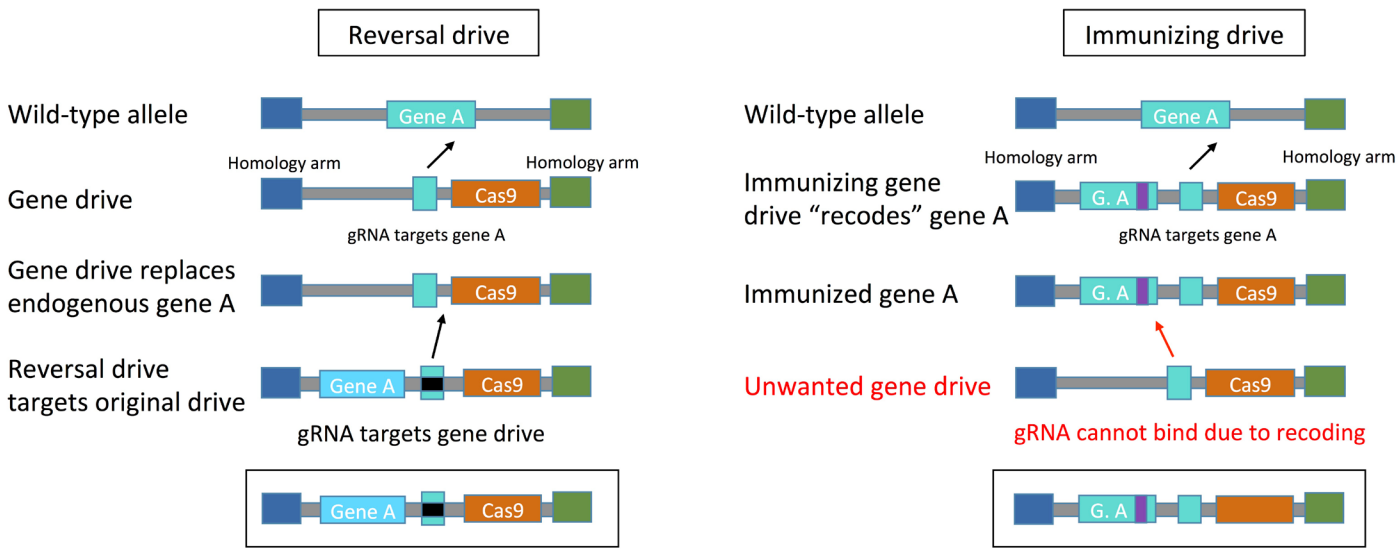
Visual depiction of common containment strategies. For optimal biocontainment, multiple strategies should be used together. Figure adapted from Esvelt et al. under a CC-BY 4.0 license.

GENERAL CRISPR SAFETY CONSIDERATIONS (CONT'D)

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 below).



- 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

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. under a CC-BY 4.0 license.

Specific questions about how biosafety relates to your research should be directed to your institution's Biosafety Committee.

Further Reading

1. Maddalo, Danilo, et al. "In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system." Nature 516(7531) (2014): 423-37. PubMed PMID: 25337876. PubMed Central PMCID: PMC4270925.
2. Gantz, Valentino M., et al. "Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*." Proc Natl Acad Sci U S A 112(49) (2015): E6736-43. PubMed PMID: 26598698. PubMed Central PMCID: PMC4679060.

GENERAL CRISPR SAFETY CONSIDERATIONS (CONT'D)

3. Akbari, Omar S., et al. "BIOSAFETY. Safeguarding gene drive experiments in the laboratory." Science 349(6251) (2015): 927-29. PubMed PMID: 26229113.
4. Esvelt, Kevin M., Andrea L. Smidler, Flaminia Catteruccia & George M. Church. "Concerning RNA-guided gene drives for the alteration of wild populations." eLife (2014): e03401. PubMed PMID: 25035423. PubMed Central PMCID: PMC4117217.
5. DiCarlo, James E., et al. "Safeguarding CRISPR-Cas9 gene drives in yeast." Nat Biotechnology 33(12) (2015): 1250-55. PubMed PMID: 26571100.

