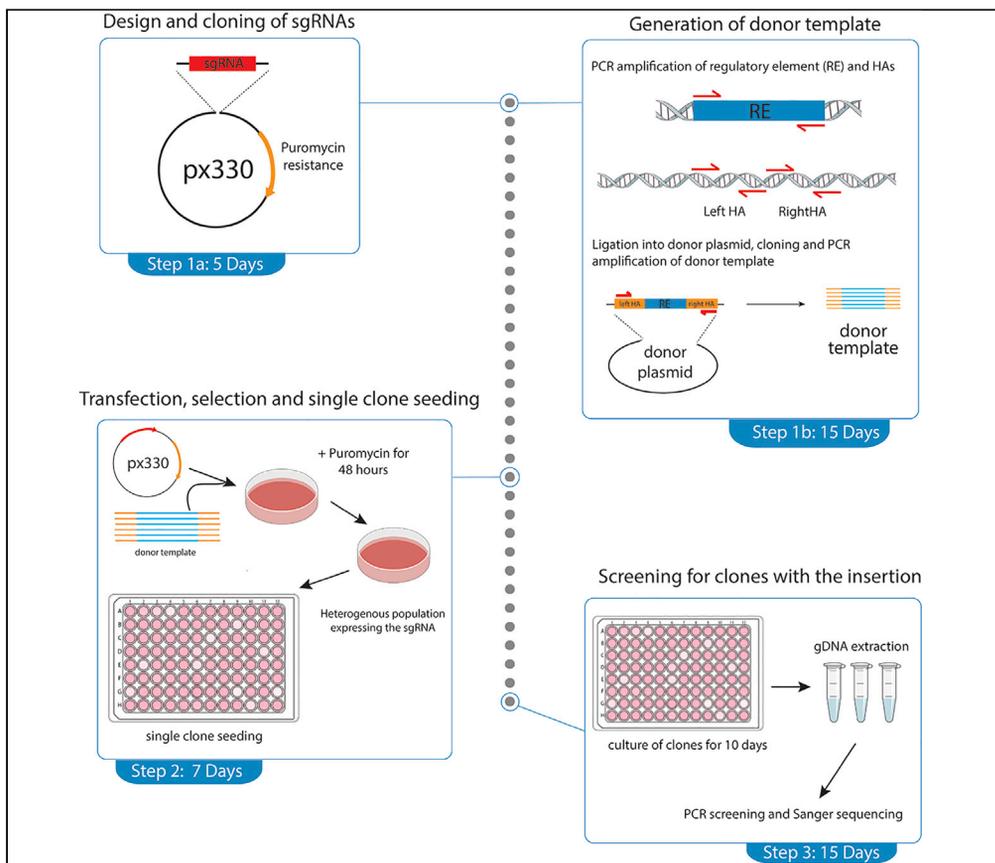


Protocol

Protocol to study sufficiency of *cis*-regulatory elements in mouse embryonic stem cells using a CRISPR-mediated knockin approach



Tomas Pachano,
Alvaro Rada-Iglesias

tomas.pachano@imp.ac.at (T.P.)
alvaro.rada@unican.es (A.R.-I.)

Highlights

CRISPR-Cas9 knockin of CREs in mouse embryonic stem cells

Scar-free knockin of CREs in genomic context

Design, transfection, and screening of CREs in mouse embryonic stem cells

Useful to study sufficiency of CREs in gene expression control

cis-regulatory elements (CREs) orchestrate the spatiotemporal control of gene expression. The regulatory activity of CREs is typically assessed by reporter assays, in which CREs are studied outside their endogenous context. To circumvent this problem, we developed a CRISPR-Cas9 knockin approach to study CREs in a scar-free genomic context. Here, we describe the design, transfection, and screening protocol to insert CREs in mouse embryonic stem cells. Our strategy can provide important insights into the sufficiency of CREs for gene expression control.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Pachano & Rada-Iglesias,
STAR Protocols 3, 101492
September 16, 2022 © 2022
The Author(s).
<https://doi.org/10.1016/j.xpro.2022.101492>



Protocol

Protocol to study sufficiency of *cis*-regulatory elements in mouse embryonic stem cells using a CRISPR-mediated knockin approachTomas Pachano^{1,2,3,*} and Alvaro Rada-Iglesias^{1,4,*}¹Institute of Biomedicine and Biotechnology of Cantabria (IBBTec), CSIC/Universidad de Cantabria, Albert Einstein 22, 39011 Santander, Spain²Research Institute of Molecular Pathology (IMP), Vienna Biocenter (VBC), Campus-Vienna-Biocenter 1, 1030 Vienna, Austria³Technical contact⁴Lead contact*Correspondence: tomas.pachano@imp.ac.at (T.P.), alvaro.rada@unican.es (A.R.-I.)
<https://doi.org/10.1016/j.xpro.2022.101492>

SUMMARY

cis-regulatory elements (CREs) orchestrate the spatiotemporal control of gene expression. The regulatory activity of CREs is typically assessed by reporter assays, in which CREs are studied outside their endogenous context. To circumvent this problem, we developed a CRISPR-Cas9 knockin approach to study CREs in a scar-free genomic context. Here, we describe the design, transfection, and screening protocol to insert CREs in mouse embryonic stem cells. Our strategy can provide important insights into the sufficiency of CREs for gene expression control.

For complete details on the use and execution of this protocol, please refer to Pachano et al. (2021).

BEFORE YOU BEGIN

The combination of CRISPR-Cas9 technologies and custom DNA templates can be used to efficiently insert DNA sequences into a specific region of the genome, a method referred to as knockin (Adli, 2018; Doudna and Charpentier, 2014; Hsu et al., 2013). Briefly, a gRNA is used to direct a Cas9-mediated DNA double-strand break (DSB) at a targeting site. Then, cells repair the DSB by either non-homologous end joining or by homology-dependent repair (HDR). The latter can be used to rationally insert a desired sequence by transfecting the cells with a donor DNA template, which contains certain homology with the DSB region.

