Genome editing *in vivo*
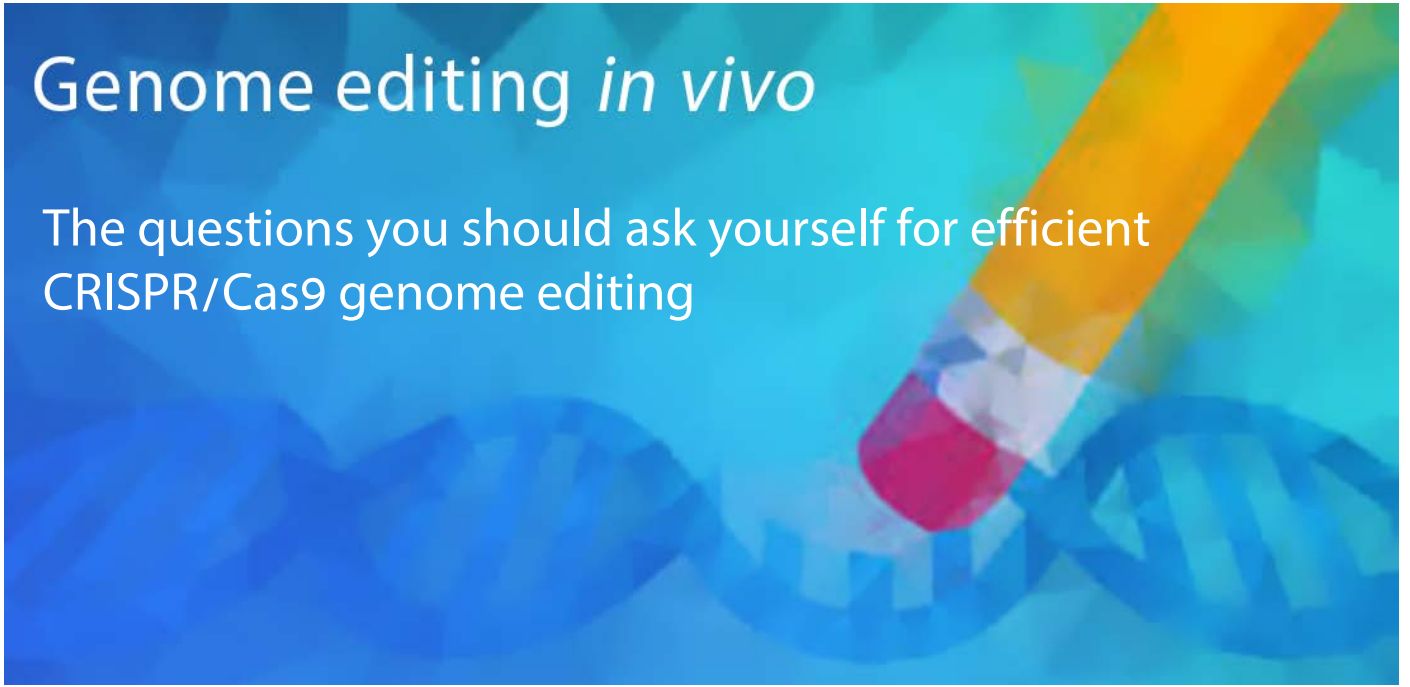
The questions you should ask yourself for efficient CRISPR/Cas9 genome editing

CRISPR/Cas9 FAQs

Since Cong et al. Science publication in February 2013 that first demonstrated the use of type II CRISPR/Cas9 system to engineer eukaryote genome, Cas9 nucleases directed by short RNAs have emerged as a revolution for genome editing, allowing precise gene modification in virtually any species. Enthusiasm from the international scientific community for this technology is illustrated by the huge and growing number of publications including the acronym «CRISPR». In 2014 and 2015, more than 1700 papers can be retrieved, which is incredible for such a recent discovery. During this very short 2 years period, improvements followed one another, including new Cas9 proteins or use of Non Homologous End Joining inhibitors. However, most of these developments were validated with very low number of cases and thus need confirmations. Moreover, all PROs and CONs, possibilities, bias like mosaicism or off target events are far from being completely understood. Correctly designing a CRISPR/Cas9 gene editing experiment can still be challenging, especially for complex modifications.

CELPHEDIA genetic engineering workgroup gathers many experts using CRISPR/Cas9 genome editing in various animal models such as mouse, rat, zebrafish and drosophila. By sharing knowledge, experience and an extensive bibliographic review, the group has written together recommendations for CRISPR/Cas9 system. It covers current advances on several technical aspects, from the RNA guide optimized design to the genotyping analysis and the validation of the newly generated models.

Authors

Guillaume Pavlovic[1] , Pascale Mercier[3] , Valérie Risson[4] , Philippe Schmitt[5] , Marie-Christine Birling[1] , Francina Langa[6] , Suzy Markossian[4] , Lydia Teboul[7] , Erwana Harscoet[8], Karelia Lipson Ruffert[9] , Sandra Offner[10], Fabien Angelis[11] , Isabelle Barde[10] , Matthieu Bringart[12] , Jean-Paul Concordet[13] , Frédéric Fiore[11] , Térésa Jagla[14] , Laurent Tesson[15], Jean-Stéphane Joly[16] and Yann Hérault[1,2,5].