TRANSGENIC ORGANISMS, CAS9 GENE DRIVES, AND APPROPRIATE SAFEGUARDS

By Kevin Esvelt | May 22, 2015

This post was contributed by Kevin Esvelt, a Wyss Technology Development Fellow at the Wyss Institute and Harvard Medical School.

Scientists making transgenic organisms with Cas9 should be aware of the potential hazards of creating "gene drives" capable of spreading through wild populations. Whereas most genomic changes impose a fitness cost and are eliminated by natural selection, gene drives distort inheritance in their favor and consequently can spread even when costly.

If even a single organism carrying a synthetic gene drive were to escape the laboratory, the drive could eventually spread through the entire wild population with unpredictable ecological effects. Because the consequences of such a mistake would necessarily extend far beyond the laboratory and seriously damage public trust in scientists, experiments involving potential gene drives should be conducted with extreme caution.

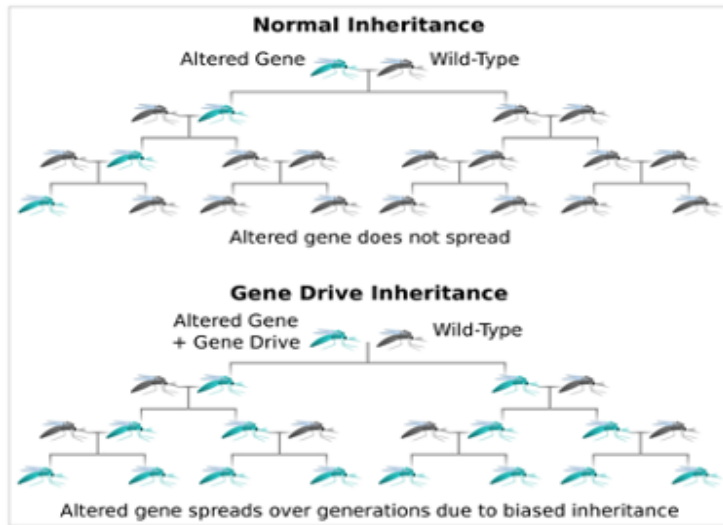


Image credit: Kevin Esvelt

Why Does Making Transgenic Organisms with Cas9 Risk Creating a Gene Drive?

The simplest way to make a gene drive is to insert an endonuclease gene within its own cut site (1). In heterozygotes, the endonuclease cuts the wild-type chromosome and its own gene is used as a repair template, thereby converting the heterozygote into a homozygote and ensuring that all future offspring inherit the endonuclease gene. As we pointed out last year (2), delivering a DNA cassette encoding the cas9 gene and single guide RNAs (sgRNAs) with appropriate flanking homology into a germline cell can create an RNA-guided gene drive. In principle, this event could happen even if there isn't any homology available. Consequently, the best way to avoid accidentally creating a gene drive is also the easiest: don't use a DNA vector that encodes both Cas9 and sgRNA.

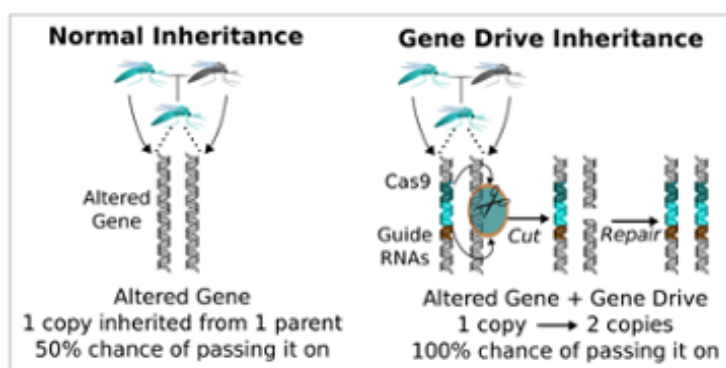


Image credit: Kevin Esvelt

What if I Want to Work with Gene Drives?

Then do so! It's a tremendously promising field that could address many pressing global problems in health, agriculture, and conservation. But please be aware that there are many easy-to-implement confinement strategies that are robust to human error. Used in combination with standard barrier protocols, they can reduce the risk of accidental escape to a negligible level.

