nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-020-0383-8

Check for updates

CRISPR-based engineering of gene knockout cells by homology-directed insertion in polyploid *Drosophila* S2R+ cells

Baolong Xia[1], Gabriel Amador[1,2], Raghuvir Viswanatha[1], Jonathan Zirin[1,2], Stephanie E. Mohr[1,2] and Norbert Perrimon[1,2,3]✉

Precise and efficient genome modifications provide powerful tools for biological studies. Previous CRISPR gene knockout methods in cell lines have relied on frameshifts caused by stochastic insertion/deletion in all alleles. However, this method is inefficient for genes with high copy number due to polyploidy or gene amplification because frameshifts in all alleles can be difficult to generate and detect. Here we describe a homology-directed insertion method to knockout genes in the polyploid *Drosophila* S2R+ cell line. This protocol allows generation of homozygous mutant cell lines using an insertion cassette which autocatalytically generates insertion mutations in all alleles. Knockout cells generated using this method can be directly identified by PCR without a need for DNA sequencing. This protocol takes 2–3 months and can be applied to other polyploid cell lines or high-copy-number genes.

Introduction

The CRISPR/Cas system can introduce DNA double-strand breaks at specific genomic sites targeted by guide RNAs (sgRNAs)[1–3]. Following the generation of double-strand breaks, the DNA repair machinery mediates repair by either the error-prone non-homologous end-joining (NHEJ) pathway or the high-fidelity homology-directed repair (HDR) pathway[4]. Previous gene knockout methods harnessed NHEJ to stochastically introduce insertions/deletions (indels) at the cleavage site in the coding region, a subset of which cause frameshifts, followed by identification of modified cells in which all alleles contain frameshift mutations in the target site[5–7]. Because of the stochastic nature of this process, the probability of obtaining frameshifts in all alleles decreases as copy number increases. Moreover, to identify frameshifts, DNA sequencing should cover all the alleles for each gene knockout clone, which is costly and time consuming.

HDR enables precise genome editing: a repair template containing homologous sequences flanking the DNA cleavage site generates a specific genome modification. Here we describe a gene knockout strategy based on homology-directed insertion[8]. The insertion cassette contains both sgRNAs and homology arms flanking the sgRNA target sites. Once sgRNA-bound Cas9 cleaves the target site, the sgRNA-containing cassette can insert into the target locus by HDR. After the cassette inserts into one allele, the sgRNA expressed from the insertion cassette can lead to further mutation of other alleles until all alleles have been replaced by the cassette. As a result, all alleles are mutated by the insertion cassette, resulting in gene knockout. Using this method, knockout cells carry homozygous insertion mutations, which can be directly identified by a simple PCR strategy without DNA sequencing.

Development of the protocol

This approach is an updated version of the previous gene knockout method established and used in research studies by our laboratory[9,10]. In our previous knockout protocol, we transfected plasmids conferring Cas9 and sgRNA expression into *Drosophila* S2R+ cells along with a third plasmid expressing GFP under the control of an actin promotor as a transfection marker. Next, we sorted single GFP-positive cells into micro-well plates by fluorescence-activated cell sorting (FACS) and established single-cell clones using conditioned medium. We screened these clones by high-resolution melt analysis to identify cell lines with indels[11]. Frameshifts caused by indels were further identified

[1]Department of Genetics, Harvard Medical School, Boston, MA, USA. [2]Drosophila RNAi Screening Center, Harvard Medical School, Boston, MA, USA. [3]Howard Hughes Medical Institute, Boston, MA, USA. ✉e-mail: perrimon@genetics.med.harvard.edu

