MERCK

Sigma-Aldrich®
Lab & Production Materials

Gene Editing 101

Contents

Sigma-Aldrich®
Lab & Production Materials

Understanding CRISPR

CRISPR Gene-editing Explained

The CRISPR (Clustered, Regularly Interspaced, Short Palindromic Repeats)–Cas system evolved in microbes as a defence mechanism to protect against invasive phages. This system is the basis for a set of gene-editing tools that are enabling advances in a wide range of research interests from health and diagnostics to agriculture and energy.

Gene editing is a specific and targeted change to a DNA sequence and involves the addition, removal or modification of the DNA. The CRISPR system accomplishes gene editing through two main components:

1. a guide RNA (gRNA).
2. a bacterially-derived nuclease (e.g. Cas9).

The gRNA is a specific RNA sequence designed to recognize and direct the nuclease to the target DNA region and consists of two parts: CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA). crRNA is a 17-20 nucleotide sequence complementary to the target DNA and therefore varies depending on the target gene. In contrast, the tracrRNA is an invariable sequence that serves as a scaffold attaching the Cas nuclease to the crRNA.

The first CRISPR editing systems utilized a two-part gRNA complex consisting of a separate crRNA and tracrRNA, but it is now standard to use a single gRNA (sgRNA) approach, which combines the crRNA and the tracrRNA into one RNA molecule. Figure 1 shows the general schematic of the CRISPR gene-editing complex.

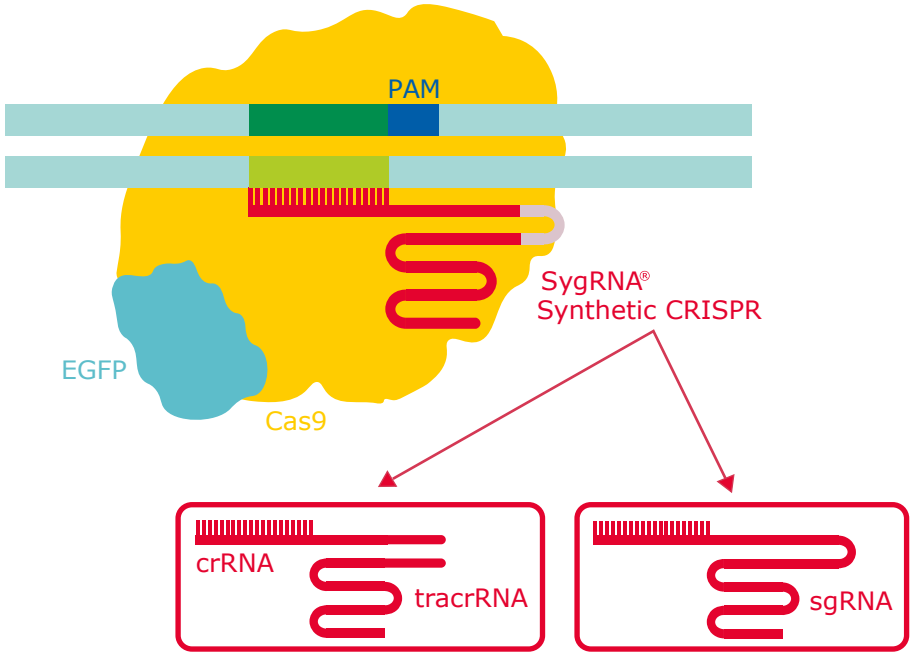


Figure 1.
Schematic diagram of the CRISPR gene editing system showing the different types of synthetic gRNA for CRISPR.
Credit: SigmaAldrich.com/CRISPR

Mechanism of CRISPR Gene-Editing

The CRISPR complex performs gene editing in three distinct steps: targeting, cleavage and repair.

Targeting of the CRISPR Complex

The crRNA is, by design, complementary to the target DNA and allows the gRNA to guide the CRISPR complex to the correct genomic location to perform gene editing. For successful binding of the CRISPR complex to the DNA, a protospacer adjacent motif (PAM) must be present downstream of the target site.

PAMs are short sequences, between 3 and 8 nucleotides in length and the precise nucleotide sequence of a PAM varies depending on the nuclease used, with nucleases isolated from different species requiring different PAMs. [1] The most commonly used nuclease, Cas9 derived from *Streptococcus pyogenes* (spCas9), recognizes a PAM sequence of 5'-NGG-3' (where 'N' is any nucleotide).

Cleavage of the DNA

Once the CRISPR complex arrives at and binds the target location, the nuclease can cut the DNA. The CRISPR complex contains two separate nuclease domains, each cutting one specific strand of DNA. The HNH nuclease domain cleaves the strand complementary to the gRNA, while the RuvC nuclease domain is responsible for cutting the non-complimentary strand.

The two nuclease domains work in unison to create the double strand break (DSB), which occurs three nucleotides upstream of the PAM.

Repair of DSB

Once the nuclease cleaves the DNA, the native cellular DNA repair machinery attempts to repair the DSB through one of two mechanisms:

1. non-homologous end-joining (NHEJ).
2. homology-directed repair (HDR).

Gene Knockout Via NHEJ

NHEJ tends to be the primary cellular DSB repair mechanism. It can introduce insertion or deletion errors (indel) into DNA as the DNA ends are re-ligated in the absence of a homologous DNA template. Indels resulting in frameshift mutations and premature stop codons can produce a loss-of-function (LOF) mutation, and this is the primary means by which CRISPR is used to disrupt (knockout) a gene.

Gene Knock-In Via HDR

CRISPR can utilize HDR to perform the replacement and expression of a specific genetic sequence (knock-in). In addition to the main CRISPR components, HDR-mediated CRISPR editing requires a DNA donor template containing the new desired sequence flanked by regions of homology. Introduction of this donor template, along with the CRISPR components, allows the cells to repair the DSB via homologous recombination. The result is the incorporation of the new sequence into the target gene.