Traditionally, knockins are performed with a selection cassette. That is, the inserted DNA contains a gene that expresses a fluorescence marker and/or an antibiotic resistance. Consequently, cells that are positive for the integration can be selected from the population, reducing the time and effort required for establishing clones with knockin alleles. However, in some cases it might not be convenient to have an actively expressing gene next to the studied sequence, as it might affect the chromatin and genetic landscape at the insertion site. To circumvent this problem, knockins can be associated with the Cre-loxP strategy to remove the drug resistance or the fluorescence markers after selection. However, this method still leaves a small “scar” in the targeted locus (i.e., loxP sequence). Moreover, the expression of the selection marker might also change the chromatin configuration of the endogenous locus and nearby genes and these chromatin changes might be preserved even after the selection marker is removed. The caveats of using selection markers can be particularly problematic when studying regulatory elements, as the activity of CREs might be sensitive to changes in the chromatin and/or genetic composition of the knockin site. Here, we describe a scar-free CRISPR-



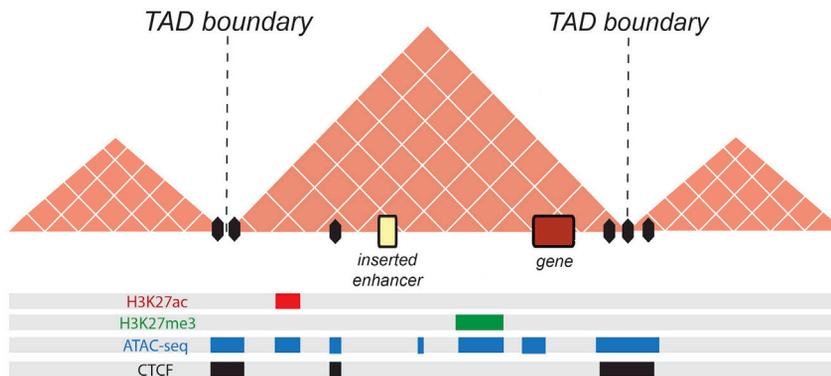


Figure 1. Example of how to choose the appropriate knockin coordinates for an enhancer insertion

To avoid a potential effect of the local chromatin state, the enhancer should be inserted in a region devoid of histone marks, CTCF or other regulatory elements. Importantly, it should be inserted into the same TAD as the studied gene. The pink triangles correspond to the TADs found around the locus of interest. The regions enriched in different chromatin features are displayed as colored bars.

Cas9 mediated knockin method to insert regulatory elements in a specific genomic locus with high efficiency and without the use of selection markers.

Before beginning, consider the CRE under study and the biological questions that you want to address, as this will determine which genomic locus should be used as an insertion site. When choosing the insertion site take into consideration: (i) the nearby genes, (ii) the presence of other regulatory elements (i.e., enhancers or silencers), (iii) accessible chromatin regions (i.e., ATAC-seq peaks), (iv) chromatin marks associated with regulatory functions (e.g., H3K27me3, H3K9me3, H3K27ac), (v) distribution of CTCF binding sites and (vi) 3D organization of the locus (i.e., topological associating domain (TAD) structure) (Figure 1). The protocol below describes the specific steps for generating knockins of enhancers using CRISPR-Cas9 in mouse embryonic stem cells (mESCs). However, other elements (e.g., silencers, insulators) can be studied in the same fashion. For using this protocol in other cell types or organisms, some steps might have to be adapted accordingly.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
NEB 5-Alpha	New England Biolabs	Cat: C2987
Chemicals, peptides, and recombinant proteins		
10× CutSmart Buffer	New England Biolabs	Cat: B7204
SpeI-HF	New England Biolabs	Cat: R0133L
MluI-HF	New England Biolabs	Cat: R3198L
NdeI	New England Biolabs	Cat: R0111L
NotI-HF	New England Biolabs	Cat: R0189L
BbsI	New England Biolabs	Cat: R0539L
T4 DNA Ligase	New England Biolabs	Cat: M0202S
T4 DNA Ligase Buffer 10×	New England Biolabs	Cat: B0202S
KCl	Roth	Cat: 0281.1
TRIS	Roth	Cat: 5429.1
MgCl ₂	Roth	Cat: KK36.2
NP40	Sigma	Cat: I3021
Tween 20	Roth	Cat: 9127.2
Proteinase K	Promega	Cat: mc5008

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
LB medium	Sigma-Aldrich	Cat: L2542
Glycerol	Millipore	Cat: CAS 56-81-5
Ampicillin	Millipore	Cat: CAS 69-52-3
OptiMEM	Gibco	Cat: 31985062
Trypsin	Gibco	Cat: 25200072
dNTPs	Thermo Fisher Scientific	Cat: 18427013
FBS	Gibco	Cat: 16141-061
Antimycotic/Antibiotic	Hyclone	Cat: SV30079.01
Beta-mercaptoethanol 55 mM	Gibco	Cat: 21985-023
Glutamax	Gibco	Cat: 35050-038
MEM NEAA	Gibco	Cat: 11140-035
Knockout DMEM	Gibco	Cat: 10829-018
LIF	Miltenyi Biotec	Cat: 130-095-777
Gelatin	Sigma-Aldrich	Cat: G9391-500G
Critical commercial assays		
Accuprime GC-rich Polymerase	Thermo Fisher Scientific	Cat: 12337016
Kapa HiFi HotStart ReadyMix	Roche	Cat: KK2602
QIAquick PCR & Gel Cleanup Kit	QIAGEN	Cat: 28506
FuGENE HD Transfection Reagent	Promega	Cat: E2312
PureYield Plasmid Miniprep System	Promega	Cat: A1222
Q5® High-Fidelity 2x Master Mix	New England Biolabs	Cat: M0492L
GoTaq polymerase	Promega	Cat: M7123
PureLink™ Genomic DNA Mini Kit	Thermo Fisher Scientific	Cat: K182001
Experimental models: Cell lines		
WT (E14) mESC	Wysocka Lab	n/a
Recombinant DNA		
Plasmid pX330A_hCas9_long_chimeric_gRNA_G2P	A gift from Leo Kurian Lab (CMMC, University of Cologne, Germany)	Available upon request
Oligonucleotides		
Right HA fwd	Sigma-Aldrich	Custom sequence with 5' NdeI site
Right HA rev	Sigma-Aldrich	Custom sequence with 5' SpeI site
Left HA fwd	Sigma-Aldrich	Custom sequence with 5' MluI site
Left HA rev	Sigma-Aldrich	Custom sequence with 5' NotI site
Donor template fwd	Sigma-Aldrich	Custom sequence
Donor template rev	Sigma-Aldrich	Custom sequence
Cloning_enhancer fwd	Sigma-Aldrich	Custom sequence with 5' SpeI site
Cloning_enhancer rev	Sigma-Aldrich	Custom sequence with 5' MluI site
Software and algorithms		
Benchling CRISPR Design and analyze guides	www.benchling.com	n/a
Other		
C1000 Touch PCR thermal cycler	Bio-Rad	Cat: 1851148
Wide Mini-Sub Cell GT Horizontal Electrophoresis System	Bio-Rad	Cat: 1704468
ZOE™ Fluorescent Cell Imager	Bio-Rad	Cat: 1450031
TC20 Automated Cell Counter	Bio-Rad	Cat: 1450102

MATERIALS AND EQUIPMENT

mESC growth medium

Component	Final concentration	Amount
FBS	15%	88 mL
Antimycotic/Antibiotic	1 ×	5.9 mL
Beta-mercaptoethanol 55 mM	0,1 mM	1.3 mL

(Continued on next page)

Continued

Component	Final concentration	Amount
Glutamax	2 mM	5.9 mL
MEM NEAA	1x	5.9 mL
Knockout DMEM		500 mL
LIF	0.045 pg/mL	0.007 mL

Store at 4°C for up to 2 months.