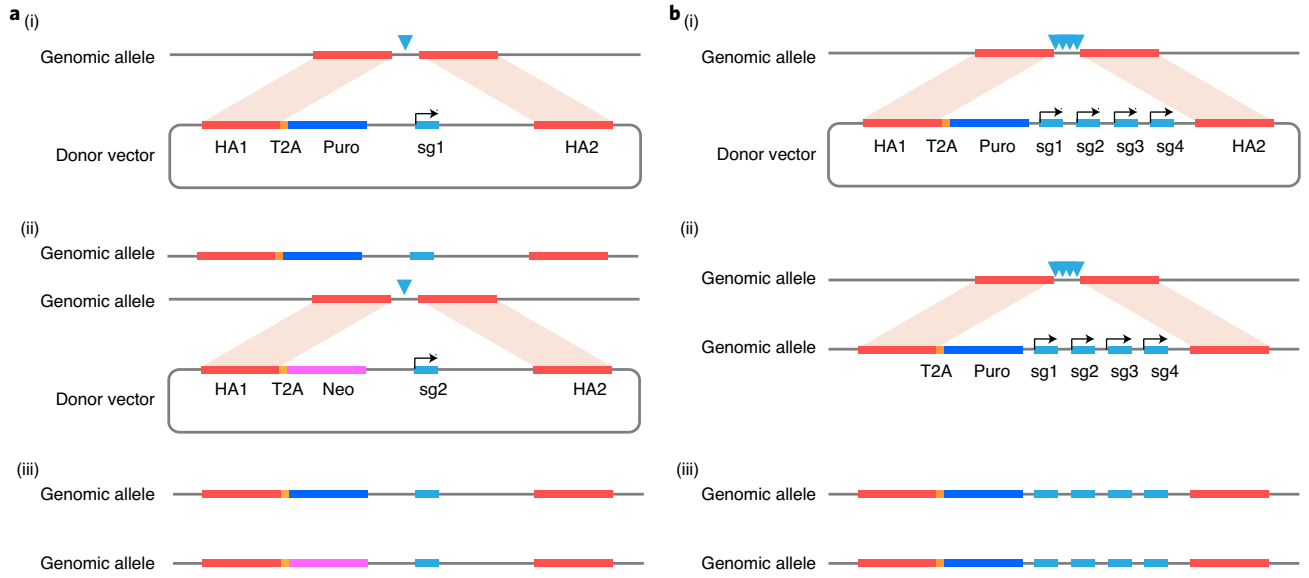

Fig. 1 | Overview of two homology-directed gene knockout methods. a, Procedure A: scheme comprising two rounds of homology-directed insertion. (i) For the first round of homology-directed insertion, the donor vector contains upstream and downstream homology arms (HA1 and HA2), a puromycin resistance gene (Puro) and sgRNA. The resistance gene does not have its own promoter and is driven by the promoter of the target gene and processed by T2A peptide cleavage. (ii) After puromycin selection, at least one allele has the insertion mutation. The second donor vector contains the same upstream and downstream homology arms (HA1 and HA2), a neomycin resistance gene (Neo) and a second sgRNA, which should target alleles that did not receive the cassette in the first round. (iii) After two rounds of targeting, at least two cassette insertion alleles can be generated (that is, insertion of one or the other cassette) and replacement of all alleles with cassettes results in gene knockout. **b**, Procedure B: a single round of homology-directed insertion with multiple sgRNAs. (i) The donor vector contains upstream and downstream homology arms (HA1 and HA2), a puromycin resistance gene and multiple sgRNAs. (ii) After the cassette inserts into one allele, sgRNA expressed from the insertion cassette can result in further mutation of other alleles. (iii) Replacement of all alleles by the cassette results in gene knockout. We note that with either procedure, a combination of cassette insertion and NHEJ-mediated indel alleles can also result in gene knockout.

by PCR amplification followed by TOPO cloning and Sanger sequencing of the PCR products. It is notable that extensive TOPO cloning and Sanger sequencing were required to identify cells with frameshifts in all alleles. These validation approaches can be costly and time consuming, especially for polyploid cells.