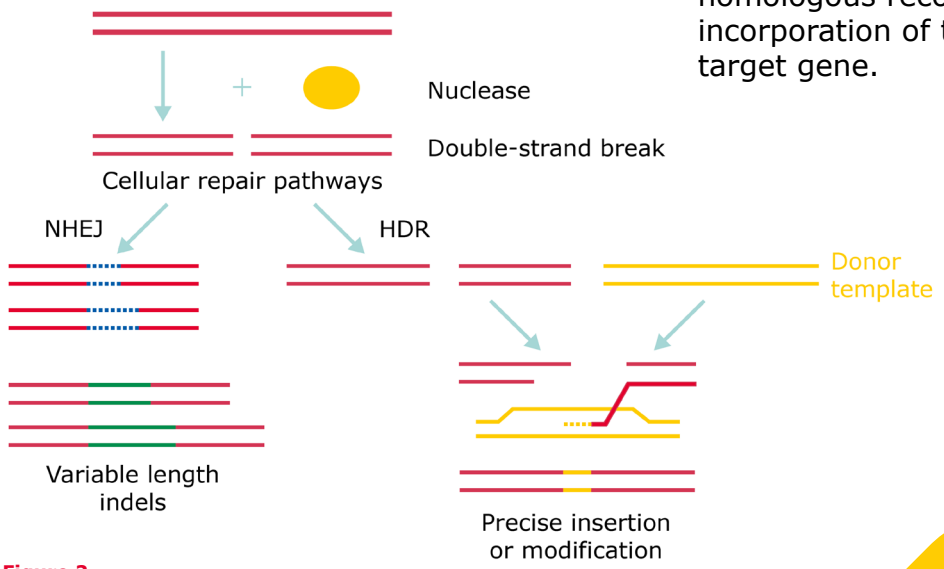


Figure 2.
Nuclease-mediated DSB are repaired using the native cellular DNA repair machinery via NHEJ or HDR.

Sigma-Aldrich®
Lab & Production Materials

Designing and Performing CRISPR

Guide RNA (gRNA) Design

Designing effective gRNAs, that enable efficient editing while minimizing unwanted off-target effects, involves the consideration of multiple elements.

Presence of a PAM Motif

CRISPR nucleases require PAMs to cleave their target DNA. Therefore, a PAM specific to the nuclease used must be present immediately downstream of the intended gRNA binding site.

GC-content

The GC content of gRNAs can affect the activity of the CRISPR complex. GC contents that are too low or too high can result in decreased activity. [2,3] A GC content of 40-60% is optimal for gRNAs. [2]

Chromatin Accessibility

Chromatin accessibility is a significant determinant of sgRNA binding *in vivo* with successful binding occurring more frequently in regions of DNA with open chromatin. [4]

Target Location

Successful generation of a LOF mutation requires the gRNA to target an exon that is essential for protein function.

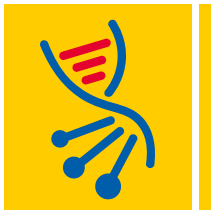
Off-target Complementarity

Ideal gRNA sequences are unique to the target DNA. However, gRNAs may still bind other regions, even if complementarity is not 100%. When possible, select gRNAs with at least 3 base pairs of mismatch from any other gRNA sequences in the genome.

Purchase predesigned CRISPR gRNAs

Our guide RNAs are guaranteed predesigned CRISPR gRNAs targeting a range of human and mouse genes.

CRISPR Design Tools



CRISPR Design

The CRISPR design tool allows the creation of custom gRNAs against genes, microRNA and long non-coding RNA using genome information from 25 different species.

Off-target Search

Determining if designed gRNAs have potential binding sites other than the target DNA reduces the off-target effects of custom gRNAs.

Donor Design

The creation of donor DNA allows CRISPR to generate precise gene knock-ins.

Considerations When Choosing Cas9 Format

There are three main formats for delivering CRISPR nucleases:

1. DNA plasmid.
2. mRNA.
3. RNP (ribonucleoprotein).

When selecting a CRISPR delivery format, it is essential to consider the transfection method to be used, efficiency, cost, specificity, whether you want transient or stable expression and how the delivery will be validated.

Transient methods work best for individual gene knockouts as they are associated with fewer off-target effects. In contrast, stable gRNA integration is needed for large-scale screening applications to enable recovery and quantification at the end of the screening process.

Selection marker genes, incorporated into the vector that expresses the CRISPR components, allow validation of the delivery of CRISPR components into the cell and can offer a mechanism to isolate only those cells that contain the CRISPR machinery. There are two primary selection markers used:

1. Fluorescent proteins. These allow the enrichment of transduced cells via fluorescence-activated cell sorting (FACS).
2. Antibiotic resistance genes. These enable the selection of the transduced cells using an appropriate antibiotic.

DNA Plasmids

The nuclease and gRNA can be encoded together in a single DNA plasmid or separately in two individual plasmids. This delivery method is simple, cost-effective and can be transfected into cells using the majority of available transfection methods.

CRISPR Plasmids can express both fluorophores and antibiotic resistance genes, allowing either of these methods to validate successful delivery of CRISPR components or to isolate successfully transfected cells.

It is possible to use fluorophore and antibiotic resistance selection markers in combination, as plasmids can encode both genes simultaneously.

mRNA

In vitro translated Cas9 mRNA coupled with a purified gRNA is an alternative to DNA plasmids. This option is more expensive, and the choice of nuclease is limited, however, using nuclease in an mRNA format overcomes issues with promoter compatibility and random integration of the CRISPR machinery into the host genome.