Primers

Primer name	Restriction enzyme
right HA fwd	NdeI
Right HA rev	SpeI
Left HA fwd	MluI
Left HA rev	NotI
Donor template fwd	–
Donor template rev	–
Cloning_enhancer fwd	SpeI
Cloning_enhancer rev	MluI

STEP-BY-STEP METHOD DETAILS

Design and cloning of gRNAs

⌚ Timing: 5 days

This section describes how to design and clone the CRISPR-Cas9 vector that will be used for inserting the DNA of interest.

1. Design gRNAs.

- a. Create an account in www.benchling.com, log in, start a new project and import a DNA sequence by selecting the desired chromosomal region and genome.

Note: Alternatively, other gRNA design software can also be used (e.g., https://eu.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM).

- b. Once on the imported sequence, select the CRISPR Menu on the right and then click on “Design and analyze guides” and select: single guide, 20 bp length, the corresponding genome and the “NGG” PAM.

Note: These settings are suitable for the Cas9 that we use (human codon-optimized *Streptococcus pyogenes* Cas9 (hSpCas9)). In case of using a less common Cas nuclease, make sure the appropriate PAM sequences are selected.

- c. In the “Sequence map” tab, select the region in which you want to design a gRNA. Then, in the “Design CRISPR” tab, click on the “+” symbol.
- d. Select a gRNA with high “On-target” (Doench et al., 2016) and “Off-target” (Hsu et al., 2013) scores.

Note: Scores above 60 are considered adequate.

2. Cloning of gRNAs.

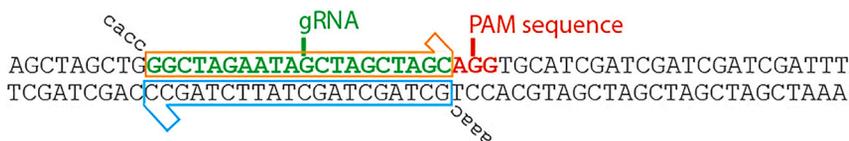


Figure 2. Example of how to design the gRNA

The “forward” primer should contain a cacc overhang (in this example: 5’ caccGGCTAGAATAGCTAGCTAGC 3’). The “reverse” primer should start with a aaac overhang (in this example 5’ aaacGCTAGCTAGCTATTCTAGCC 3’).

For each selected gRNA order two complementary oligos with different overhangs: a “forward” oligo with a cacc 5’ overhang and a “reverse” oligo with a aaac 5’ (Figure 2).

Note: These overhangs are used to generate restriction sites for the *BbsI* enzyme, and different sequences should be added in case other restriction enzymes are used.

- Order forward gRNA oligo with a 5’ cacc overhang.
- Order reverse gRNA oligo with a 5’ aaac overhang.
- Mix 1 μ L of each oligo (100 μ M stock in TE) together with 8 μ L of H₂O (10 μ L final volume).
Anneal using the following conditions:
 - 5 min at 98°C.
 - Cool at 5°C/min rate down to 25°C.
 - Dilute reaction 1:200 in H₂O. Annealed oligos can be stored at –20°C.

3. Prepare the CRISPR-Cas9 vector.

Note: CRISPR-Cas9 is used to generate a DSB at a specific site, which dramatically improves the insertion of a donor DNA template by HDR. To this purpose, a single guide RNA (gRNA) that can recognize a 20-nucleotide target sequence is expressed together with a Cas9 enzyme. Then, both the gRNA and Cas9 form a complex that specifically binds to the target site and generates a DSB. In this protocol, both the gRNA and Cas9 are transfected into cells and transiently expressed from the same plasmid.

Note: A variety of CRISPR-Cas9 expressing vectors are available on Addgene. In this example, we used the *px330A_hCas9_long_chimeric_gRNA_G2P* vector, which is not listed on Addgene. This vector expresses hSpCas9; GFP fused by T2A to a puromycin-resistance gene; and the inserted gRNA (<https://benchling.com/s/seq-sHYI3xBkQ3weqGOaQZDZ?m=slm-vUjeRzZzAmSvmf3h868G>). Other vectors can be used, but it is important that they express an antibiotic resistance gene to select for positively transfected cells. We recommend puromycin resistance since it enables a more rapid selection in comparison with other antibiotics. For example, pSpCas9(BB)-2A-Puro (PX459) V2.0 from Addgene (Plasmid #62988).

- Digest vector for gRNA insertion with the corresponding restriction enzyme. For the *px330A* vector, we use *BbsI*.

Vector	1 μ g
Restriction enzyme	2 μ L
Restriction enzyme buffer 10 \times	2 μ L
H ₂ O	to 20 μ L
Overnight (from 12 to 24 h) at 37°C.	

- Run the digested product in a 1.5% agarose gel.

- c. Gel purify the linearized vector using a commercial kit (e.g., QIAquick PCR & Gel Cleanup Kit).
- d. Ligate annealed oligos with CRISPR vector:

Annealed oligos (from step 2.c)	1 μ L
Digested vector (from step 3.a)	50 ng
Ligase	1 μ L
Ligase buffer 10 \times	1 μ L
H ₂ O	To 10 μ L
30 min at 25°C.	

- e. Transformation:
 - i. Add 2.5 μ L of the ligated reaction to 30 μ L of chemically competent cells (NEB 5-Alpha).
 - ii. Chill on ice.
 - iii. Heat-shock for 1 min at 37°C and then 1 min on ice.
 - iv. Add 1 mL of LB and incubate 30 min at 37°C.
 - v. Plate transformed bacteria on LB-ampicillin plates and incubate at 37°C overnight (12–20 h).
 - vi. Pick a colony and grow overnight (12–20 h) at 37°C in 4 mL of LB-ampicillin.
 - vii. Take 1 mL of the culture and make a glycerol stock (20% glycerol + 80% LB-ampicillin) for long-term storage at –80°C.
 - viii. Purify the plasmid by performing a miniprep using 3 mL of the bacteria culture and a commercial kit of your choice (e.g., PureYield Plasmid Miniprep System; <https://www.promega.es/resources/protocols/technical-bulletins/101/pureyield-plasmid-miniprep-system-protocol/>).
 - ix. Quantify the plasmid and verify the sequence by Sanger sequencing.