In this updated knockout protocol, instead of using stochastic formation of indels by NHEJ, we developed a method based on precise insertion mutation by HDR. The donor vector contains two ~1 kb homology arms, sgRNAs and a promoterless selectable marker. The promoterless selectable marker is designed to be in-frame with the target gene. The selectable marker can only be expressed after it inserts in-frame into the location we targeted, thus enriching correctly edited cells and excluding isolation of antibiotic-resistant cells generated by either non-inserted copies of the plasmid or by random insertion into the genome. In order to convert all alleles, we first tried performing two rounds of homology-directed insertion (procedure A and Fig. 1a). Procedure A was expected to be most appropriate for diploid cells, where, after two rounds of homology-directed insertion and selection, in theory all cells should be knockout cells. However, after two rounds of homology-directed insertion, non-insertion alleles remained in the polyploid *Drosophila* S2R+ cells. Thus, we subsequently developed another homology-directed insertion method (procedure B and Fig. 1b). With this approach, the donor vector contains four different sgRNAs to decrease the probability of sgRNA-resistant allele formation. After the cassette inserts into one allele, it can further mutate other alleles until all alleles have been replaced by the cassette. As a result, all alleles are mutated by the insertion cassette, resulting in gene knockout. With procedure B, we established three *Tnks* knockout cell lines and two *Apc* knockout cell lines (Supplementary Data 1 and 2, unpublished observations). All of these cell lines are available at the *Drosophila* Genomics Resource Center (DGRC #299, #300 and #301 for the *Tnks* knockout cell lines, DGRC #271 and #272 for the *Apc* knockout cell lines).

Applications

For polyploid cells, true gene knockout can only be achieved when every copy is mutated. As high copy numbers would hamper gene knockout efficiency, it is difficult—although not impossible—to

generate gene knockout cells in polyploid cell lines via a conventional NHEJ frameshift approach[12,13]. *Drosophila* S2R+ cells are pseudotetraploid cells, with four or more copies for most genes[14]. The gene knockout protocol with a homology-directed insertion cassette that we developed in *Drosophila* S2R+ cells should be applicable to other polyploid cell lines or for targeting any gene present in high copy number.

Limitations

This protocol has several limitations. First, we used a promoterless selectable marker in the insertion cassette. Although this should filter out nonspecific insertion and benefit precise targeting, the promoterless marker method requires that the target gene be expressed. Thus, silent genes or genes expressed only very lowly are not accessible by this protocol[15]. The method is also only feasible if the target gene is not essential for S2R+ cell viability/proliferation. Second, in this protocol, allele conversion by the insertion cassette relies on HDR. However, HDR efficiency depends on species, cell type and cell cycle stage[16]. In addition, although we used four different sgRNAs in procedure B to decrease the probability of sgRNA-resistant allele formation and increase the probability of allele conversion by the inserted cassette, we still observed resistant alleles in many cells. This protocol could be further improved by combining the approach with other methods to increase HDR efficiency[17–19]. Third, detection of knockouts using this protocol relies on PCR amplification of the target loci. However, large deletion alleles caused by CRISPR/Cas9 might not be amplified by PCR[20]. Although a large deletion can inactivate the target gene, it is worth noting that this might also induce genetic compensation. Further characterization, such as detection of RNA or allele copy number, can be applied.

Comparison with other methods

With previous frameshift indel-mediated knockout methods[5–7,9,10], after Cas9 and sgRNAs are delivered into cells, sgRNA-bound Cas9 causes DNA double-strand breaks at the target site, and random insertions or deletions can be introduced into this cleaved site during DNA repair without a template. Because the formation of indels during this process is stochastic, it results in a mixed population with different allele sequences. For polyploid cells or genes present in high copy number, gene knockout with this approach is inefficient because frameshifts in all alleles can be difficult to generate and detect. Moreover, the frameshift indel-mediated knockout approach relies on frameshift mutation. Recent studies showed that the premature stop codons generated by frameshifts can activate the RNA surveillance pathway and induce genetic compensation by upregulating paralog expression[21,22], complicating phenotypic analysis of putative gene knockouts.

Compared with the frameshift indel-mediated knockout approach, the insertion cassette-mediated knockout method presented in this protocol has three distinct advantages. First, the promoterless selectable marker in the inserted cassette can be used to enrich for the gene-targeted cells. Second, knockout cells generated by this method carry homozygous insertion mutations that can be directly identified by PCR without DNA sequencing, which allows us to screen a large number of colonies. Third, because this method disrupts gene loci without generating a premature stop codon, a genetic compensation response is unlikely to happen.

In this protocol, we describe two procedures for gene knockout using an insertion cassette. In procedure A, we use two sequential rounds of homology-directed insertion. In procedure B, we use a homology-directed insertion cassette containing four sgRNAs. Procedure A can be applied to diploid cells. After two sequential rounds of homology-directed insertion and selection, in theory all cells would be knockout cells. For polyploid cells, non-insertion alleles are likely to remain after two rounds of insertion, due to formation of alleles resistant to the single sgRNA present in each cassette. Because procedure B is based on the use of multiple sgRNAs, this protocol is useful for knockout in polyploid cells or for knockout of genes present in high copy number.