Antibiotic selection is not possible for nucleases in mRNA format. Still, custom fluorophore-tagged gRNAs are available, which allow identification and selection of cells containing CRISPR components. CRISPR nuclease mRNA can be delivered into target cells using many physical transfection methods, although the range of options is more limited than for plasmid formats.

RNPs

RNP complexes couple synthetic gRNAs with a Cas9 protein. They offer faster gene editing because the functional nuclease is immediately available in the cell. RNP complexes are quickly degraded by cellular proteases making the editing activity short-lived. The reduced time of available RNP in the cells also reduces off-target effects. Overall, RNP provides a practical and straightforward method for achieving high levels of gene knockouts with efficiency rates reaching 70-80%. [5]

Purified nuclease protein is more expensive than plasmid or mRNA formats, and there are fewer suitable transfection methods than for plasmid transfections.

The clearance of RNPs from the cell is observable in real-time in transfected cells using tagged RNP complexes such as Cas9-GFP Fusion Proteins, as well as fluorophore-tagged gRNAs.

Transfection Methods for CRISPR

There are multiple ways to deliver CRISPR nucleases (e.g., Cas9) and gRNAs into cells, including lentiviral transduction, PiggyBac integration, and transient transfection (DNA, RNA, or RNP).

Transfection methods for the delivery of CRISPR cargo into target cells fall into one of three main categories:

1. Physical delivery methods.
2. Viral vectors.
3. Non-viral delivery vehicles.

Table 1 describes these methods in more detail.

The choice of delivery method depends on the cargo format, cell type used, desired efficiency, required throughput, and cost.

CRISPR Format

The cargo type used will limit the transfection method available. Plasmid formats can be delivered using the majority of transfection methods including electroporation, viral and lipid delivery methods. In contrast, mRNA and RNP formats are more limited, with viral and other methods being unsuitable (Table 1).

Cell Type

The suitability of a transfection method is dependent on if the delivery is performed *in vitro* or *in vivo*. For *in vitro* delivery, immortalized cell lines are amenable to most transfection methods, whereas primary cells such as T cells are more difficult to transfect. See CRISPR Cell Choice for more details on cell types.

Transfection Efficiency and Throughput

Delivery methods often exhibit a trade-off between the proportion of cells that can be successfully transfected (transfection efficiency) and the total number of cells that are transfected (throughput). Microinjection is highly efficient as every cell will be individually injected with the CRISPR complex, but the method is very low throughput. Therefore, microinjection is appropriate for single-gene CRISPR experiments analyzed using methods requiring only small cell numbers but unsuitable for large-scale screening studies.

Table 1.
Summary of main CRISPR delivery methods. [6]

Delivery method	Delivery format	Advantages	Disadvantages
Physical Delivery Methods			
Microinjection	Plasmid; mRNA; RNP	High efficiency	Requires skill and precision; low throughput
Electroporation	Plasmid; mRNA	A well-established method; allows delivery to many cells simultaneously	Not all cell types are amenable
Nucleofection	Plasmid; mRNA	Allows delivery directly into the nucleus	Requires specialist kits
Hydrodynamic delivery	Plasmid; RNP	Non-viral; low cost; ease of use	Delivery is random; limited cell type
Viral Delivery Vectors			
Adenovirus	Plasmid	High efficiency	Immunogenic; cargo size is limited
AAV	Plasmid	Low immunogenicity	Cargo size is limited
Lentiviral	Plasmid	Persistent expression	Prone to genomic integration and mutation induction
Non-Viral Delivery Vehicles			
Liposomes/lipoplexes	Plasmid; mRNA; RNP	Virus-free; ease of use; low cost	Degradation of cargo by endosomal pathway
Cell-penetrating peptides	RNP	Virus-free, delivery of intact protein	Non-uniform penetration

CRISPR Cell Choice

There are several factors to consider when choosing an appropriate cell type for a CRISPR experiment, including transfection efficiency, haplotype and biological relevance of the model.

Transfection Efficiency

The transfection efficiency of cells correlates directly with the proliferation rate, as the nuclear membrane dissolves during division allowing transfected CRISPR components to enter the nucleus. Immortalized cell lines proliferate profusely in culture, are easy to transfect and exhibit high transfection efficiencies. These qualities make immortalized cells ideal for high-throughput experiments, including large-scale CRISPR screens.

Primary cells, such as T cells, are hard to transfect as they do not exhibit the same level of proliferation as immortalized cells. Transfection efficiency is further reduced in primary cells because the viral vector elicits an immune response or the plasmid delivery ends up killing the target cell. In such cases, an RNP system is used.

Haplotype

The ploidy of your chosen cell line dictates the number of mutations needed to obtain a homozygous LOF mutant. Many transformed or cancer cell lines possess more complex genomic configurations (e.g., triploid), making it more challenging to achieve a gene knockout. Haploid cell lines, such as HAP1 cells, enable the generation of a complete knockdown or homozygous mutation, as there is only one copy of the gene to target.

Biological Relevance

Cells should be selected to reflect the disease or system studied to ensure results are biologically relevant. Immortalized cell lines accumulate mutations over time, and their complex haplotypes make them less representative of normal biology and inadequate models of disease. Primary cells, such as T cells or cancer cells obtained from patients, provide results that more easily translate to the tissue or disease studied. Table 2 provides a summary of the benefits and considerations of different cell types.

Table 2.
Considerations when choosing cells for CRISPR gene editing.