Design and cloning of donor template

⌚ Timing: 10–15 days

This section describes how to design and clone the CRE that will be used as a donor template for the CRISPR-Cas9 knockin.

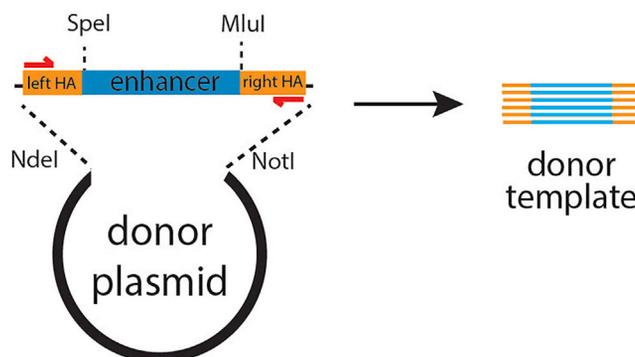


Figure 3. Generation of donor template

First, the HAs and the enhancer of interest are cloned into a donor plasmid using overlapping restriction enzymes. Then, the complete cassette is amplified by PCR and gel purified. The resulting dsDNA product is finally used to transfect the cells and to insert the candidate enhancer in the desired genomic locus by HDR. The red arrows correspond to the forward and reverse primers used to amplify the donor template.

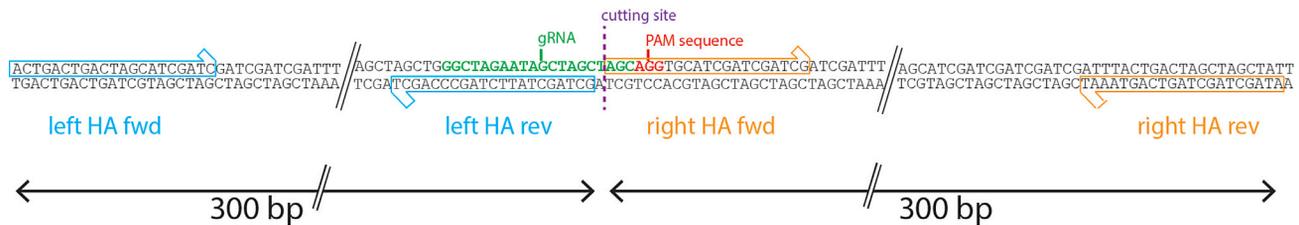


Figure 4. Example of how to design the HAs

HAs are amplified using primers flanking the gRNA recognition site. The right HA should extend from the cutting site (dashed line). The left HA should extend up to +1 from the cutting site.

Note: For obtaining the donor template, it is necessary to first clone each DNA fragment (enhancer and homology arms) into a donor plasmid and then amplify all of them together by PCR (Figure 3). Alternatively, it is also possible to synthesize the enhancer with the HAs and restriction sites and then directly clone this into the donor plasmid.

Note: To generate the donor template, we used a “donor vector” to separately insert the HAs and the enhancer sequence. Once all the elements are inserted in the vector, a PCR is performed to obtain a dsDNA with the enhancer sequence flanked by the two HAs. Since the donor vector will be only used as a PCR template to obtain the desired dsDNA, any vector with a multiple cloning site (MCS) can be used. Importantly, the MCS should contain at least 4 different and compatible restriction enzyme sites. In this example, we used a simple vector that contains a MCS, a bacterial origin of replication and a kanamycin resistance gene under the EM7 promoter (<https://benchling.com/s/seq-oEmdhAQmCYW2OsMVhtjh?m=slm-40wylnJzrZ4TzwErasYI>). For the rest of the protocol, we will consider the restriction enzymes present in this vector, which are NdeI-SpeI-MluI-NotI (from 5' to 3').

4. Design primers for amplification of homology arms (HAs) and the enhancer of interest.
 - a. Design two sets of primers to amplify the left and right HAs from genomic DNA (gDNA). Each HA should be designed to cover approximately 300 bp flanking the site of integration (Figure 4).

△ **CRITICAL:** To improve the efficiency of the knockin by HDR, it is recommended to leave 1 bp between both HAs, preferentially the +1 position from the cutting site (Figure 4) (Suzuki et al., 2016). In addition, each primer should contain a restriction site present in the multiple cloning site (MCS) of the donor plasmid (also with the same order). In this example, we used the primers “right HA fwd”, “right HA rev”, “left HA fwd” and “left HA rev” which are listed in the key resources table with their respective restriction sites (Figure 4). In addition, order “donor template fwd” and “donor template rev” primers, which will be used for amplifying the donor template (Figure 3).

- b. Design primers to amplify the enhancer of interest (“cloning_enhancer fwd” and “cloning_enhancer rev”). The forward primer should contain an overhang with the same restriction site as the “left HA reverse primer” and the reverse primer should contain the same restriction site as the “right HA forward primer”.
5. Cloning of HAs.
 - a. Amplify left and right HAs with a high fidelity polymerase (e.g., Kapa Hifi or Q5; see key resources table). Pure gDNA is preferred for this reaction (e.g., extracted with PureLink™ Genomic DNA Mini Kit). These are the conditions we used for the Kapa Hifi polymerase:

PCR components

Kapa Hifi HotStart ReadyMix	12.5 μ L
10 μ M Forward Primer	0.75 μ L
10 μ M Reverse Primer	0.75 μ L
gDNA	50 ng
H2O	To 25 μ L

PCR cycling conditions

Steps	Temperature	Time	Cycles
Initial Denaturation	98°C	3 min	1
Denaturation	98°C	20 s	35 cycles
Annealing	68°C	15 s	
Extension	72°C	10 s	
Final extension	72°C	1 min	1
Hold	4°C	Forever	

- b. Visualize the PCR product by electrophoresis in a 1.5% agarose gel and purify the band with a gel purification kit (e.g., QIAquick PCR & Gel Cleanup Kit).
 - c. Digest overnight (12–24 h) at 37°C approximately 1 μ g of each purified reaction with the chosen restriction enzymes. In our case, the left HA is typically digested with SpeI and NdeI; and the right HA with MluI and NotI.
 - d. Purify the digested PCR products with a column PCR purification kit (e.g., QIAquick PCR & Gel Cleanup Kit).
6. Prepare the donor vector.