[1]PHENOMIN-ICS, CNRS UMR7104, INSERM U964, UdS, Illkirch; [2]IGBMC, CNRS UMR7104, INSERM U964, UdS, Illkirch; [3]IPBS CNRS UMR 5089, Toulouse; [4]IGFL, SFR Biosciences CNRS/INRA (UMS 3444/US8), Université de Lyon, ENS, Lyon; [5]CELPHEDIA, CNRS, Illkirch; [6]Centre d'Ingénierie Génétique Murine, Institut Pasteur, Paris; [7]The Mary Lyon Centre, Medical Research Council, Harwell, Oxfordshire OX11 0RD, UK; [8]INRA UMR1198, Biologie du Développement et Reproduction, Jouy-en-Josas; [9]SEAT-TAAM PHENOMIN-TAAM-SEAT CNRS UPS44,Villejuif; [10]School of Life Sciences and Frontiers in Genetics Program, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; [11]PHENOMIN-CIPHE, INSERM US012, CNRS UMS3367, UM2 Aix-Marseille Université, Marseille; [12]INSERM-US006 CREFRE (Centre Régional d'Exploration fonctionnelle et Ressources Expérimentales), Toulouse; [13]TACGene, Muséum National d'Histoire Naturelle, CNRS UMR 7196, INSERM U1154, Paris; [14]Genetics, Reproduction and Development, INSERM U931, CNRS UMR6247, Clermont University, Clermont-Ferrand; [15]INSERM UMR 1064-ITUN, Centre Hospitalier Universitaire de Nantes, Faculté de Médecine, Nantes; [16]TEFOR, CNRS, Gif-sur-Yvette
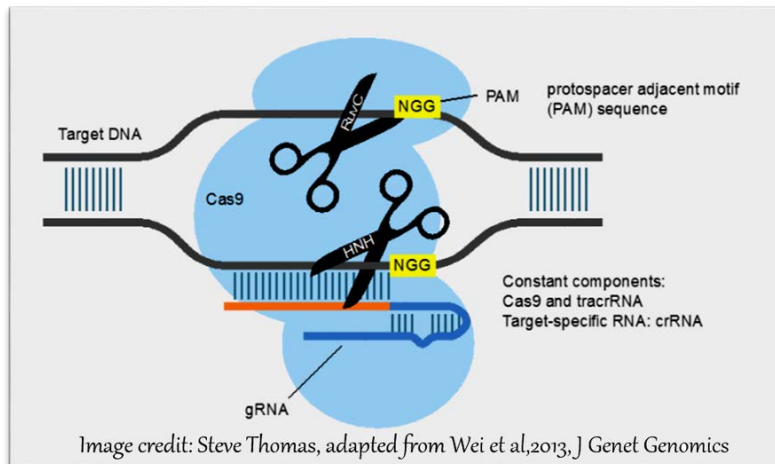
Table of contents

What is the CRISPR/Cas9 system?

The CRISPR/Cas system is a natural adaptive immune system used by various bacteria and archaea to mediate defense against viruses and other foreign nucleic acids. Prokaryotes can specifically incorporate short sequences from invading genetic elements (virus or plasmid) into a region of its genome that is distinguished by Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs). When these sequences are transcribed and precisely processed into small RNAs, they guide a multifunctional protein complex (Cas proteins) to recognize and cleave incoming foreign genetic material (for review see (Doudna & Charpentier 2014)).

Since discovery, these prokaryotic CRISPR/Cas systems have been modified and adapted to allow genetic engineering in animals, plants and other eukaryotes.

The sgRNA (fusion of tracr and crRNAs) directs Cas9 to the target sequence where it unwinds and cleaves the double stranded DNA.

Six types of CRISPR systems were described to date (Shmakov et al. 2015; Makarova et al. 2015). The Type II CRISPR system is the most widely adopted because of its simplicity, efficiency and multiplexing possibilities. In this system, a combination of a single "guide RNA" (sgRNA), complementary to the DNA target site, and a Cas9 protein, which acts as a site specific nuclease, is used to generate a double-stranded DNA break at the selected targeted sequence (Jinek et al. 2012). Genomic target sequence binding by Cas9 and its subsequent cutting requires a Protospacer Adjacent Motif (PAM) (often -NGG) flanking the 3′ site of the chosen 20 bp target.

For basic information on mechanisms and applications of CRISPR, please see AddGene CRISPR Practical Guide

CRISPR a tool for any species?

The CRISPR/Cas9 system has been used to achieve efficient genome editing in a variety of species and cell types, including human cell lines, bacteria, zebrafish, yeast, mouse, fruit fly, roundworm, rat, rabbit, common crops, pig, and monkey.