Cell Type	Examples	Considerations for delivery format	Advantages	Limitations
Immortalized cell lines	HeLa HEK293	Suitable for Plasmid, mRNA or RNP formats	High transfection efficiency; *in vitro* proliferation	Complex haplotypes; limited biological relevance
Primary Cells	T cells Fibroblasts Epithelial cells	Viral transfection of plasmid may elicit an immune response (T cells), RNP format is preferred	High biological relevance	Difficult to transfect; may be unsuitable for screens; primary immune cells may exhibit immune response against transfected DNA; limited proliferation *in vitro*
Stem cells	iPSC ASC ESC	Suitable for plasmid, mRNA, or RNP formats	High biological relevance; *in vitro* proliferation	Difficult to transfect; may be unsuitable for screens
Haploid (or near-haploid) cell lines	HAP1 eHAP1 KBM-7	Suitable for plasmid, mRNA, or RNP formats	Simple haplotype; *in vitro* proliferation	Limited biological relevance; may revert to diploid in culture

ASC, adult stem cell; ESC, embryonic stem cell; iPSC, induced pluripotent stem cell.

Screening with CRISPR

CRISPR/Cas9 gene editing offers the scalability of RNA interference with reduced off-target effects and higher gene knockdown efficiency. [7] Large-scale single sgRNA libraries, synthesized using array-based technologies, enable genome- and subgenome-scale CRISPR/Cas9 screening applications. [8,9]

Options for Scaling CRISPR

The approach used to study gene-function relationships determines the scale of a CRISPR experiment. Reverse genetics approaches rely on disrupting individual genes to determine the associated phenotypes.

By contrast, forward genetics approaches interrogate multiple genes to identify those that contribute to a particular phenotype. The latter is scalable from a small group of genes to the entire genome.

Single Target Manipulation

CRISPR/Cas9 can edit individual genes in cell lines, [8] primary cells, [10] and whole organisms. [11] The primary goal of these experiments is to understand how perturbation of a single gene relates to a particular phenotype.

Individual-scale CRISPR experiments are cost-efficient and easy to implement compared to multi-gene investigations but are limited to the study of individual genes. [9]

Custom Libraries of Specific Targets

CRISPR screens allow the identification of genes involved in producing a specific phenotype by targeting multiple genes simultaneously. For example, subgenome-scale CRISPR screens can identify kinases that promote tumor growth [12] or membrane proteins that regulate phagocytosis. [13]

Predesigned screens are available that target genes within a given pathway or family, such as kinases. Alternatively, custom libraries allow selection of custom target libraries, based on groups of genes expected to regulate the studied phenotype. Both predesigned screens and custom screens are available in arrayed or pooled format.

Whole-Genome Screens

Genome-scale CRISPR screens enable unbiased and comprehensive detection of all genes that contribute to a particular phenotype. Applications include identifying genes essential for cell survival, [8] drug resistance, [8] and cell-autonomous responses to distinct stimuli. [14]

However, whole-genome screens are also more costly and resource intensive than their subgenome-scale counterparts because of the number of genes investigated. [9] Whole-genome CRISPR screens are available in both pooled and arrayed formats.

Available CRISPR Screening Products

Sanger Whole Genome KO Arrayed CRISPR Libraries contain Over 74,000 individual gRNAs and are available for both mouse and human genomes.

Whole genome knockout pooled libraries provide a more cost-effective way to screen either mouse or human genomes.

For smaller scale screens, 10x genomics compatible custom CRISPR lentiviral pools are available.

Whole Genome CRISPR Activator Libraries utilize Synergistic Activation Mediators (SAM) to allow gain-of-function screening across human and mouse genomes.

Arrayed vs Pooled Screens

Arrayed screens use multi-well microplates with each well receiving sgRNAs targeting an individual gene. This directly links gene-targeting events to the observed phenotype.

In contrast, pooled CRISPR screens involve delivering a single preparation of tens to thousands of sgRNAs to a large population of cells in a concentration that ensures an individual cell receives only one sgRNA. Cell populations with the desired phenotype are selected and next-generation sequencing and data deconvolution is used to determine the correlation between genomically-integrated sgRNA sequences and phenotypes.

The choice between arrayed and pooled formats depends upon assay optimization, library synthesis and delivery, the type of cells used, cost and the endpoint readout(s). Figure 3 summarizes these considerations.

Assay Optimization and Data Analysis

Pooled screens are more cost-effective and simpler to set up than arrayed screens because reagents are not prepared individually. [15] However, pooled screens require complicated data deconvolution, whereas arrayed screens offer more straightforward gene-phenotype correlations. [9]

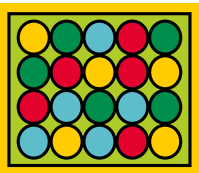
Cell Type

Pooled screens require cells with high proliferative rates to provide enough cells to achieve and maintain full coverage of the library throughout the screening process (>500 targeted cells per targeted gene). [16] The requirement for high proliferation rates means pooled screens are less suitable for primary cell types (e.g., neurons), which are better analyzed using arrayed screens.

However, small-scale pooled screens can be performed in whole organisms, particularly within the context of cancer research. [17]

Figure 3.
Pooled screens vs arrayed screens.

Pooled Screens vs Arrayed Screens

 Arrayed		 Pooled
Complex, labor intensive, costly	Assay Setup	Simple, cost-effective
Simple, direct phenotype-genotype correlation	Data analysis	Complex - requires deconvolution to identify phenotype-genotype correlation
Primary, immortalized	Suitable cell types	Immortalized
Various	Delivery method	Lentiviral only
$$$	Cost	$
Various	Endpoint readout	Limited to proliferation and viablility

These whole organism screens can be achieved through one of two methods:

1. Indirectly by transplanting cultured cells that have been treated to a pooled CRISPR screen and sequencing the resulting tumors.
2. Directly, by transducing the pooled CRISPR screen *in vivo* and capturing the gRNAs from the resulting tumors. [17]

Endpoint Readouts

Pooled screens limit the types of phenotypes that can be measured to simple readouts, including cell proliferation and viability measurements, [9, 15] although fluorescent reporters/markers can be used as well. [14] Arrayed screens can implement virtually any endpoint readout, including morphological, electrophysiological, or biochemical (e.g., luminescence) measurements. [9, 15]

Advanced CRISPR Formats and Their Applications

Modified CRISPR nucleases provide gene-editing capabilities with decreased off-target effects and expand the application of CRISPR to gene activation, repression, and epigenetic modification. Two of the most versatile alternative formats of CRISPR are nickases and catalytically inactive nucleases.