Note: The CRISPR-mediated DSB triggers the HDR process to mediate a strand-exchange process with the DNA repair template, which contains a homologous sequence of the target site and the desired cassette (Figure 5). Repair templates can be used in different flavors: plasmids, short circularized DNAs, double-stranded linearized DNA (dsDNA), single stranded DNA (ssDNA) and dsDNA with ssDNA overhangs (Dokshin et al., 2018; Miura et al., 2018; Suzuki et al., 2016; Zhang et al., 2017). In this method, we use a linearized dsDNA since, in our hands, it shows greater efficiency compared to the other donor types.

Note: The homology arms (HAs) of the repair template should flank both sites of the integration cassette, since it improves the HDR rate (Suzuki et al., 2016). Therefore, the enhancer sequence of interest will be flanked by a left and right HA (Figure 5).

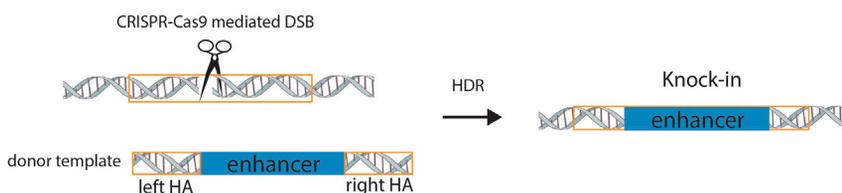


Figure 5. Inserting an enhancer into an endogenous locus via HDR

The CRISPR-Cas9 system is used to induce a DSB, which improves the insertion efficiency of a donor template with the enhancer of interest into a specific locus by HDR.

- a. Digest the donor vector with the first pair of restriction enzymes (i.e., NdeI-SpeI).

Vector	1 μ g
Restriction enzyme 1	2 μ L
Restriction enzyme 2	2 μ L
Restriction enzyme buffer 10 \times	2 μ L
H ₂ O	to 20 μ L

Overnight (from 12 to 24 h) at 37°C.

- b. Run the digested product in a 1.5% agarose gel.
- c. Gel purify the linearized vector using a commercial kit (e.g., QIAquick PCR & Gel Cleanup Kit).
- d. Ligate the left HA into the digested donor plasmid:
- Mix 200 ng or more of digested and purified left HA with 50 ng of the digested donor plasmid, 1 μ L of ligase and 1 \times ligase buffer. Add H₂O up to a final volume of 10 μ L.
 - Incubate 30 min at 25°C.
 - Transform chemically competent cells as previously described (step 3.e), amplify and purify the vector using a miniprep kit (e.g., PureYield Plasmid Miniprep System) and Sanger sequence it to confirm that the left HA has been successfully cloned. For sequencing the generated plasmid you can use any forward primer located between 50 to 200 bp upstream of the left HA.
- e. Ligate the right HA:
- Digest the donor plasmid containing the left HA with the restriction enzymes that correspond to the right HA (i.e., MluI and NotI).

Vector	700 ng
Restriction enzyme 1	2 μ L
Restriction enzyme 2	2 μ L
Restriction enzyme buffer 10 \times	2 μ L
H ₂ O	to 20 μ L

Overnight (from 12 to 24 h) at 37°C.

- Run the reaction in a 1.5% agarose gel and purify the band using a commercial kit.
 - Ligate the right HA as previously described for the left HA (Design and cloning of donor template: step 6.d).
 - Transform chemically competent cells as previously described (step 3.e), amplify and purify the vector using a miniprep kit and sequence it to confirm that the right HA has been successfully cloned.
- f. Digest overnight (12–24 h) at 37°C 700 ng of the donor plasmid containing both HAs with the restriction enzymes corresponding to “left HA rev” and “right HA fwd” (i.e., SpeI and MluI).
- g. Gel purify (e.g., QIAquick PCR & Gel Cleanup Kit) the digested plasmid and store at –20°C.
7. Cloning the enhancer of interest into the donor plasmid.
- a. Select the enhancer sequence you want to study:
- Amplify the enhancer sequence with a high fidelity polymerase (e.g., Kapa HiFi or Q5) and primers “Cloning_enhancer fwd” and “Cloning_enhancer rev”. These are the conditions we used for the Kapa HiFi polymerase:

PCR components

Kapa HiFi HotStart ReadyMix	12.5 μ L
10 μ M Forward Primer	0.75 μ L
10 μ M Reverse Primer	0.75 μ L
gDNA	50 ng
H ₂ O	To 25 μ L

PCR cycling conditions

Steps	Temperature	Time	Cycles
Initial Denaturation	98°C	3 min	1
Denaturation	98°C	20 s	35 cycles
Annealing	68°C	15 s	
Extension	72°C	1 min/kb	
Final extension	72°C	1 min	1
Hold	4°C	forever	

- ii. Visualize the PCR product by electrophoresis in a 1.5% agarose gel and purify the band with a gel purification kit.
- iii. Digest overnight (12–24 h) approximately 1 µg of the previous PCR product (i.e., the PCR product of the enhancer of interest) with *SpeI* and *MluI*.

Vector	1 µg
Restriction enzyme 1	2 µL
Restriction enzyme 2	2 µL
Restriction enzyme buffer 10×	2 µL
H ₂ O	to 20 µL

Overnight (from 12 to 24 h) at 37°C.

- iv. Purify the reaction with a PCR purification kit (e.g., QIAquick PCR & Gel Cleanup Kit).
- b. Ligate the enhancer sequence with the donor plasmid containing the HAs:
 - i. Mix 200 ng or more of digested and purified enhancer sequence with 50 ng of the previously digested donor plasmid, 1 µL of ligase and 1× ligase buffer. Add H₂O up to a final volume of 10 µL.
 - ii. Incubate 30 min at 25°C.
 - iii. Transform chemically competent cells as previously described (step 3.e), amplify and purify the vector using a miniprep kit and sequence it to confirm that the candidate enhancer and the HAs have been successfully cloned.
8. Generation of the donor template.
 - a. Amplify the donor template with the previously designed “donor template fwd” and “donor template rev” primers using a high fidelity polymerase. We typically use Kapa Hifi polymerase in this step.
 - b. Visualize the PCR product by electrophoresis in a 1.5% agarose gel and purify the band with a gel purification kit.

△ **CRITICAL:** Even though the PCR product is purified from the gel, the amplification of the donor template should not result in additional (unspecific) bands/products. The presence of unspecific bands might significantly affect the efficiency of subsequent integration events. If you see unspecific bands, try changing the conditions of the PCR (e.g., increasing the annealing temperature).