Experimental design

Gene target suitability

The promoterless selectable marker is in frame with the target gene and driven by the endogenous promoter of the target gene. Thus, for the strategy to work, it is important to ensure that the target gene is expressed in S2R+ cells (for example, using public data searchable with the *Drosophila* Gene Expression Tool[23,24], http://www.flyrnai.org/tools/dget/web/). The method is also only feasible if the gene is not essential for S2R+ cell viability/proliferation.

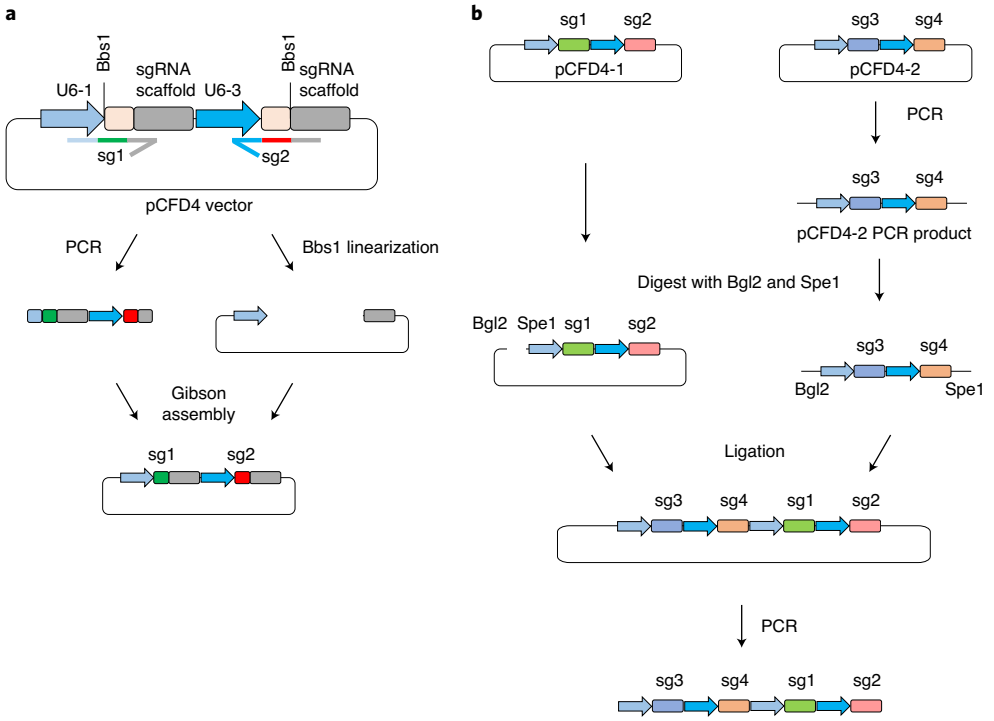

Fig. 2 | Construction of a four-sgRNA cassette using two pCFD4 vectors. a, Construction of pCFD4-1 vector for expression of two tandom sgRNAs. pCFD4 vector contains U6-1 and U6-3 promoters and two sgRNA scaffolds. The forward PCR primer contains the spacer sequence of sg1 and the reverse primer contains the reverse complementary sequence of the spacer sequence of sg2. After PCR amplification using pCFD4 as a template, the PCR product is ligated with Bbs1-linearized pCFD4 vector by Gibson Assembly. The pCFD4-2 vector can be constructed in the same way. **b**, pCFD4-1 contains sg1 and sg2; pCFD4-2 contains sg3 and sg4. The procedure involves PCR amplifying sg3 and sg4 from pCFD4-2, then digesting the pCFD4-1 vector and pCFD4-2 PCR product with Bgl2 and Spe1, and ligating to form a cassette containing four sgRNAs.