CRISPR Nickases

Wild type CRISPR nucleases contain two active nuclease domains, which work together to cleave both strands of DNA to create a double-strand break. Inactivation of one domain, such as the RuvC nuclease domain through the introduction of a D10A point mutation, results in a Cas9 nuclease that can cleave only the target strand. [18] A double-strand break can be created by pairing together two nickases with different, but adjacent, gRNAs (Figure 4).

The use of paired nickases removes the possibility of off-target effects since both Cas9 nickases must nick their targets to generate a DSB. Wild type Cas9 nuclease creates a blunt-end double-strand break. In contrast, the two separate gRNAs employed by paired nickases result in a DSB with cohesive ends, providing greater control over gene insertion and integration. The increased precision and reduced off-target effects of CRISPR nickases make them ideal for performing gene editing in therapeutic applications, such as genome editing in primary T cells and excision of viral DNA (e.g., hepatitis B virus, HBV) in humans. [19]

Cas9-D10A Nickase Products

Cas9-D10A nickases are available in mRNA, plasmid and protein formats.

Cas9-D10A nickase plasmids expressing either GFP or RFP fluorophores are available to enable visualization and selection of transfected cells.

CRISPR Nickase EMX1 Positive Control allows validation of the nickase in your system.

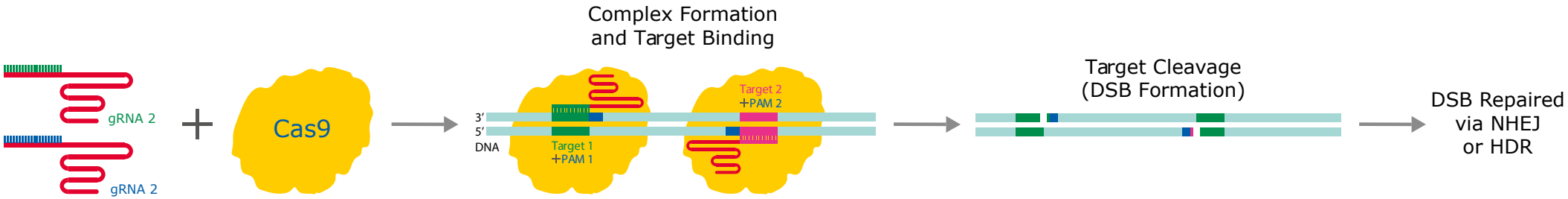


Figure 4.
Schematic depicting how paired nickases create a DSB.

Nuclease-Deficient Cas9

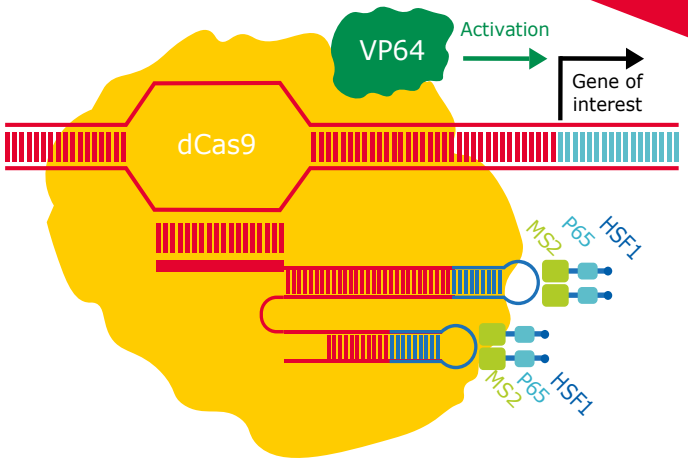
The nuclease-deficient version of Cas9 (dCas9) has point mutations in both nuclease domains (D10A and H840A), making it catalytically inactive. While dCas9 cannot cleave DNA, it maintains the ability to precisely target DNA, allowing it to serve as a cargo-delivery system to specific sequences in the genome. dCas9, coupled to transcriptional activators or repressors, enables modulation of gene expression without creating a DSB, which can induce cytotoxicity and genomic instability, particularly in cancer cell lines. [20, 21]

Nuclease-deficient Cas9 mimics endogenous gene repression and activation processes, making it more biologically relevant than traditional CRISPR knock-ins or knockouts.

CRISPR Activation

CRISPR activation (CRISPRa) uses a modified CRISPR-dCas9 system, with transcriptional activators attached to either the dCas9 or the gRNA to increase expression of a target gene. The most common CRISPRa system uses dCas9 fused to the polymeric viral transactivation domain, VP64. This synthetic transcription factor is then targeted to a specific promoter sequence by the guide RNA, resulting in gene activation (Figure 5).

The CRISPR Synergistic Activation Mediator (SAM) complex is a highly effective CRISPRa system, which recruits three unique transcriptional activator domains to



CRISPRa

Figure 5.
The CRISPR Synergistic Activation Mediator (SAM) complex.

the targeted gene promoter. [22,23] This system utilizes the VP64 transcriptional activator (a multimeric form of VP16) fused directly to dCas9, and two other transcriptional activator domains (p65 and HSF1) recruited via protein-binding RNA aptamers engineered into the stem-loop regions of the sgRNA. [23] Coupling the RNP with these additional co-activators further increases the potency of the CRISPRa system and can increase the expression of challenging targets like OCT4 hundreds of folds. [24]

CRISPRa can also serve as a complementary approach to loss-of-function (LOF) genetic screens by enriching for a different set of genes responsible for a phenotype. [19] CRISPRa screens show little-to-no off-target activity and can be used to study genes that are difficult to characterize using only LOF approaches because of their high copy number.