Which nuclease to use: CRISPR/Cas9 vs. ZFNs and TALEs

Zinc Finger Nucleases (ZFNs) and Transcription Activator–Like Effector (TALEs) proteins are -like CRISPR/Cas-nucleases that cleave specific genomic sequences. Both ZFNs and TALEs were successfully used in numerous species to obtain targeted genetic alterations. CRISPR-associated (Cas) systems have emerged in 2013 as potentially convenient and efficient alternatives. Indeed, unlike ZFNs and TALEs, design of CRISPR/Cas9 is very easy and only requires synthesis of a short RNA sequence corresponding to sgRNA. The Cas9 nuclease can be delivered as DNA, RNA or protein (see Cas9 and guide RNA: What to microinject?).
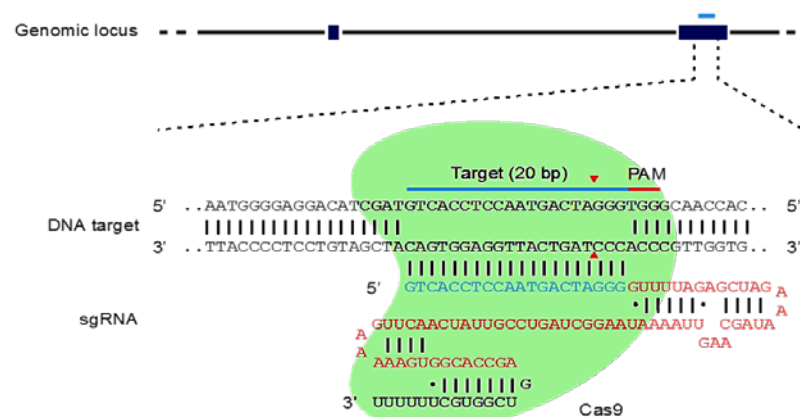
Design of CRISPR/Cas9 genome editing experiment

Genomic DNA target sequence

- The target sequence must be complementary to the sgRNA sequence, 20 bp long and directly flanked on the 3′end by the appropriate Protospacer Adjacent Motif (PAM) sequence. The PAM sequence is present in the DNA target sequence but NOT in the guide RNA sequence. When the sgRNA recognises a DNA target sequence flanked by the PAM sequence, Cas9 will cleave 3-4 nucleotides upstream from the PAM sequence. Efficiency and specificity depend on the choice of an appropriate target sequence in the genomic DNA.
- The PAM sequence at the immediate 3′ end of the DNA target sequence is a short nucleotide sequence specific to each Cas9 complex and varies according the bacterial origin. The most widely used system comes from Streptococcus pyogenes and the PAM sequence is a 5′-NGG (Jinek et al. 2012). Sites flanked by 5'-NAG PAMs are targeted but with a lower efficiency.
- Cas9 can be derived from other bacteria, such as Neisseria meningitidis where PAM sequence is -NNNNGATT (Zhang et al. 2013), Campylobacter jejuni where PAM sequence is -NNNNACA or Streptococcus thermophilus where PAM sequence is - NNAGAAW (Cong et al. 2013; Garneau et al. 2010).
- Components (sgRNA, Cas9) derived from different bacteria will not function together.

Design of guide RNAs (sgRNAs)

The main features of single guide RNAs in genome editing experiments are: efficiency, specificity and target cleavage position.

Extracted from Ran et al. 2013, Nat Protoc.
Schematic of the RNA-guided Cas9 nuclease (Ran, Hsu, Wright, et al. 2013). The Cas9 nuclease from Streptococcus pyogenes (in green) is targeted to genomic DNA (shown for example is the human EMX1 locus) by an sgRNA consisting of a 20-nt guide sequence (blue) and a scaffold (red). The guide sequence pairs with the DNA target (blue bar on top strand), directly upstream of a requisite 5′-NGG adjacent motif (PAM). Cas9 mediates a double strand breaks ~3 bp upstream of the PAM (red triangle).

- For efficiency, no strict recommendations are yet defined. Some studies have pointed out sequence features that may be critical such as high G/C content proximal to the PAM (Ren et al. 2014), GG proximal to the PAM (Farboud & Meyer 2015). Some even have proposed efficiency prediction scores (Doench et al. 2014).
- For knock-out experiments, cuts occur and are repaired, generally introducing an indel in sequences. The presence of short sequence homologies on both sides of the cleavage site may promote frequent recurrence of identical deletions that will or will not be in-frame. Bae et al. 2014 have proposed an "out-of-frame" score to evaluate designs. Generally, it is important to plan the experiment so genotyping will be facilitated. One very convenient possibility for KO experiments is to use 2 sgRNAs and select mutants carrying the expected deletion of defined sequence.

- In knock-in experiments, especially when point mutations are desired, the proximity of the cleavage site to the intended mutation impacts greatly on efficiency. We recommend introducing additional base changes in the donor oligonucleotide that will prevent further processing of the repaired allele.
- As all the sgRNA do not show the same efficiency for generation of double strand breaks, we advise to test multiple sgRNAs (see validation of guide RNA) and to inject two sgRNAs per target site.
- For specificity, recent experimental studies have shown that despite early concerns about the CRISPR-Cas9 system, cleavage directed by sgRNAs is generally highly specific to the target sequence (Cho et al. 2014). Off-target cleavage has been detected at sequences related to the target that can be identified and scored by bioinformatic analysis (Hsu et al. 2013; Tsai et al. 2015). Practical solutions for improving specificity potentially include using shorter guide RNAs (Fu et al. 2014) and appending GG at the 5'end (Ran, Hsu, Wright, et al. 2013).