Transfection, selection, and establishment of clonal cell lines

⌚ **Timing:** 7 days

This section describes how to perform the knockin of CREs in mESC.

Note: The donor template and the CRISPR-gRNA-expressing vector will be used to generate knockins of the enhancer of interest in the previously selected genomic site.

9. Transfection of mESCs:
 - a. Prepare a gelatin coated well (6 well plate (wp)) by adding 0.1% gelatin (just enough to cover the whole well) and incubate them 2 h at 37°C.
 - b. Thaw a frozen stock of mESC, centrifuge cells at 300 × g for 3 min, remove the supernatant and resuspend in 1 mL of mESC growth medium (see [key resources table](#)).
 - c. Remove gelatin from the gelatin-coated plate and add 2 mL of mESC growth medium. Then, plate the resuspended cells.
 - d. Grow mESCs until approximately 80% confluent and prepare two gelatin-coated wells of a 12 wp (step 9.a).
 - e. Prepare the transfection reaction by mixing 2.5 μL of HD fusion reagent (see [key resources table](#)) with 250 ng of CRISPR-gRNA vector (px330 vector with the cloned gRNA), 500 ng of donor template and 80 μL of OptiMEM.
 - f. Incubate 15 min at 25°C.
 - g. While incubating, trypsinize cells with 500 μL trypsin (see [key resources table](#)) and plate a 1/20 dilution into two gelatin-coated wells of a 12wp (step 9.a) with 400 μL of mESC growing medium.
 - h. Once the 15 min incubation time finishes, add the transfection reaction into one of the wells. The other well should remain as a non-transfected control.
 - i. Incubate at 37°C overnight (12–24 h).

Note: For increasing efficiency, the transfection is performed as the cells are split.

10. Select transfected cells.
 - a. The transfection efficiency can be initially evaluated by visualizing the number of GFP positive cells in the transfected population. For an optimal transfection efficiency, you should see at approximately 80% of GFP positive cells. If you are using a CRISPR plasmid without GFP, this step can be omitted.
 - b. Change medium in both wells (transfected and negative control). The medium should contain 1.7 μg/mL of puromycin.

Note: This concentration is the optimal for selecting transfected cells in 48 h, but you could increase or decrease the concentration if the negative control is not dying or the negative control dies before 48 h, respectively.

- c. Incubate cells at 37°C until all the cells in the non-transfected well are dead (dead cells detach from the well surface). This should take approximately 48 h.
- d. Grow the transfected cell population in mESC medium without puromycin until confluent.

Note: This will take between 3 and 6 days depending on the initial transfection efficiency.

- e. Trypsinize the cell population with 250 μL of trypsin (see [key resources table](#)), inactivate with 500 μL of mESC growth medium and split in half: plate half of the cells in a gelatin-coated well of a 12 well plate (step 9.a) and the other half in a 1.5 mL tube.
- f. Extract gDNA from the 1.5 mL tube:
 - i. Spin down cells at 300 × g for 3 min (25°C) and remove supernatant.
 - ii. Resuspend cells in 25 μL of gDNA extraction buffer (50 mM KCl, 10 mM TRIS pH 8.3, 2.5 mM MgCl₂, 0.45% NP40, 0.45% Tween 20).
 - iii. Add 1 μL of 20 μg/mL of Proteinase K.
 - iv. Incubate 5 min at 65°C.
 - v. Incubate 1 min at 98°C.

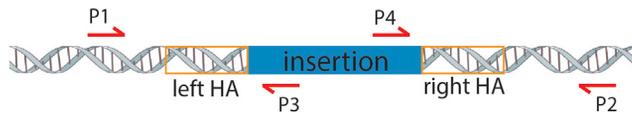


Figure 6. Primers used for PCR-based genotyping of transfected cells

Combination of P1+P3 and P4+P2 should only amplify a product if the desired insertion is present in the transfected cell population, but no product should be detected using WT gDNA. Primers P1+P2 should amplify a large and a short product corresponding to the inserted and WT alleles, respectively. In WT cells, primers P1+P2 should only amplify the short product. In the transfected cell population, both products should be ideally observed. However, only the short product is often observed in the transfected cells as it is preferentially amplified over the larger one.

- vi. Dilute 1:2 with H₂O.
- vii. Measure concentration with Nanodrop.
11. Screen for the presence of the desired knockin in the population by PCR genotyping.
 - a. Perform PCR reactions with primer pairs that either flank the homology arms or are located within the inserted enhancer (Figure 6).

Note: Ideally, PCR products should be in the range of 300–700 bp. Use gDNA from the transfected population and the GoTaq polymerase (M7123 Promega) with the following conditions:

PCR reaction master mix

Reagent	Amount
DNA template	200 ng
DNA Polymerase	0.07 μ L
Primer 1 (P1) 10 μ M	0.125 μ L
Primer 2 (P2) 10 μ M	0.125 μ L
GoTaq Green Master Mix 5 \times	2.5 μ L
MgCl ₂	2 μ L
dNTPs (10 mM)	0.25 μ L
ddH ₂ O	to 12.5 μ L

PCR cycling conditions

Steps	Temperature	Time	Cycles
Initial Denaturation	98°C	2 min	1
Denaturation	98°C	30 s	35 cycles
Annealing	59°C	30 s	
Extension	72°C	1 min/kb	
Final extension	72°C	5 min	1
Hold	4°C	forever	

- b. Visualize the PCR products by electrophoresis in an agarose gel to determine whether the desired insertion can be detected within the transfected cell population.

Note: That primers P1+P2 might only amplify a product corresponding to the WT allele even if the insertion is present in the population due to the smaller size of the WT product (Figure 6).

12. Single cell seeding.

- a. Grow the transfected population until it is confluent.
- b. Trypsinize the transfected population with 250 μ L trypsin, inactivate with mESC growth medium and split into three 1.5 mL tubes.
- c. Two tubes should be used for freezing the population in cryovials.

Note: This can be used as a backup if you encounter any problem in the following steps and you need to perform a new single cell seeding.

- i. Spin down cells at $300 \times g$ for 3 min.
- ii. Resuspend in 1 mL of Freezing medium (30% FBS + 20% DMSO + 50% mESC growth medium).
- iii. Save at -150°C .

Pause point: Alternatively, you can freeze the three tubes of cells and store them at -150°C until the next step is planned and started.

- d. The other tube is used to seed single cells in two full gelatin-coated 96wp (as step 9.a):
 - i. Dilute the cells 1:3 with medium and count the cells with a cell counter (we use TC20 Automatic cell counter from Bio-Rad).