Available CRISPRa Products

SAM CRISPRa is available as lentiviral transduction particle kits or DNA plasmid kits.

Whole Genome SAM CRISPRa Pooled Lentiviral Libraries provide genome wide gain-of-function screening and are available for both human and mouse genomes.

CRISPR Interference

CRISPR interference (CRISPRi) provides a highly efficient method for silencing genes without altering the underlying DNA sequence. CRISPRi uses dCas9 fused to a transcriptional repressor, such as the well-characterized Krüppel associated box (KRAB) domain. KRAB recruits proteins that result in epigenetic repression surrounding the target DNA (e.g. DNA methylation). When targeted to promoters of genes, this activity prevents the recruitment of the transcriptional machinery, effectively silencing gene function.

CRISPRi using dCas9-KRAB recapitulates two mechanisms of endogenous transcriptional repression: the recruitment of repressor complexes and sterically hindering RNA polymerase binding (Figure 6). [25, 26]

The dCas9-KRAB CRISPRi system can successfully silence genetic components in non-dividing neurons, which are notoriously difficult to edit at the genomic level. [27] CRISPRi exhibits better targeting specificity compared to other conventionally used gene silencing approaches such as RNA interference (RNAi). [28]

Epigenetic Modification Using dCas9

CRISPR-dCas9 can be used to alter epigenetic modifications surrounding the targeted DNA, allowing the direct study of the epigenome. Epigenetic effectors, such as histone acetyltransferases, can be targeted to genes of interest in order to study the epigenetic control of these loci as well as induce or repress expression in a more native manner.

The dCas9-p300 CRISPR gene activator system uses the catalytic histone acetyltransferase domain of the human p300 protein to unwind DNA, allowing for endogenous activation of gene expression. In practice, OCT4 transcript levels have been increased at least two-fold when activated using p300-dCas9. [29]

CRISPR-based epigenome modification can also be utilized in high-throughput screens to identify and study functional regulatory elements for specific loci. [30]

CRISPR Epigenetic Modification Products

CRISPR dCas9-p300 Activator is available in plasmid format with co-expression of either GFP or RFP allowing for easy visualization of successful transfection.

Recombinant dCas9-3X FLAG-Biotin protein is available seperately.

A validated positive control against OCT4 provides confirmation of your CRISPR dCas9-p300 system.

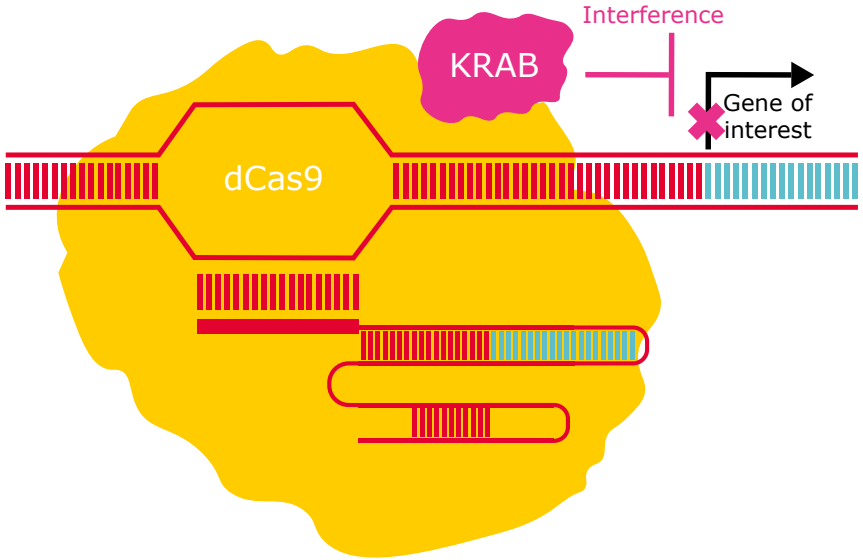


Figure 6.
CRISPRi using the KRAB domain.

Controls for CRISPR Gene Editing

Choosing the right controls for your gene-editing experiment is essential in determining the validity of your results and facilitates troubleshooting. Many different types of controls are available for CRISPR. As a minimum, every CRISPR experiment should have both a positive and negative control.

Positive Controls

A positive control demonstrates that the CRISPR reagents and delivery method used are working as expected in your experimental setup. In cases where the experiment produces a negative result (or no observed change), it allows confirmation that this is not due to experimental design. A typical positive control consists of all the necessary CRISPR reagents required to perform gene editing along with a gRNA that can successfully target another gene in the cell.

A range of validated controls in a variety of delivery formats are available from the Sigma-Aldrich® portfolio that can verify your CRISPR system.

Negative Controls

A negative control demonstrates that the observed change is a direct result of the introduced mutation and not due to other non-specific effects. A typical negative control consists of all the necessary CRISPR reagents to perform gene editing but uses a gRNA that does not recognize any sequence in your experimental system. Off the shelf negative controls that have been designed not to target any region in the human, rat, or mouse genome are available from the Sigma-Aldrich® portfolio.

Non-targeting gRNAs, or gRNAs that do not target any gene in the genome of the cell, can be used to confirm that your LOF phenotype is not a technical artifact.