For selection of sgRNAs, several web sites are available with specific features of interest (MacPherson & Scherf 2015; Ishida et al. 2015) such as

- Use of experimentally validated off-target score (http://crispr.mit.edu/),
- Efficiency score (http://crispr.mit.edu/, http://www.flyrnai.org/crispr2/),
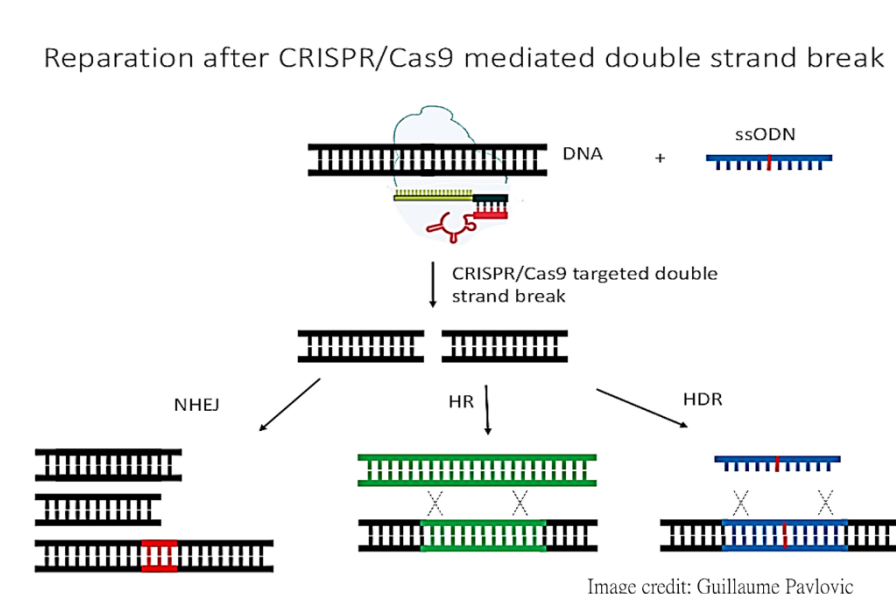- Out-of-frame score (http://www.rgenome.net/).

We recommend trying the **CRISPOR bioinformatic program** on the TEFOR web site for the design of sgRNAs and the evaluation of off-target sites when using the CRISPR/Cas9 system. The CRISPOR tool covers a large number of genomes and Streptococcus pyogenes Cas9 as well as the selection of guide RNAs for Cas9 with different PAMs. In addition, primer design for PCR of the targeted region and primer sequences for sgRNA expression are proposed. We recommend using the specificity score rather than the efficiency score for selection of the sgRNA. Indeed in the mouse, sgRNA with efficiency score as low as 1 have worked really well (PHENOMIN-ICS unpublished data).

Design of donor DNA

For knock-in experiments, a set of definitive guidelines has yet to be defined. Design and efficiency will depend on the exact intended sequence modification and the biological system.

Use of ssODNs donors

The use of single-stranded DNA oligonucleotides (ssODNs) is an effective and simple method for making small edits in the genome, such as the introduction of point mutations in the mouse and rat genomes (Chen et al. 2011).

Reparation after CRISPR/Cas9 mediated double strand break

Image credit: Guillaume Pavlovic

- ssODN can be used to carry point mutations or short sequences such as Tag sequences, LoxP sites and/or silent mutations that introduce a restriction site to facilitate genotyping of mutant animals.
- Homology arms flanking the mutation should be between 30 bases to 100 bases of homology either side, 40-60 bases on each side appearing to be optimal ([Yang et al. 2013](#); [Yang et al. 2014](#)). The point of cleavage by Cas9 should be as close as possible to the intended mutation (up to 40 bp Cas9 cut side and target site has been shown to work efficiently).
- To prevent further sequence cleavage and modification, the PAM sequence in the ssODN donor should be mutated (but avoid changing to NAG or NGA that can also act as a PAM, although less efficiently).
- A silent mutation can also be introduced to create a restriction site to facilitate subsequent genotyping.
- Using the sequence strand identical to the guide RNA target has been recommended to prevent potential DNA-RNA complex formation.
- When an ssODN is used to introduce modification, the mechanism is called Homology Directed Repair (HDR) in contrast to homologous recombination (HR) that occurs when longer homology arms are used (double stranded plasmid). Currently, the mechanisms involved in HDR are not well understood.

Use of plasmid donor DNA

To achieve the insertion of a larger sequence, such as fluorescent protein coding sequences or other cDNA, or precise deletions. A plasmid donor DNA can be used. Homology arms flanking the cleavage site are generally 500-1 000 bp long, but the optimal length of these homology arms is still under discussion. If possible, it is preferable to use isogenic DNA as the frequency of homologous recombination can be affected by the presence of vector-target mismatches. Use of circular vs linear plasmid can also influence efficiency, as well as the occurrence of multicopy integration. The latter might be better verified by Southern blotting or copy counting assays (quantitative PCR (qPCR) or Droplet Digital™ PCR (ddPCR™)).