CRITICAL: The cellular density should be between 5×10^5 and 2×10^6 (if the density is higher, dilute with medium and count again).

- ii. Make serial dilutions to obtain a cellular density of approximately 150 cells in 20 mL.

CRITICAL: Cell counting is only performed in the previous step. As a recommendation, perform 5 serial dilutions until obtaining the desired cellular density.

Note: As an example, if your initial count was 1×10^6 : transfer 200 μ L of the counted cells in 2.8 mL medium (this will be your dilution 1 and you would have $6,666 \times 10^4$ cells/mL). Then, transfer 500 μ L of dilution 1 in 4.5 mL medium (this will be your dilution 2 and you would have 6666 cells/mL). Transfer 500 μ L of dilution 2 in 4.5 mL medium (this will be your dilution 3 and you would have approx. 666 cells/mL). Transfer 1 mL of dilution 3 in 4 mL medium (this will be your dilution 4 and you would have approx. 133 cells/mL). Transfer X in final volume 20 mL (dilution 5). X should be calculated according to the initial cell count. In this example, $X = (150 \text{ cells} / 130 \text{ cells/mL}) \times 1 \text{ mL} = 1.13 \text{ mL}$.

- iii. Use the final dilution (150 cells in 20 mL) to plate 100 μ L in each of the wells of two 96 wp.
- e. Incubate both plates at 37°C .

Screening for clones with the insertion

⌚ Timing: 15 days

This section describes how to obtain mESC clones with the integrated CRE.

13. Expand single clones.
 - a. Culture clones (the seeded wells from step 12.e) for 5 days.
 - b. Use a microscope to look for wells that only have one colony.
 - c. Mark wells that have only one colony and supplement with 100 μ L of mESC growth medium.

Note: As the cells are seeded at very low density (<1 cell/well), expect some wells in each plate without any colony.

- d. Culture clones for another 5 days.
 - e. After the 5 days, trypsinize cells with 50 μ L of trypsin, resuspend with 100 μ L of mESC growth medium and transfer cells from each of the marked wells into larger wells of a gelatin-coated 48 wp.
 - f. Grow until confluent (between 2 to 5 days).
14. Screen clones by PCR genotyping.
- a. Trypsinize clones with 100 μ L of trypsin, resuspend with 200 μ L of mESC growth medium and split into two: half of the cells should be plated into wells of a gelatin-coated 48 wp with fresh medium, the other half should be used for gDNA extraction following the previously described protocol (step 10.f).
 - b. Perform PCRs with the same primers and conditions described above (Figure 6, step 11). Include PCRs with WT gDNA (extracted from WT mESC following the previously described protocol (step 10.f)) and the gDNA from the transfected population (extracted from transfected mESC from step 10.f) as negative and positive control templates, respectively.
 - c. Visualize the PCR reaction by agarose gel (1% agarose) electrophoresis to determine whether the desired insertion can be detected within the transfected cell population.
- △ CRITICAL:** in heterozygous clones, primers P1+P2 might only amplify a product corresponding to the WT allele due to its smaller size (Figure 6).
- d. Analyze the PCR products of the clones that are positive for the insertions by Sanger sequencing.
 - e. Expand or freeze the positive clones and discard the negative ones. For freezing clones, use the same protocol as before (step 12.c).

EXPECTED OUTCOMES

The expected outcome is the generation of clonal cell lines containing the desired insertion of an enhancer into a specific genomic position. This protocol was optimized to achieve a high number of recombination events without a selection marker. The efficiencies of integration events will depend on the targeted locus and the insert, but should range between 1% and 10% of homozygous clones. Once obtained, these clonal lines can be used to study the function and mechanism of action of the selected enhancer element.

We recently used this approach to investigate the role of orphan CpG islands in the cis-activation capacity of poised enhancers (see (Pachano et al., 2021) for more details).

LIMITATIONS

The described approach whereby knockins of selected enhancers can be achieved without the use of selection markers has the advantage of studying regulatory elements without the secondary effects that the selection marker might have over the studied enhancers. However, it is also more time-consuming, especially if multiple homozygous clones for the desired insertion are required. Therefore, the choice between the described protocol or an alternative one in which the regulatory element is inserted together with a constitutively expressed selection marker should depend on the biological question that wants to be answered.

One limitation to consider is that some genomic sites are more difficult to target than others and, therefore, it might take more time to get clones with the desired knockins. Moreover, the knockin sequences might also influence the efficiency of integration (Paix et al., 2017). For the combinations of genomic sites and cassettes that we tested, the more difficult combinations gave approximately 1% of homozygous insertions and about 5% of heterozygous. In case homozygous insertion is needed, it is also possible to retarget the WT allele in a heterozygous clonal line. However, this

requires careful sequencing of the WT allele as small indels might be generated by the CRISPR-Cas9 targeting.

Another limitation is the size of the insert. We noticed that longer sequences have lower integration rates. The longest DNA sequence we inserted was about 1.8 kb. In the case of longer insertions, some parameters might have to be changed. For example, increasing the length of the HAs from 300 bp to 900 bp might improve integration efficiency (Zhang et al., 2017).

TROUBLESHOOTING

Problem 1

There are no GFP positive cells after CRISPR transfection (steps 9 and 10).

Potential solution

For improving the transfection efficiency, three different steps should be considered:

The template DNA or the CRISPR plasmid might have low quality. It is crucial that both DNAs used in the transfection have high purity (A260/A280 should be approximately 1.8). If necessary, purify them again using PureLink™ Genomic DNA Mini Kit to ensure they are clean and without any contaminant that could affect the transfection. Also, using DNA extraction kits with the option to remove endotoxins could benefit the transfection efficiency.

Try reducing the number of cells, as the efficiency of transfection depends on the density of cells plated.

Try with a different amount of DNA, as the efficiency of the transfection may depend on the amount of template DNA and plasmid. You can try with more or less amount of DNA and assess if this improves the amount of GFP positive cells. It is important to maintain a ratio of approximately 1:2 (vector:donor template) to ensure a high number of recombination events. Reducing the donor template might improve the amount of GFP positive cells but might also significantly reduce the amount of recombination events and, thus, of cells with the intended insertions.

Problem 2

The PCR products obtained with the Kapa HiFi polymerase (steps 5.a and 7.a) or the GoTaq polymerase (steps 11.a and 14.a) show unspecific bands.

Potential solution

Adjust the annealing temperature within the range of +/- 5°C from the suggested starting temperature (i.e., 68°C for Kapa HiFi and 59°C for GoTaq polymerase).