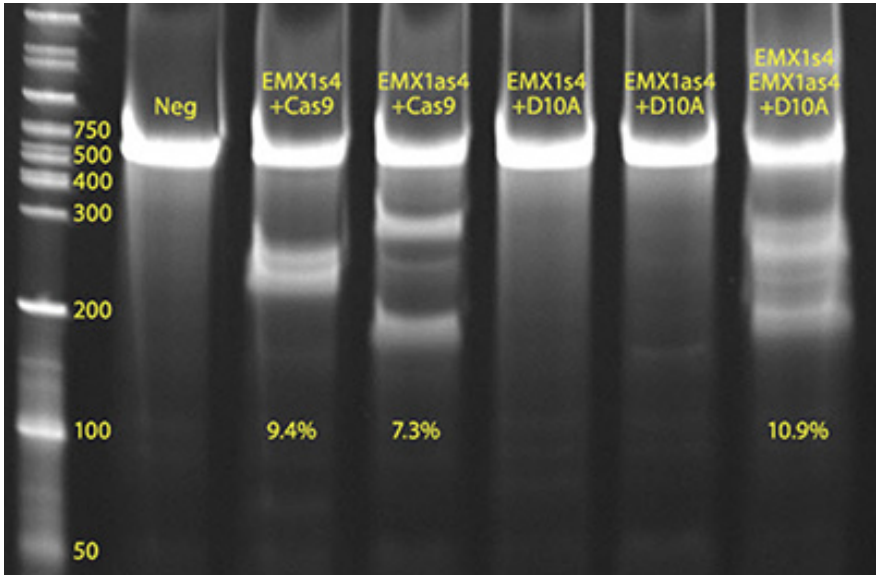


Figure 7.
CEL I image of CRISPR01 positive control (lane labeled EMX1s4+Cas9) and CRISPR02 positive control (lane labeled EMX1s4, EMX1as4+D10A). Credit: SigmaAldrich.com/CRISPR

Determination of Successful Gene Editing

There are many options to measure the success of a gene-editing experiment, including Sanger sequencing, mismatch detection assays, next-generation sequencing (NGS), phenotypic assessment, and measuring mRNA and protein levels for the targeted gene. These methods differ in their sensitivity, scalability, resolution, and cost. [31]

For example, although NGS offers exceptionally high sensitivity and resolution, it is costly and requires significant technical expertise to carry out. Mismatch detection, on the other hand, is easy to perform but lacks the sensitivity of Sanger sequencing and NGS.

Determining the presence of an indel is often considered best practice. However, simply measuring the changes made to the genome is not sufficient to determine if they have disrupted gene function and created a gene knockout that causes a phenotypic response. It is also important to measure protein levels, ideally using a well-validated antibody.

Validating Successful Genetic Targeting

Confirmation of successful gene targeting requires the detection of the insertions or deletions (indels) introduced by the CRISPR experiment. [5]

Sanger DNA Sequencing

Sanger DNA sequencing is a reliable and highly sensitive method that precisely identifies introduced mutations. This method is considered the gold-standard for gene-editing validation, despite the time-consuming and labor-intensive nature of the technique, which requires multiple preparation steps to establish a clonal cell population before sequencing.

Next-Generation Sequencing

Next-generation sequencing (NGS) does not require the establishment of a clonal cell population harboring the mutation, and identifies rare mutations in a cell sub-population. A high error rate and cost per run are the current main barriers to this validation technique. The continued advancement of sequencing technology and falling costs means that NGS will soon become the dominant validation option for CRISPR editing.

Surveyor Nuclease Assay

The Surveyor™ nuclease assay is an easy method based on the principle of mismatch detection. [32] In this method, a pool of PCR products from the edited alleles is denatured and rehybridized, after which the Surveyor nuclease selectively detects and cuts those pairs that are mismatched due to the presence of an indel on one of the strands introduced by the cellular NHEJ repair process. Cleaved DNA amplicons that incorporate a mismatch are detected on an agarose gel. This assay has limited resolution, as it can miss small indels and is unable to distinguish CRISPR-derived indels from naturally occurring polymorphisms. [31]

TIDE Assay

Tracking of Indels by Decomposition (TIDE) provides a more sensitive indel detection method. This method involves:

1. Amplification of genomic region targeted by the nuclease from DNA isolated from transfected cells.
2. Sanger sequencing of PCR products.
3. Analysis of sequencing to determine the presence of indels.

TIDE reduces the total cost of validation, as it allows Sanger sequencing to be performed on mixed cell populations. However, this pooled approach means that it cannot distinguish two alleles of the same length, and it struggles with rare alleles. Moreover, the reliability of TIDE is dependent on the quality of the PCR products and Sanger sequences.

The Importance of Orthogonal Validation

While technical validation that the gene-editing experiment was successful is essential, it does not identify false positives and false negatives, whether the editing has led to the desired phenotype or guarantee the identified genes are useful therapeutic targets.

They are also relatively simple assays that do not convey the complicated nature of mutations and transcriptional dysregulation observed in many diseases, nor whether small-molecule or protein-mediated inhibition will deliver the same response *in vivo.*

Therefore, implementing CRISPR knockout experiments should be complemented by more nuanced gain-of-function assays, including CRISPR knock-ins and CRISPR activation.

Table 3.
Validation for CRISPR applications.

Experimental Application	Main Approach	Complimentary Techniques	Related Products
Generation of Cellular and Animal Models	gRNA gene KO	KI experiments with donor DNA template Rescue/overexpression experiments using CRISPRa Rescue/overexpression experiments using ORF	Guaranteed predesigned gRNAs Sanger QuickPick KO gRNAs CRISPRa Screening ORF clones
Functional Genomic Screens	Pooled or Arrayed whole genome CRISPR KO libraries	Gene modulation libraries: CRISPRa and CRISPRi	Sanger Arrayed KO library GeCKO KO pooled libraries Sigma KO pooled libraries SAM CRISPR activation library Weissman CRISPR inhibition library (coming soon)
Transcriptional Modulation	RNAi	CRISPR activation and CRISPR inhibition	CRISPRa Single cell compatible CRISPR pools RNAi