TRANSGENIC ORGANISMS, CAS9 GENE DRIVES, AND APPROPRIATE SAFEGUARDS (CONT'D)

One major reason we chose to publish the concept and likely capabilities of RNA-guided gene drives in the summer of 2014 (after extensive discussions with biosecurity experts, ecologists, and national authorities) was to detail these robust and easy-to-implement confinement strategies (2). Our hope was that anyone else who came up with the same idea or wanted to build a gene drive in a model organism would find our work and consequently use these precautions in their own experiments (2,3). While that didn't work quite as well as anticipated (4), greater awareness will certainly improve safety.

What Safeguards and Confinement Strategies Are Available?

Molecular confinement involves building gene drives that can spread through populations of transgenic laboratory organisms but not wild organisms. For example, an sgRNA-only drive will spread exclusively through populations that already express Cas9 from an unlinked locus, while a Cas9+sgRNA drive targeting a synthetic sequence will only spread in transgenic laboratory populations with that sequence (2). Both methods are easy to implement and have been tested in yeast (5).

Ecological confinement involves performing experiments in a geographic area where escaped organisms won't be able to find mates (2). For example, ongoing experiments attempting to build gene drives in tropical mosquito vectors of diseases such as malaria and dengue are currently being performed in regions that don't have resident populations of the relevant mosquito species.

Reproductive confinement involves working with laboratory organisms that can't reproduce with wild ones. For example, *Drosophila* lines with compound autosomes are completely infertile when mated to wild fruit flies (6). It's also worth noting that gene drive experiments are less hazardous in organisms that seldom reproduce sexually because the drive must be much more efficient and minimally harmful in order to spread.

Barrier confinement seeks to keep the organisms in the laboratory. It varies by organism, but your local biosafety officer should be familiar with appropriate measures. Barriers should be a component of all gene drive confinement strategies, but they should not be relied on exclusively because historical studies of pathogen research have conclusively shown that barrier protocols are vulnerable to human error. And with gene drives, one mistake can be enough.

Reversal drives are designed to overwrite a previous gene drive and thereby undo the genetic changes driven by the earlier intervention (2). While an initial reversal drive cannot restore the exact original sequence, it can restore the original protein-coding sequence using a recoding strategy; a subsequent drive can restore the wild-type sequence (save for the residual sgRNAs and possibly Cas9 gene). An immunizing reversal drive is a variant that also spreads through the wild population and immunizes it against the first drive. Laboratories interested in building candidate gene drives intended for eventual release should consider building an appropriate immunizing reversal drive at the same time to mitigate the potential effects of an accidental release (3).

Which Gene Drive Confinement Strategies Should Be Used?

There are ongoing efforts to develop formal guidelines to answer this question. Until then, using multiple confinement strategies is strongly advised. The effects are multiplicative, and there are very few gene drive experiments that can't be performed with two or more confinement methods. Since molecular, ecological, and reproductive confinement typically require very little effort to implement (depending on the species in question), why not use them whenever applicable?

Against this, consider the cost of an accidental release. Science relies on popular support, which in turn

TRANSGENIC ORGANISMS, CAS9 GENE DRIVES, AND APPROPRIATE SAFEGUARDS (CONT'D)

depends on widespread trust in scientists to develop new technologies with care and humility. Nothing would damage that trust more than an accidental gene drive release, which would represent proof positive that at least some scientists could not be trusted to handle such a powerful technology responsibly. An accidental release would be particularly dangerous for molecular biology and genetics because the underlying advance enabling RNA-guided gene drives is Cas9 genome editing.

Finally, gene drives are an inherently collective technology: because they necessarily alter the shared environment, they must never be used without popular support (2). An accidental release could delay real-world applications against scourges such as malaria and dengue for many years. Since more than a thousand children died of malaria today, let's not risk a potential way to dramatically reduce that number for the sake of laboratory convenience.

Disclaimer: Recommendations concerning the appropriate degree of confinement are those of the author and do not represent a formal stance taken by Addgene or its staff. Additional information about gene drives can be found in the referenced publications.

Further Reading

1. Burt, Austin. "Site-specific selfish genes as tools for the control and genetic engineering of natural populations." Proceedings of the Royal Society of London B: Biological Sciences 270.1518 (2003): 921-928. PubMed PMID: 12803906. PubMed Central PMCID: PMC1691325.