sgRNA design

In our previous study[9], based on luciferase assays and high-resolution melt assays, we optimized the CRISPR system for *Drosophila* cell culture with regards to sgRNA efficiency and minimizing off-targets, and established an online sgRNA design tool (www.flyrnai.org/crispr/). We use this online design tool for *Drosophila* sgRNA design. Nevertheless, we cannot exclude the possibility that NHEJ-mediated or other types of events could occur at off-target genomic locations. For procedure A, we use two sgRNAs 100−200 bp apart for two rounds of insertion. For procedure B, we use four different sgRNAs within a 300−500 bp region and avoid selecting sgRNAs with overlapping target sites to reduce the chance of resistant allele formation. The ideal sgRNA target sites are at or immediately upstream of an important domain of the target gene (using UniProt to analyze protein domains, http://www.uniprot.org/). For promoterless donor vector design, we ensure the sgRNA target sites are >1,000 bp away from the transcription start site, such that there should be no promoter sequence on the homology arm included in the donor plasmid. Testing the efficiency of individual sgRNAs with the T7 endonuclease I assay is recommended[25]. Briefly, denature and reanneal the PCR products of the target gene loci from wild-type cells and sgRNA-cleaved cells, then digest the annealed product with T7 endonuclease I. As T7 endonuclease I recognizes and cleaves mismatch DNA, the sgRNA efficiency can be evaluated based on the cutting efficiency of T7 endonuclease I.

To construct a cassette containing four different sgRNAs in procedure B, we ligate two sets of two sgRNAs in pCFD4 (ref. [26]) (Fig. 2a,b). pCFD4 vector construction follows the protocol (http://www.crisprflydesign.org/wp-content/uploads/2014/06/Cloning-with-pCFD4.pdf) (Fig. 2a). The forward primer contains the spacer sequence of the first sgRNA and the reverse primer contains the reverse complementary sequence of the spacer of the second sgRNA. After PCR amplification using pCFD4 as a template and insertion into linearized pCFD4, we obtain expression vectors containing two tandem sgRNAs. The four-sgRNA cassette is constructed by ligating a pCFD4 vector containing two sgRNAs with a PCR-amplified pCFD4 sgRNA insert region containing a different set of two sgRNAs (Fig. 2b). Alternatively, a four-tandem-sgRNA cassette can be constructed using pCFD5 with a tRNA-flanked sgRNA expression system[27].

Donor vector design

The donor vector contains two 1 kb homology arms, a promoterless selectable marker and sgRNAs. The T2A-Puro^r/Neo^r-SV40 polyA promoterless selectable marker cassette was directly cloned from existing vectors in our laboratory (pLib7.1, pLib7.2). It can also be cloned from other plasmids containing puromycin or neomycin resistance genes using PCR primers that amplify the open reading frame of the selectable marker and add the T2A sequence. We ligate homology arms, the selectable marker cassette and the sgRNA cassette to a pCR-Blunt vector (Gibson Assembly).

Cas9 provision

We use the S2R+-MT::Cas9 cell line to provide spCas9 activity[28] (DGRC cell cat. no. 268). Cas9 is constitutively expressed in this cell line. This protocol can also be applied using Cas9-expressing vector co-transfection or inducible Cas9 to avoid constitutive expression of Cas9.

Antibiotic selection

The antibiotic resistance genes (Puro^r/Neo^r) do not have their own promoters in the vector. Thus, they can only be expressed after they correctly insert into the target frame. Moreover, antibiotic selection can take effect early after vector transfection, without the need to wait for loss of unintegrated copies of the donor plasmid. Antibiotic selection should not be done with hygromycin, which is already present in S2R+-MT::Cas9 cells. To estimate the efficiency of antibiotic selection, wild-type S2R+ cells are treated with antibiotics as a control. Normally, for wild-type S2R+ cells, almost all cells die in 2 weeks with puromycin or G418 (a neomycin analogue) selection.

Single cell cloning

To eliminate heterogeneity after antibiotic selection, we sort single cells into individual wells in 96-well plates. It is not necessary to change the medium for 2−3 weeks during single-cell cloning. We avoid adding antibiotic to the conditioned medium because this decreases single-cell cloning efficiency. After 2−3 weeks, cell clones comprised of hundreds of cells can be observed. Using conditioned medium, the efficiency of forming clones from FACS-isolated single cells is 20–50%.