Problem 3

Cannot detect insertions in the transfected population (step 11).

Potential solution

If the PCR product corresponding to the insertion event cannot be seen in the population, new transfections can be performed increasing the ratio of donor template (e.g., 1:5 gRNA:donor). Alternatively, this could also occur because the PCR conditions are not optimized. In this case, the design of new genotyping primers might be required. It is important to note that, for most insertions, the PCR primers (P1+P2) amplifying the whole insert will preferentially amplify the WT allele due to its smaller size, thus precluding detection of the whole insert in the transfected population. Lastly, insertions might not be detected in the transfected population because the gRNA used has a low cutting efficiency. To assess for cutting efficiency of the gRNA, you could use the T7 endonuclease I and Cell assay (i.e., Surveyor assay) (Qiu et al., 2004).

Problem 4

Insert is detected in the transfected population, but only mildly (i.e., weak PCR band) (step 11).

Potential solution

Insertions might not be strongly detected in the transfected population because the gRNA used has a low cutting efficiency. To obtain an efficient gRNA, you could assess for cutting efficiency of several gRNAs by using the T7 endonuclease I and Cell assay (i.e., Surveyor assay) (Qiu et al., 2004). Then, select the gRNA with higher cutting efficiency.

Problem 5

Cannot detect insertion in the clones (step 14).

Potential solution

If the insert is detected in the population but no positive clones are found, the knockin efficiency might be too low. This could be improved by doing new transfections with an increased ratio of donor template (e.g., 1:5 gRNA:donor). Alternatively, one can first detect the presence of the insert in small subpopulations and then derive clones from the positive subpopulations:

Plate between 5 to 10 cells/well in 20 wells of a 96wp with mESC growth medium. Briefly, repeat step 12.d but seeding 5–10 times more cells in the final dilution (5–10 times X). Then, let the subpopulations grow for 7 days in mESC growth medium and then split the cells from each well into two wells of a 48wp. Let the splitted cells grow for 3–4 days. After this, keep one well and use the other to extract gDNA. Use this gDNA to genotype the subpopulations by PCR. Then, seed single cells from positive subpopulations (i.e., subpopulations that contain the corresponding insert) in a 96wp as described in previous sections (step 12). Finally, screen for positive clones as described above (step 14).

Problem 6

When using the primers amplifying the whole insert (P1+P2), PCR products cannot be detected despite detecting a clear product (P1+P3; P2+P4) using an internal primer (step 14).

Potential solution

PCR amplification of the whole insertion might be difficult to achieve depending on the insert (e.g., CG rich amplicons). Consider the following:

WT alleles will be preferentially amplified in heterozygous clones due to their smaller sizes. Alternatively, an internal primer located on one end of the insertion and another primer located on the other end but outside the insertion can be used.

For the detection of whole inserts in homozygous clones, a potential solution is to try different polymerases. In our hands, we were able to amplify all the investigated inserts, even those with particularly high CG contents, with the Accuprime GC-rich polymerase.

Using high-quality gDNA significantly improves the efficiency of the PCRs, especially for the larger products. Try purifying the gDNA from the clones by phenol-chloroform extraction and ethanol purification and repeat the genotyping PCRs using the purified gDNA as templates.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Alvaro Rada-Iglesias (alvaro.rada@unican.es).

Materials availability

Plasmids and cell lines used in this study are available upon request.

Data and code availability

This study did not generate or analyze datasets of code.

ACKNOWLEDGMENTS

Work in the Rada-Iglesias laboratory is supported by the EMBO Young Investigator Programme, grant PGC2018-095301-B-I00 funded by MCIN/AEI /10.13039/501100011033 and by "ERDF A way of making Europe," grant RED2018-102553-T (REDEVNEURAL 3.0) funded by MCIN/AEI /<https://doi.org/10.13039/501100011033>, grant ERC CoG "PoisedLogic" (862022) funded by the European Research Council, and grant "ENHPATHY" H2020-MSCA-ITN-2019-860002 funded by the European Commission.

AUTHOR CONTRIBUTIONS

Methodology, T.P.; writing, T.P. and A.R.-I.; funding acquisition, A.R.-I.; supervision, A.R.-I.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* 9, 1911. <https://doi.org/10.1038/s41467-018-04252-2>.
- Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34, 184–191. <https://doi.org/10.1038/nbt.3437>.
- Dokshin, G.A., Ghanta, K.S., Piscopo, K.M., and Mello, C.C. (2018). Robust genome editing with short single-stranded and long, partially single-stranded DNA donors in *Caenorhabditis elegans*. *Genetics* 210, 781–787. <https://doi.org/10.1534/genetics.118.301532>.
- Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346, 1258096. <https://doi.org/10.1126/science.1258096>.
- Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. <https://doi.org/10.1038/nbt.2647>.
- Miura, H., Quadros, R.M., Gurumurthy, C.B., and Ohtsuka, M. (2018). Easi-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors. *Nat. Protoc.* 13, 195–215. <https://doi.org/10.1038/nprot.2017.153>.
- Pachano, T., Sánchez-Gaya, V., Ealo, T., Mariner-Faulí, M., Bleckwehl, T., Asenjo, H.G., Respuela, P., Cruz-Molina, S., Muñoz-San Martín, M., Haro, E., et al. (2021). Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness. *Nat. Genet.* 53, 1036–1049. <https://doi.org/10.1038/s41588-021-00888-x>.
- Paix, A., Folkmann, A., Goldman, D.H., Kulaga, H., Grzelak, M.J., Rasoloson, D., Paidemarry, S., Green, R., Reed, R.R., and Seydoux, G. (2017). Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc. Natl. Acad. Sci. USA* 114, E10745–E10754. <https://doi.org/10.1073/pnas.1711979114>.
- Qiu, P., Shandilya, H., D'Alessio, J.M., O'Connor, K., Durocher, J., and Gerard, G.F. (2004). Mutation detection using Surveyor nuclease. *Biotechniques* 36, 702–707. <https://doi.org/10.2144/04364pf01>.
- Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., Wu, J., Zhu, J., Kim, E.J., Hatanaka, F., Yamamoto, M., Araoka, T., Li, Z., et al. (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* 540, 144–149. <https://doi.org/10.1038/nature20565>.
- Zhang, J.-P., Li, X.-L., Li, G.-H., Chen, W., Arakaki, C., Botimer, G.D., Baylink, D., Zhang, L., Wen, W., Fu, Y.W., et al. (2017). Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* 18, 35. <https://doi.org/10.1186/s13059-017-1164-8>.