2. Esvelt, Kevin M., et al. "Concerning RNA-guided gene drives for the alteration of wild populations." elife 3 (2014): e03401. PubMed PMID: 25035423. PubMed Central PMCID: PMC4117217.

3. Oye, Kenneth A., et al. "Regulating gene drives." Science 345.6197 (2014): 626-628. PubMed PMID: 25035410.

4. Gantz, Valentino M., and Ethan Bier. "The mutagenic chain reaction: A method for converting heterozygous to homozygous mutations." Science 348.6233 (2015): 442-444. PubMed PMID: 25908821.

5. DiCarlo, James E., et al. "RNA-guided gene drives can efficiently and reversibly bias inheritance in wild yeast." bioRxiv (2015): 013896. bioRxiv.

6. Fitz-Earle, M., D. G. Holm, and D. T. Suzuki. "Genetic control of insect populations: I. Cage studies of chromosome replacement by compound autosomes in *Drosophila melanogaster*." Genetics 74.3 (1973): 461-475. PubMed PMID: 4200686. PubMed Central PMCID: PMC1212962.

CRISPR CONCERNS BEYOND THE LAB BENCH

By Mary Gearing | January, 2016

CRISPR has proven to be a robust and flexible system for genome editing, and the biotechnology sector has certainly noticed. At least four companies (Caribou Biosciences, Editas Medicine, CRISPR Therapeutics, and Intellia Therapeutics) are based primarily on CRISPR technology, with numerous others using CRISPR in combination with other approaches. Since CRISPR technology has developed so quickly, there are still many attributes of the system that scientists don't understand, and ethical and societal implications that haven't been fully explored. CRISPR has the potential to revolutionize agriculture and gene therapy, but before adopting this system for non-research applications, one must fully consider the potential ramifications of this technology.

CRISPR's Agricultural and Therapeutic Applications

Compared to human editing, we don't hear as much about CRISPR in agriculture, but that doesn't mean there's not controversy. CRISPR may replace or supplement previous genome editing techniques used in agriculture, because it makes precise editing much easier. However, these edits are less traceable than traditional manipulations, which worries some regulators. Regulating CRISPR usage in plants and animals meant for human consumption will likely not be straightforward, and it remains to be seen how governments will adapt to this disruptive technology.

The potential therapeutic applications of CRISPR are even more appealing than the agricultural applications. Given the limited success of gene therapy, researchers have turned to CRISPR as the next great hope. Will it be possible to use CRISPR for postnatal genome editing to treat genetic disease? Three studies published concurrently in *Science* represent the first time CRISPR has been used to treat grown animals with a genetic disease. Duchenne muscular dystrophy (DMD), caused by various mutations in exons of the dystrophin protein, can be modeled in mice. Using complementary approaches in DMD mice, the teams were able to edit out the affected exon, preventing its incorporation into the dystrophin protein and improving muscle function. Importantly, the teams did not observe high levels of off-target cleavage throughout the genome. It's estimated that 80% of DMD sufferers could benefit from such a therapy, and these studies represent an important step toward clinical translation (read more in a previous section). Other monogenic diseases like sickle-cell anemia and beta-thalassemia are often mentioned as potential targets for CRISPR therapy.

The International Summit on Human Gene Editing

Given the potential and controversy surrounding CRISPR, the U.S. National Academy of Science and National Academy of Medicine have created a new initiative to help policy makers, researchers/clinicians, and the public understand human gene editing technology and its implications, with the goal of informing logical decision making about this technology. One important part of this plan is the International Summit on Human Gene Editing, held jointly by the U.S. Academies, the Chinese Academy of Science, and the U.K. Royal Society in December 2015.

Over a three-day program, the International Summit brought together experts to discuss various topics, including basic research, human germline editing, governmental regulation of editing, and national/international implications. At the conclusion of the program, the organizing committee issued a statement summarizing their conclusions. This statement divides human editing into three categories: basic/preclinical research, clinical somatic editing and clinical germline editing, with different recommendations for each category.