Identification of homozygous insertion mutations

To avoid false-negative results and nonspecific PCR products, we designed nested PCR primers to detect non-insertion alleles. These PCR primers target sequences flanking the sgRNA target sites. If all alleles include the insertion cassette, then no PCR product from non-insertion alleles will be detected. PCR products from non-insertion alleles from wild-type S2R+ cells serve as a control. To confirm that HDR occurred, a forward primer is designed based on the insertion cassette, and a reverse primer is designed based on the flanking genomic sequence outside of the homology arm region. Failure to observe these PCR products from wild-type S2R+ cells serves as a control. To further verify the cassette is correctly inserted into the target locus, we sequenced the target loci of the gene in knockout cell lines (Supplementary Data 2, unpublished observations).

Materials

Reagents

Vector construction

- Plasmid pCFD4 (Addgene, cat. no. 49411)
- Plasmids pLib6.4, pLib7.1, pLib7.2 (available on request)
- pCR-Blunt vector (Invitrogen, cat. no. K2700-20)
- PCR primers or oligos used in this study are listed in Table 1 (Integrated DNA Technologies)
- Nuclease-free water (Thermo Fisher Scientific, cat. no. AM9930)
- BbsI (NEB, cat. no. R0539; supplied with 10× NEBuffer 2.1)
- SpeI (NEB, cat. no. R3133; supplied with 10× CutSmart Buffer)
- Bgl2 (NEB, cat. no. R0144; supplied with 10× NEBuffer 3.1)
- QIAquick PCR purification kit (Qiagen, cat. no. 28104)
- QIAquick gel extraction kit (Qiagen, cat. no. 28704)
- T4 DNA ligase (NEB, cat. no. M0202, supplied with 10× T4 DNA ligase buffer)
- Phusion High-Fidelity DNA polymerase (NEB, cat. no. M0530, supplied with 5× Phusion GC Buffer and 5× Phusion HF Buffer) ▲ **CRITICAL** To minimize PCR error during vector construction, high-fidelity DNA polymerase is highly recommended. Other high-fidelity DNA polymerases, such as