Conditional alleles

The generation of floxed alleles is an important goal. However the best strategy to achieve this is still under debate. Some propose that the best strategy, at least in mammals, is the use of HR plasmid, as the insertion will ensure that both loxP are targeted to the same allele (ISTT source). Others suggest that even if insertion of 2 ssODNs can occur on different alleles, the overall efficiency of using oligonucleotides as templates is such that this strategy remains more efficient.

The various Cas9

CRISPR/Cas9 is a very large family of RNA/nuclease complex ([Makarova et al. 2015](#)). Moreover, many new versions of CRISPR molecules have been engineered since 2013 and more are continuously developed. They can be grouped in 4 categories.

The Streptococcus pyogenes wild type Cas9

It is the most frequently used nuclease. For most of the applications, this Cas9 is preferably used by scientists and supports a very good NHEJ efficiency.

Cas9 nickases

Cas9 nickase proteins are modified version of the *Streptococcus pyogenes* wild type Cas9 obtained either by inactivation of either the RuvC domain or the HNH domain. As the nuclease activity occurs via these 2 separate RuvC and HNH

nuclease domains, each of them cutting independently one single strand generating a single-stranded DNA lesion, called "nick", on opposite strands (Jinek et al. 2012). Inactivation of one of the domain will lead in a single nick instead of a double strand break (Ran, Hsu, Lin, et al. 2013).

Two different versions of Cas9 nickase are commonly used (Ran, Hsu, Lin, et al. 2013)

- Cas9D10A: aspartate to alanine substitution leads to inactivation in the RuvC domain.
- Cas9H840A: histidine to alanine substitution leads to inactivation in the HNH domain.

When using a nickase, two guide RNAs are designed so they binds opposite strands of the target and the result of nickase activity is a staggered-cut double strand break. The literature suggests that these overhanging breaks are predominantly repaired by high fidelity base excision repair pathway or HDR events, but rarely by NHEJ. This thought to lead to a decrease of the off target effects (Ran, Hsu, Lin, et al. 2013; Mali et al. 2013), although it is the experience of some members of this group that indels can occur with high frequency.

The Cas9 nickase approach shows two major limitations:

- 2 efficient sgRNAs are necessary to get a staggered-end double strand break.
- The requirement for 2 closely spaced sgRNA/PAM sequences in the same target region of the genome narrows the choice of possible sgRNA design.

RNA-guided endonucleases from other bacterial strains

New Cas9 nucleases from other bacterial strains are continuously being described. They may have a potential interest due to their different PAM specificity or smaller size for integration in Adeno-Associated Viruses (Hou et al. 2013; Ran et al. 2015; Makarova et al. 2015).

Very recently, a new class 2 CRISPR system called CRISPR-Cpf1 has been described (Zetsche et al. 2015). The practical value of this new tool needs further evaluation.

Other engineered Cas9

A Split Cas9 was engineered for inducible genome editing and transcription modulation (Zetsche et al. 2015). The fusion of catalytically inactive Cas9 to FokI nuclease was generated to reduce off targets (Guilinger et al. 2014; Tsai et al. 2014).

Which Cas9 should be used?

For most applications, the wild type Cas9 is sufficient and is preferred when efficiency is the main objective. Whereas the nickase Cas9 should be preferred when the specificity (absence of high risk potential off-target site) is mandatory.

Cas9 and guide RNA: What to microinject?

The Cas9 nuclease can be delivered as DNA, mRNA or protein. The single guides can be delivered as DNA (included in a plasmid), or RNA. Each of these forms were successfully used by scientists. The delivery method should be adapted to your objective.

- Expression plasmids are easier to prepare and to handle. However, the risk of random integration of DNA in the genome cannot be ignored. If DNA integrants are selected, depending on the promoter used, expression of the CRISPR/Cas9 can be constitutively maintained.

- To induce double strand breaks within the targeted site, mRNA seems more efficient than DNA, probably because higher concentrations can be injected in embryos without toxic effects (Horii et al. 2014). The resulting expression and activity will be transient (but not limited to the 1 cell-stage).
- Cas9 protein shows efficiency similar to Cas9 mRNA, which could be of particular interest for fast-developing species like Zebrafish or Xenopus because of their rapid pace of development. In mouse embryos, microinjection of Cas9 protein instead of mRNA could also diminish mosaicism (Singh et al. 2015). Moreover, direct microinjection of recombinant Cas9 protein can also reduce the potential off-target effects by regulating the amount and the activity of Cas9. Additional experiments are yet to be done to confirm these two last points.

Concentration of material to microinject

It is assumed that increasing Cas9 concentration leads to higher off-target activity (Hsu et al. 2013; Pattanayak et al. 2013), whereas decreasing Cas9 concentration significantly improves the on- to off-target ratio, at the expense of on-target cleavage efficiency.