The committee stated that basic/preclinical research on both somatic and germline editing should proceed, but with appropriate regulations. In the case of embryo/germline editing, these cells should not be used to

CRISPR CONCERNS BEYOND THE LAB BENCH (CONT'D)

establish a pregnancy. Ideally, this balance will allow researchers to better understand the biology behind such modifications at a basic level. However, some scientists fear that modifying even embryos not intended for implantation will be highly controversial, as seen in previous debates over embryonic stem cell research.

Discussing clinical somatic applications, the committee notes that these applications have enormous promise and potential value. However, the statement emphasizes that risks, such as off-target editing, must be weighed along with benefits, and that this work should be done within regulatory frameworks for gene therapy. The recent development of high fidelity Cas9 enzymes eSpCas9 and SpCas9-HF, as well as further research into improving CRISPR editing, will help bring clinical somatic applications closer to reality. Editas Medicine is hoping to begin a CRISPR clinical trial as early as 2017, targeting a rare form of blindness called Leber congenital amaurosis.

Although *in vivo*, non-heritable gene editing has been relatively well-received, clinical germline modification remains highly controversial. The International Summit committee writes that "it would be irresponsible to proceed with any clinical use of germline editing unless and until (i) the relevant safety and efficacy issues have been resolved … and (ii) there is broad societal consensus about the appropriateness of the proposed application." The committee cites six key concerns associated with clinical germline editing (see table below) - these concerns are not simply health-related, but also include profound societal, moral, and ethical questions. The statement does not advocate a complete ban on germline editing, as some scientists had predicted, but suggests that the possibility of germline editing should be revisited periodically as "scientific knowledge advances and societal views evolve."

Concerns Associated with Clinical Human Genome Editing	Examples
Risks of inaccurate or incomplete editing	Mosaicism, harmful off-target mutations
Unpredicted effects of genetic edits	Interaction with other variants/the environment decreases fitness - e.g. sickle-cell disease and malaria
Irreversibility of genetic edits once introduced into a population	Edits cannot be tracked once introduced into a population - may combine unfavorably through reproduction
Implications for individuals and future generations carrying genetic alterations	Edited individuals may struggle with feelings of "otherness"
Potential for permanent genetic enhancements affecting social structure	Designer babies, genetic underclass of unmodified individuals
Moral and ethical considerations of altering human evolution	Potential conflicts with religion - idea of "playing God"

Six key concerns associated with human clinical germline editing raised by the International Summit of Human Gene Editing. Adapted from Section 3 of the International Summit Statement.

In the closing paragraph of the statement, the committee calls for an ongoing international forum related to human gene editing. This forum should help inform policymakers and allow nations to coordinate regulation of editing. In addition to including researchers and policymakers, the forum would seek expertise and opinions

CRISPR CONCERNS BEYOND THE LAB BENCH (CONT'D)

from many other groups, including health care providers, patients and their families, faith leaders, industry representatives, and the general public.

With CRISPR technology still in its infancy, we are a long way from societal consensus about how to use these tools, especially when it comes to clinical germline editing. While some researchers are pro-germline editing, others worry that the germline controversy will prevent acceptance of clinical somatic applications. As society struggles with these ethical dilemmas, clear scientific communication will be necessary to ensure that the public understands both the potential benefits and risks of human CRISPR applications. As the International Summit has recommended, open dialogue among researchers, government, health care, industry, religious leaders, and the general public is key to helping society develop a consensus on how to appropriately use this powerful technology.