References

1. Chatterjee, P, *et al*. Minimal PAM specificity of a highly similar SpCas9 ortholog. Sci Adv. (2018) 4: eaau0766. doi: 10.1126/sciadv.aau0766

2. Liu, X, *et al*. Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. Sci Rep. (2016) 6: 19675. doi: 10.1038/srep19675

3. Doench, JG, *et al*. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol. (2014) 32: 1262–7. doi: 10.1038/nbt.3026

4. Wu, X, *et al*. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol. (2014) 32: 670–6. doi: 10.1038/nbt.2889

5. Kosicki, M, *et al*. Dynamics of indel profiles induced by various CRISPR/Cas9 delivery methods. Prog Mol Biol Transl Sci. (2017) 152: 49-67. doi: 10.1016/bs.pmbts.2017.09.003

6. Lino, CA, *et al*. Delivering CRISPR: a review of the challenges and approaches. Drug Deliv. (2018) 25: 1234-57. doi: 10.1080/10717544.2018.1474964.

7. Evers B, *et al*. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. Nat Biotechnol. (2016) 34: 631-3. doi: 10.1038/nbt.3536.

8. Shalem O, *et al*. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. (2014) 343: 84-7. doi: 10.1126/science.1247005

9. Costa JR, *et al*. Genome Editing Using Engineered Nucleases and Their Use in Genomic Screening. In: Sittampalam GS, Grossman A, Brimacombe K, Arkin M, Auld D, Austin CP, et al, editors. Assay Guidance Manual. Bethesda (MD) 2017.

10. Schumann K, *et al*. Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. Proc Natl Acad Sci U S A. (2015) 112: 10437-42. doi:10.1073/pnas.1512503112

11. Wang H, *et al*. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. (2013) 153: 910-8. doi: 10.1016/j.cell.2013.04.025

12. Sredni ST, *et al*. A functional screening of the kinome identifies the Polo-like kinase 4 as a potential therapeutic target for malignant rhabdoid tumors, and possibly, other embryonal tumors of the brain. Pediatr Blood Cancer. (2017) 64: e26551. doi: 10.1002/pbc.26551

13. Pluvinage JV, *et al*. CD22 blockade restores homeostatic microglial phagocytosis in ageing brains. Nature. (2019) 568: 187-92. doi: 10.1038/s41586-019-1088-4

14. Parnas O, *et al*. A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. Cell. (2015) 162: 675-86. doi: 10.1016/j.cell.2015.06.059

15. Shalem O, *et al*. High-throughput functional genomics using CRISPR-Cas9. Nat Rev Genet. (2015) 16: 299-311. doi: 10.1038/nrg3899

16. Joung, J, *et al*. Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. Nat Protoc. (2017) 12: 828–63. doi: 10.1038/nprot.2017.016

17. Chow RD, Chen S. Cancer CRISPR Screens In Vivo. Trends Cancer. (2018) 4: 349-58. doi: 10.1016/j.trecan.2018.03.002

18. Lamberth, J, *et al*. Highly efficient and specific genome editing in human cells with paired CRISPR-Cas9 nickase ribonucleoproteins. BioRxiv. (2019) 766493. doi: 10.1101/766493

19. Karimova, M, *et al*. CRISPR/Cas9 nickase-mediated disruption of hepatitis B virus open reading frame S and X. Sci Rep. (2015) 5: 13734. doi: 10.1038/srep13734

20. Rayner E, *et al*. CRISPR-Cas9 Causes Chromosomal Instability and Rearrangements in Cancer Cell Lines, Detectable by Cytogenetic Methods. CRISPR J. (2019) 2: 406-16. doi:10.1089/crispr.2019.0006

21. Schwertman P, *et al*. Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers. Nat Rev Mol Cell Biol. (2016) 17: 379-94. doi: 10.1038/nrm.2016.58

22. Yang J, *et al*. Genome-Scale CRISPRa Screen Identifies Novel Factors for Cellular Reprogramming. Stem Cell Reports. (2019) 12: 757-71. doi: 10.1016/j.stemcr.2019.02.010

23. Konermann S, *et al*. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature. (2015) 517: 583-8. doi: 10.1038/nature14136

References

24. Chavez, A, *et al*. Comparison of Cas9 Activators in Multiple Species. Nat Methods. (2016) 13: 563-567. doi: 10.1038/nmeth.3871

25. Gilbert LA, *et al*. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell. (2014) 159: 647-61. doi: 10.1016/j.cell.2014.09.029

26. Palmer AC, *et al*. Transcriptional interference by RNA polymerase pausing and dislodgement of transcription factors. Transcription. (2011) 2: 9-14. doi: 10.4161/trns.2.1.13511

27. Zheng, Y, *et al*. CRISPR interference-based specific and efficient gene inactivation in the brain. Nat Neurosci. (2018) 21: 447-54. doi: 10.1038/s41593-018-0077-5

28. Kampmann, M. CRISPRi and CRISPRa screens in mammalian cells for precision biology and medicine. ACS Chem Biol. (2018) 13: 406-16. doi: 10.1021/acschembio.7b00657

29. Ji, Q, *et al*. Engineered zinc-finger transcription factors activate OCT4 (POU5F1), SOX2, KLF4, c-MYC (MYC) and miR302/367. Nucleic Acids Res. (2014) 42: 6158-67. doi: 10.1093/nar/gku243

30. Klann, T, *et al*. CRISPR–Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. Nat Biotechnol. (2017) 35: 561–8. doi: 10.1038/nbt.3853

31. Germini, D, *et al*. A comparison of techniques to evaluate the effectiveness of genome editing. Trends Biotech. (2018) 36: 147-59. doi: 10.1016/j.tibtech.2017.10.008

32. Qiu, P, *et al*. Mutation detection using Surveyor™ nuclease. Biotechniques. (2004) 36: 702-7. doi: 10.2144/04364PF01