A wide range of concentrations have been used efficiently, especially for RNA molecules (less toxicity in rodent embryos than DNA):

- DNA plasmid- generally encoding both Cas9 and sgRNAs (2-10 ng per µl)
- Cas9 mRNA (5-200 ng per µl).
- sgRNA (2.5 to 50 ng per µl)
- protein (0.3-6 µM)
- For HDR mediated gene correction or insertion: donor ssODN (1-100 ng per µl)

RNAs, proteins and oligonucleotides can be injected into the cytoplasm or the nucleus and lead to successful mutagenesis. Efficiency in either technique seems to be related to specific expertise in microinjection facilities. The use of CRISPR/Cas9 system seems to work with all genetic backgrounds in mice and rats. As in classical transgenesis experiments, the critical point is the availability, robustness and development of embryos of pure inbred strains, especially C57BL/6s, as well as microinjectionist proficiency.
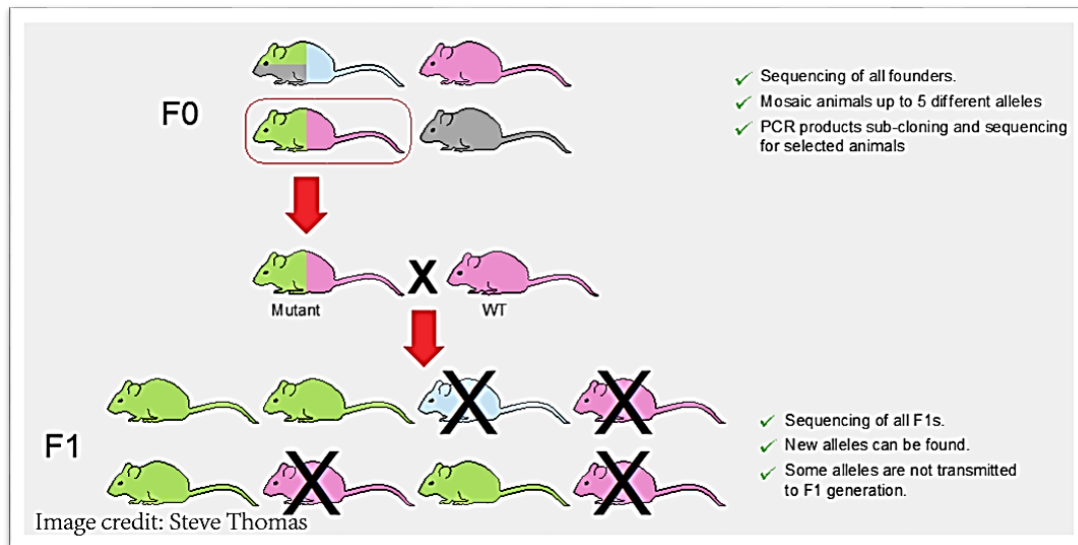
Validation of guide RNAs

- We recommend validating each gRNA efficiency, independently and in combination in case of a multiple gRNA injection.
- RNA preparations (home-made or commercial) should be tested by electrophoresis (capillary preferentially, or agarose). Note that for experiments involving co-delivery of a donor DNA, stability of the mix RNA/DNA should also be ascertained.
- sgRNA/enzyme complex cleaving activity is easily validated in vitro on PCR synthetized targets. Commercial kits are available (eg Surveyor , Guide-It).
- More advanced validation of reagents involves assaying in the context in which they will be employed (analysis on transfected cells or after in vitro culture of micro-injected embryos). It remains unclear whether activity in one context is predictive of that in others. The choice of rapid or in-depth activity validation depends on the context of use, as most sgRNAs show at least some activity.

The mosaicism question

A critical challenge when injecting CRISPR reagents into zygotes is mosaicism, partly due to a slow rate of nuclease-induced mutagenesis (Yen et al. 2014; Zhong et al. 2015). Indeed, many different allele variants can be found in one single founder (see below).

Definition of mosaicism and trans-heterozygote

Mosaicism in genetically modified mice generated by CRISPR/Cas

In genetics, a mosaic or mosaicism denotes the presence of two or more populations of cells with different genotypes in one individual who has developed from a single fertilized egg. Mosaics may be contrasted with chimerism, in which two or more genotypes arise from the fusion of more than one fertilized zygote in the early stages of embryonic development. Heteroallelic combination or trans-heterozygote refers to a gene that carries two different mutated alleles. Organisms with one mutant and one wildtype allele at one locus are called simply heterozygous, not trans-heterozygous. In a founder animal, both mosaicism and trans-heterozygoty can occur. These events can make elucidating the genotype of founders complicated. This is why we prefer to confirm genotype at the F1 stage (after breeding of the FO with WT animal) because each F1 will carry only one F0 derived allele.

What to know about CRISPR/Cas induced mosaicism

When generating a model, you have to take into account the high probability of mosaic F0 (founder) animals

1. Mosaicism is found in multiple species.
2. Animal genotyping from biopsy (ear, tail or finger) or embryos does not completely reveal the genetic makeup of a founder.