Further Reading

1. International Summit on Human Gene Editing. On Human Editing: International Summit Statement. 2015 Dec 03.
2. Kaiser J. CRISPR helps heal mice with muscular dystrophy. Science. 2015 Dec 31. doi:10.1126/science.aae0169
3. Ledford H. CRISPR, the disruptor. Nature. 2015 June 04;522:20-24. PubMed PMID: 26040877.
4. Long C, Amoasii L, Mireault AA, McAnally JR, Li H, Sanchez-Ortiz E, Bhattacharyya S, Shelton JM, Bassel-Duby R, Olson EN. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science. 2015 Dec 31. pii: aad5725. PubMed PMID: 26721683.
5. Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Rivera RM, Madhavan S, Pan X, Ran FA, Yan WX, Asokan A, Zhang F, Duan D, Gersbach CA. *In vivo* genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science. 2015 Dec 31. pii: aad5143. PubMed PMID: 26721684.
6. Reardon S. Gene-editing summit supports some research in human embryos. Nature. 2015 Dec 3. doi:10.1038/nature.2015.18947
7. Regalado A. Engineering the perfect baby. MIT Technology Review. 2015 Mar 5.
8. Regalado A. CRISPR gene editing to be tested on people as early as 2017, says Editas. MIT Technology Review. 2015 Nov 5.
9. Regalado A. Patients favor changing the genes of the next generation with CRISPR. MIT Technology Review. 2015 Dec 2.
10. Tabebordbar M, Zhu K, Cheng JK, Chew WL, Widrick JJ, Yan WX, Maesner C, Wu EY, Xiao R, Ran FA, Cong L, Zhang F, Vandenberghe LH, Church GM, Wagers AJ. *In vivo* gene editing in dystrophic mouse muscle and muscle stem cells. Science. 2015 Dec 31. pii: aad5177. PubMed PMID: 26721686.

ACKNOWLEDGEMENTS AND FINAL WORDS

Special Thanks to Our Guest Contributors!

Alex Chavez

Alejandro (Alex) Chavez is a member of the Addgene Scientific Advisory Board with a particular interest in generating Cas9 based tools to enable facile control of DNA and RNA.

John Doench

John Doench is Associate Director of 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 and recently became a member of Addgene's Advisory Board.

Kevin Esvelt

An inventor of technologies that harness evolution, Kevin Esvelt studies ways of using molecular tools to alter populations of organisms and associated ecosystems. Because the environment is shared by everyone, he is deeply concerned with ensuring that gene drives and other collective technologies are developed safely and transparently. For more information, visit www.sculptingevolution.org.

Søren Hough

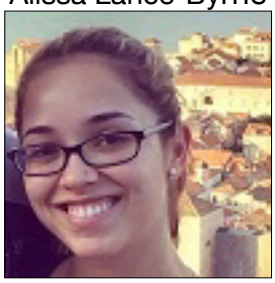
Søren Hough is the Head Science Writer at Desktop Genetics, a company that provides free sgRNA design tools for CRISPR experiments. He spent many years in several biochemistry laboratories and now works on communicating the latest science both inside and outside the field.

Aneesh Karve

Aneesh Karve is CTO at Quilt Data. Quilt is a collaborative database. Aneesh's research interests span machine learning, proteomics, and user interfaces for math.

ACKNOWLEDGEMENTS AND FINAL WORDS

Alissa Lance-Byrne

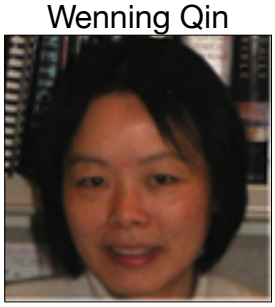


Alissa Lance-Byrne is a researcher at the Wyss Institute for Biologically Inspired Engineering. She is interested in synthetic and evolutionary biology.

Cameron R. MacPherson

Cameron R. MacPherson is Lead Data Scientist for the Milieu Intérieur project, Institut Pasteur, Paris. You can follow him on Twitter @CMacPhD.

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.

Chris Richardson

Chris Richardson is a Postdoctoral Researcher in Jacob Corn's lab investigating DNA repair pathways. Addtional Corn Lab members who helped with eBoook content include: Jordan Ray, Mark DeWitt, Gemma Cuire, and Jacob Corn.

ACKNOWLEDGEMENTS AND FINAL WORDS (CONT'D)

Marcelle Tuttle

Marcelle Tuttle is a medical student at Tufts University and former researcher at the Wyss Institute for Biologically Inspired Engineering.

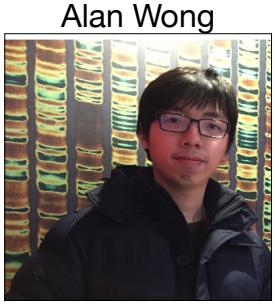
Haoyi Wang

Haoyi Wang is a professor at the Institute of Zoology, Chinese Academy of Sciences, and he is interested in using CRISPR to develop therapeutic applications that better human health.