Table 1 | PCR primers or oligos used in this study

Step	Primer	Sequence (5′–3′)	Purpose
Step 3 in procedure A	Forward oligo	GTCGNNNNNNNNNNNNNNNNNNN	Oligo for sgRNA vector construction with N20 sgRNA sequence
Step 3 in procedure A	Reverse oligo	AAACNNNNNNNNNNNNNNNNNNN	Oligo for sgRNA vector construction with N20 reverse complementary sequence
Step 10 in procedure A	pLib7-seq	CAGAGTAGAATGAAACGCCAC	Sequencing primer for sgRNA in pLib7.1 and pLib7.2
Step 11 in procedure A	Forward primer	AAGCTTGAGGGCAGAGGAAGC	Amplify cassette from pLib7.1 or pLib7.2
Step 11 in procedure A	Reverse primer	ACTCTCAGGCTCCAGGTAGGC	Amplify cassette from pLib7.1 or pLib7.2
Step 16 in procedure A	UHA forward primer	GCCGCCAGTGTGCTGGAATTNNNNNNNNNNNNNNN	PCR amplify upstream homology arm
Step 16 in procedure A	UHA reverse primer	GCTTCCTCTGCCCTCAAGCTTNNNNNNNNNNNNNNN	PCR amplify upstream homology arm
Step 16 in procedure A	DHA forward primer	GCCTACCTGGAGCCTGAGAGTNNNNNNNNNNNNNNN	PCR amplify downstream homology arm
Step 16 in procedure A	DHA reverse primer	TGATGGATATCTGCAGAATTNNNNNNNNNNNNNNN	PCR amplify downstream homology arm
Step 25 in Procedure A	M13 forward primer	GTAAAACGACGGCCAGT	Sequencing primer for pCR-blunt vector
Step 25 in procedure A	M13 reverse primer	CAGGAAACAGCTATGACC	Sequencing primer for pCR-blunt vector
Step 39 in procedure A	Forward primer	CGAAGGCAAGACCTTCTCGCTC	PCR of Tnks loci
Step 39 in procedure A	Reverse primer	GAGTAGGGCCAGCAGGCGATC	PCR of Tnks loci
Step 3 in procedure B	pCFD4-1 Forward primer	TCCGGGTGAACTTCGNNNNNNNNNNNNNNNNNNN GTTTTAGAGCTAGAAATAGCAAG	Primer for pCFD4-1 construction with N20 sequence of sgRNA1
Step 3 in procedure B	pCFD4-1 Reverse primer	TTCTAGCTCTAAAACNNNNNNNNNNNNNNNNNNN CGACGTTAAATTGAAAATAGGTC	Primer for pCFD4-1 construction with N20 reverse complementary sequence of sgRNA2
Step 3 in procedure B	pCFD4-2 forward primer	TCCGGGTGAACTTCGNNNNNNNNNNNNNNNNNNN GTTTTAGAGCTAGAAATAGCAAG	Primer for pCFD4-2 construction with N20 sequence of sgRNA3
Step 3 in procedure B	pCFD4-2 reverse primer	TTCTAGCTCTAAAACNNNNNNNNNNNNNNNNNNN CGACGTTAAATTGAAAATAGGTC	Primer for pCFD4-2 construction with N20 reverse complementary sequence of sgRNA4
Step 11 in procedure B	pCFD4-seq	GACACAGCGCGTACGTCCTTCG	Sequencing primer for sgRNAs in pCFD4
Step 13 in procedure B	pCFD4-2 forward primer	ACACTAGTATTTTCAACGTCCTCGATAGTATAG	Amplify sgRNA cassette from pCFD4-2
Step 13 in procedure B	pCFD4-2 reverse primer	TCAGATCTCAATTGATCGGCTAAATGGAACAACTC TCAGGCTCCAGG	Amplify sgRNA cassette from pCFD4-2
Step 18 in procedure B	4sgRNA forward primer	GCCTTCAACCCAGTCAGCTCTAGAAGGCCTAATTCGGTAC	Amplify four-sgRNA cassette
Step 18 in procedure B	4sgRNA reverse primer	CTGTTGCCGAGCACAATTGTC	Amplify four-sgRNA cassette
Step 22 in procedure B	pLib6.4-forward primer	AAGCTTGAGGGCAGAGGAAGC	Amplify T2A-Puro^r-SV40 polyA cassette from pLib6.4 vector
Step 22 in procedure B	pLib6.4-reverse primer	GAGCTGACTGGGTTGAAGGCTC	Amplify T2A-Puro^r-SV40 polyA cassette
Step 23 in procedure B	UHA forward primer	GCCGCCAGTGTGCTGGAATTNNNNNNNNNNNNNNN	PCR amplify upstream homology arm

Table continued

Table 1 (continued)

Step	Primer	Sequence (5′–3′)	Purpose
Step 23 in procedure B	UHA reverse primer	AGAAGGCTTCCTCTGCCCTCNNNNNNNNNNNNNNN	PCR amplify upstream homology arm
Step 23 in procedure B	DHA forward primer	ACAATTGTGCTCGGCAACAGNNNNNNNNNNNNNNN	PCR amplify downstream homology arm
Step 23 in procedure B	DHA reverse primer	TGATGGATATCTGCAGAATTNNNNNNNNNNNNNNN	PCR amplify downstream homology arm
Step 29 in procedure B	Tnks-first round PCR-forward	CGAAGGCAAGACCTTCTCGCTC	First round PCR of Tnks loci
Step 29 in procedure B	Tnks-first round PCR-reverse	GAGTAGGGCCAGCAGGCGATC	First round PCR of Tnks loci
Step 29 in procedure B	Apc-first round PCR-forward	CCTATCTGCTTTCCTTGCAG	First round PCR of Apc loci
Step 29 in procedure B	Apc-first round PCR-reverse	CGGAATGGGGTTCTCGTCCAG	First round PCR of Apc loci
Step 32 in procedure B	Tnks-second round PCR-forward	TCCCATCGACCAATACATGAC	Second round PCR of Tnks loci
Step 32 in procedure B	Tnks-second round PCR-reverse	TTCAAGCAGCTCATCCTTTCG	Second round PCR of Tnks loci
Step 32 in procedure B	Apc-second round PCR-forward	GCATTGACTTGCACCACTGAG	Second round PCR of Apc loci
Step 32 in procedure B	Apc-second round PCR-reverse	TCCTTCTTGGTATAGCCACTG	Second round PCR of Apc loci
Step 35 in procedure B	Tnks-HDR-forward	CGCTTAATGCGTATGCATTC	Test insertion cassette in Tnks loci
Step 35 in procedure B	Tnks-HDR-reverse	CGGCGCACAGTGTCCAGATC	Test insertion cassette in Tnks loci
Step 35 in procedure B	Apc-HDR-forward	CGCTTAATGCGTATGCATTC	Test insertion cassette in Apc loci
Step 35 in procedure B	Apc-HDR-reverse	CACCAGGGCCTCCATGAACTGC	Test insertion cassette in Apc loci