How to establish animal or cell lines?

- F0 animals are not established lines because of mosaicism.

- Phenotype of F0 animals should be interpreted with care: Each F0 will **generally** carry a different combination of mutated alleles.
- F0 founders should be crossed with wild type animals. Segregation events in F1s will reveal the different mutated alleles contained in F0 (some alleles may not have been detected by F0 genotyping, see Genotyping FAQ below for genotyping details).
- Each F1 allele will correspond to a well-defined line.
- Take into account the high probability of mixed cell populations for cell culture experiments (sub-clone to obtain uniform genotype clone).

How to reduce mosaicism ?

- To date, there are no clear strategies to be advised to reduce mosaicism. Thus, CRISPR nuclease encoded by DNA, RNA, or directly the protein itself, can be successfully used according to animal species.
- Direct microinjection of recombinant Cas9 protein instead of RNA may help reduce mosaicism (Singh et al. 2015) but additional experiments are clearly required to confirm this hypothesis (see Cas9 and guide RNA: What to microinject?). Be aware that the protein can degrade and so specific activity reduces with time (when a single tube is used).

Genotyping

Choice of genotyping method depends on experimental design and mutation type

Depending on the experimental design and intended mutation type, one of the following methods can be employed to define the outcome of CRISPR mediated mutagenesis:

- PCR product sequencing is our preferred option. Note that PCR amplicon should exceed the boundaries of the mutated site, as additional sequence changes can occur in the vicinity of the repair event (typically up to 200 bp away (mouse) and in some instances up to 1kb away (rabbit) deletion). For founders with high levels of mosaicism level, sequencing can be uninterpretable (presence of multiple traces on the chromatograms).
- Cloning of the PCR product: PCR has to be performed with a high fidelity DNA polymerase and PCR products have to be sequenced individually. They do not reveal the whole complexity of the genetic makeup of F0 animals as these are often mosaic.
- T7 endonuclease 1 or Cel1-based assay, High Resolution Melt and identification of heteroduplexes by capillary electrophoresis: these are more affordable and scalable options but do not yield sequence data of new alleles. Also, these do not identify homozygous changes unless pre-hybridisation with wild-type DNA is performed in addition. Also, single point mutations are not always detectable.
- High throughput sequencing: cell-batches or F0 animals are individually bar-coded and sequenced in a library format. This is applicable when large numbers of events must be screened (cell culture in some instances or large-scale micro-injection program).
- Some modifications (NHEJ) have been found up to 200 bp away of the Cas9 cleavage site. This can be in addition to the integration of the intended donor sequence. Therefore, it is worth amplifying PCR fragments larger than the expected targeted size (ideally perform 2 PCRs).
- **In the case of integration of a >100 bp knock-in target by homologous recombination**, combination of external and internal probes Southern Blot and qPCR assay are required to ascertain locus integrity in the case of complex targeting events.

Choice of method depends on mutagenesis context (tissue culture or live animals)

- Mutagenesis in cell culture: method must accommodate the screen of large numbers of events. Clonal quality of mutated population might affect the complexity of genotype.

- Mutagenesis in embryos :
  - o F0 animals (born from mutagenised embryos) are often mosaic (*see "The mosaicism question"*). Moreover, their genotyping is not entirely predictive of their offspring. Strategy for **F0 genotyping aims at identifying the 'best' founders to breed to generate F1s (containing the expected mutation). The genotyping of the F1s ascertains the definitive mutation.**
  - o F1 animals: F1 animals should be obtained by breeding F0 with wild-type stock. F1 genotypes are not necessarily predictable from that of their parents. Sequencing, and in some instances Southern blotting for more complex targeting designs, are required to define the genetic modification of all new lines.
- Identification of mutant by phenotyping is possible in some instances: i.e. eye marker in drosophila or coat color or expression of fluorescent marker in mouse/rat/rabbit.

Choice of method also depends on mutagenesis strategy

Briefly, we recommend for:

- Gene disruption by small insertion or deletion (NHEJ-mediated frame-shift mutations): any of the above method may apply, choose most convenient.
- Gene disruption using 2 sgRNAs (designed to induce DNA breaks flanking a target exon): any of the above method may apply, choose most convenient.
- Point mutation: sequencing is preferable but insertion of a silent mutation to introduce a restriction site very close to the mutation is much more convenient for genotyping purposes.
- Complex gene targeting: in depth characterization will eventually be necessary after initial screen. Typically PCR/qPCR is employed for screening and Southern blot is employed for full characterization of a new line (animal or cell).