Jordan Ward

Jordan Ward is a postdoctoral fellow with Keith Yamamoto at UCSF, and is currently studying gene regulation during nematode molting while preparing to transition to his own independent group. He received his PhD in Biochemistry from Cancer Research UK/ University of London studying mechanisms of DNA repair in the lab of Professor Simon Boulton. Learn more at his webpage or follow him on Twitter @Jordan_D_Ward.

Alan Wong



Alan Wong is currently an Assistant Professor at the School of Biomedical Sciences of the University of Hong Kong. He is particularly interested in developing and applying new technologies to interrogate and understand complex biological systems. He can be contacted at aslw@hku.hk.

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.

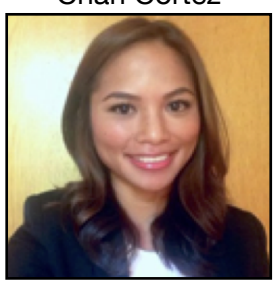
ACKNOWLEDGEMENTS AND FINAL WORDS (CONT'D)

Samantha Young

Samantha Young is a Medical Writer with a PhD in reproductive biology and genetics. She loves to communicate the latest and greatest in scientific content.

Contributions to This eBook Have Been Made by the Following Addgenies Past and Present:

Chari Cortez



Chari was previously a Scientist at Addgene. Her primary research experiences are 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 Fan is a cofounder and Chief Scientific Officer at Addgene. She also cofounded LabLife, a lab management software company. She loves that her job at Addgene connects her with scientists around the world who are working on everything from cancer to agriculture.

Tyler J. Ford

Tyler J. Ford is an Outreach Scientist at Addgene. His professional duties include helping maintain the Addgene blog (blog.addgene.org), talking to people about Addgene, and improving Addgene's services. His non-professional duties include running, biking, drawing, hiking, playing tennis, reading, and writing. Follow him on Twitter @TyFordFever.

Mary Gearing

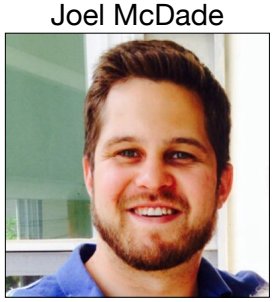
Mary Gearing is a Scientist at Addgene. She got her start as a Science Communications Intern writing for the Addgene blog and website. As a full-time Addgenie, she still enjoys blogging about CRISPR and other cool plasmids! You can follow Mary on Twitter at @megearing.

ACKNOWLEDGEMENTS AND FINAL WORDS (CONT'D)

Caroline LaManna

Caroline is currently the Associate Director of Scientific Outreach at Addgene. She received her Ph.D. in Biomedical Engineering from Boston University. Her goal is to make it easier for scientists to share reagents, communicate their research to others, and teach science in the community.

Joel McDade



Joel R. McDade is a Business Development Manager at Addgene whose interests include Genome Editing and Science Education.

Kendall Morgan

Kendall K. Morgan, Ph.D., is a science writer based in North Carolina. She writes about science, health, and medicine for many organizations and publications with the goal to help researchers share their work with each other and the world.

Marcy Patrick

Marcy is a Senior Scientist at Addgene. She received her Ph.D. in Microbiology and Immunology from the University of Michigan. She loves that her job allows her to interact with scientists from all over the world and help them easily share their reagents.

Brook Pyhtila

Brook Pyhtila is a Scientific Curator at Addgene. She received her PhD at Stanford and did her post-doctoral work at Duke. She enjoys contributing to Addgene's database and web content so that scientists can more easily do their research.

ACKNOWLEDGEMENTS AND FINAL WORDS (CONT'D)

Nicole Waxmonsky

Nicole is an Outreach Scientist and Curator at Addgene. When not riding her bike or playing broomball, Nicole enjoys using the web to increase scientific sharing and to improve access to basic science information.

If you have any questions, comments or suggestions about how Addgene can improve its educational content, please contact us at blog@addgene.org.

Looking for more information about Addgene?
Connect with us!

www.addgene.org

blog.addgene.org

www.facebook.com/addgene

www.twitter.com/addgene

www.linkedin.com/company/addgene

https://www.youtube.com/user/addgenemedia

http://blog.addgene.org/topic/podcast