Off target effects

- Cas9 complexed with a sgRNA can cleave DNA sequences that contains a one- or more base mismatch(es) with the intended target sequence (Jinek et al. 2012; Cong et al. 2013; Fu et al. 2013; Hsu et al. 2013). CRISPR/Cas nucleases can tolerate mismatches especially in the 5' region, but not in the seed region (6–11 nt that is immediately upstream of the PAM sequence; Pattanayak et al. 2013).
- To identify appropriate guide RNAs, you can use the Tefor web site (http://crispor.tefor.net/crispor.cgi) (optimised algorithm to avoid off-targets, possibility to search for sgRNAs used with different Cas9 variants, identification of off targets on the target chromosome, many different genomes available) or the MIT web site http://crispr.mit.edu/.
- Cas9 nickase: use of 2 efficient sgRNAs with Cas9 nickase is the best way to avoid or at least to minimise off-targets but at the cost of efficiency.
- *In vivo* (after pronuclear injection), off-targets seem to be less of a problem. A very recent study using whole-genome sequencing (detection of more than 95% heterozygous variants) to assess damages induced by microinjection of Cas9 or Cas9 nickase, concludes that undesired off-target mutations are very rare *in vivo* (Iyer et al. 2015). Moreover, by breeding the line with wild type animals, segregation of off-targets is easy.
- In cell culture (other than ES cells which will give rise to mice after microinjection), off-targets are a real issue as allele segregation is not possible.
- How to choose the off targets that should be analysed?
  - o This is the real difficulty. You can look at the 5 or 10 highest ranked potential off-targets but other criteria might be taken into account. For example, target site in exons, algorithm of the search program, etc. (One might ask why the next potential off-target on the list would not be the real off-target). When working in

vivo, one could choose to focus on the sites that are on the same chromosome as the target site as they would be more difficult to segregate.

- How to choose the off targets that should be analyzed?
  - o This is the real difficulty. You can look at the 5 or 10 highest ranked potential off-targets but other criteria might be taken into account. For example, target site in exons, algorithm of the search program, etc. (One might ask why the next potential off-target on the list would not be the real off-target). When working in vivo, one could choose to focus on the sites that are on the same chromosome as the target site as they would be more difficult to segregate.
- How to analyze off-targets :
  - o Start off-target analysis after establishment of the line (not at the founder stage) and only use the animals you have already checked for further breeding.
  - o Whole-genome sequencing DNAseq: a very high coverage of the genome is required for the detection of most off targets. However whole-genome sequencing results might be difficult to interpret and this method is still quite expensive.
  - o PCR for detection of NHEJ events (sequence the PCR product or capillary electrophoresis being quite sensitive, you can visualize the presence of heteroduplexes).
  - o PCR + T7 assay centered on the potential off-target sites: this approach works in most cases but is not completely reliable (it may not detect difference of 1 nt).

Validation of knock-in generated by CRISPR/Cas HR assisted recombination

- Use Long Range PCR to verify correct integration.
- Be aware of the possibility of multiple site integration at the locus; count the copy number of the donor DNA by qPCR or Southern.
- Only Southern blot analysis will support a characterization of the structure of the mutated locus.

Other resources about CRISPR

CRISPR in the Lab: A Practical Guide https://www.addgene.org/CRISPR/guide

References

Chen, F. et al., 2011. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nature Methods*, 8(9), p.753–755.

Cong, L. et al., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, 339(6121), p.819–823.

Doudna, J.A. & Charpentier, E., 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*, 346(6213), p.1258096.

Fu, Y. et al., 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology*, 31(9), p.822–826.

Garneau, J.E. et al., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, 468(7320), p.67‑71.

Guilinger, J.P., Thompson, D.B. & Liu, D.R., 2014. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nature Biotechnology*, 32(6), p.577‑582.

Hou, Z. et al., 2013. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. *Proceedings of the National Academy of Sciences of the United States of America*, 110(39), p.15644‑15649.

Hsu, P.D. et al., 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, 31(9), p.827‑832.

Iyer, V. et al., 2015. Off-target mutations are rare in Cas9-modified mice. *Nature methods*, 12(6), p.479–479.

Jinek, M. et al., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, 337(6096), p.816‑821.

Makarova, K.S. et al., 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nature Reviews. Microbiology*.

Mali, P. et al., 2013. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*, 31(9), p.833‑838.

Pattanayak, V. et al., 2013. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology*, 31(9), p.839‑843.

Ran, F.A., Hsu, P.D., Lin, C.-Y., et al., 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, 154(6), p.1380‑1389.

Ran, F.A., Hsu, P.D., Wright, J., et al., 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), p.2281‑2308.

Ran, F.A. et al., 2015. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*, 520(7546), p.186‑191.

Singh, P., Schimenti, J.C. & Bolcun-Filas, E., 2015. A Mouse Geneticist's Practical Guide to CRISPR Applications. *Genetics*, 199(1), p.1‑15.

Tsai, S.Q. et al., 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature Biotechnology*, 32(6), p.569‑576.

Yen, S.-T. et al., 2014. Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. *Developmental Biology*, 393(1), p.3‑9.

Zetsche, B., Volz, S.E. & Zhang, F., 2015. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nature Biotechnology*, 33(2), p.139‑142.

Zhang, Y. et al., 2013. Processing-independent CRISPR RNAs limit natural transformation in Neisseria meningitidis. *Molecular Cell*, 50(4), p.488‑503.

Zhong, H. et al., 2015. CRISPR-engineered mosaicism rapidly reveals that loss of Kcnj13 function in mice mimics human disease phenotypes. *Scientific Reports*, 5, p.8366.