Q5 High-Fidelity DNA Polymerase (NEB) and KAPA HiFi DNA Polymerase (Sigma-Aldrich), can also be used.

- dNTP Mixture, 2.5 mM each (Clontech, cat. no. 4030)
- 1 kb Plus DNA ladder (Life Technologies, cat. no. 10787018)
- DNA Gel Loading Dye (6×) (Thermo Fisher Scientific, cat. no. R0611)
- Gibson Assembly Master Mix (NEB, cat. no. E2611)
- One Shot TOP10 Competent cells (Invitrogen, cat. no. C4040-10)
- QIAprep Spin Miniprep Kit (Qiagen, cat. no. 27106)
- Quick-DNA Microprep Kit (Zymo Research, cat. no. D3021)
- QuickExtract DNA Extraction Solution (Lucigen, cat. no. QE09050)
- GeneMate LE Agarose (BioExpress, cat. no. E-3120-500)
- GelRed Nucleic Acid Gel Stain (Biotium, cat. no. 41003) **! CAUTION** Nucleic acid dye may cause DNA mutation: wear gloves and dispose of according to laboratory safety regulations.

Cell culture

- Schneider's *Drosophila* Medium (Gibco, cat. no. 21720001)
- Fetal bovine serum (Gibco, cat. no. 16000044)
- Penicillin–streptomycin (Gibco, cat. no. 15140148)
- PBS, pH 7.4 (Gibco, cat. no. 10010049)
- Effectene Transfection Reagent (Qiagen, cat. no. 301427)
- Puromycin (Calbiochem, cat. no. 540411)
- Geneticin (G418) (Invitrogen, cat. no. 11811-031)
- 6-well tissue culture plates (Corning, cat. no. 3516)
- 96-well tissue culture plates (Corning, cat. no. 3610)

Biological materials

• S2R+-MT::Cas9 cell line (DGRC, cat. no. 268, RRID:CVCL_UD30)

Equipment

• Thermal cycler (any brand)
• 25 °C cell culture incubator (any brand)
• Microscope (any brand)
• Centrifuge (any brand)
• Fluorescence-activated cell sorter (BD FACSAria)

Reagent setup

Complete medium

Schneider medium supplemented with 10% (vol/vol) heat-inactivated FBS and 1× penicillin–streptomycin. Prepared medium can be stored at 4 °C for up to 3 months.

Puromycin selection medium

Schneider medium supplemented with 5 μg/ml puromycin, 10% (vol/vol) heat-inactivated FBS and 1× penicillin–streptomycin. Prepared medium can be stored at 4 °C for up to 3 months.

Neomycin selection medium

Schneider medium supplemented with 500 μg/ml G418, 10% (vol/vol) heat-inactivated FBS and 1× penicillin–streptomycin. Prepared medium can be stored at 4 °C for up to 3 months.

Conditioned medium

Grow S2R+ cells in complete medium in a T75 flask to 100% confluency in a 25 °C cell culture incubator, then split S2R+ cells 1:5 to new T75 flasks and culture cells for 3–5 d to reach 100% confluence. Collect the medium, centrifuge at 1,000*g* at room temperature (20 to 25 °C) for 5 min to remove the cell pellet and filter the conditioned medium with a 0.22 μm filter. The conditioned medium can be stored at 4 °C for up to 1 month.

Procedure

Procedure A: two rounds of homology-directed insertion

Design sgRNAs for target gene ● Timing 1 d

1 Identify sgRNAs using the DRSC online design tool (www.flyrnai.org/crispr/). Input the target gene ID or symbol, select target coding sequence (CDS), 3 mismatches and submit.
2 Select two sgRNAs with high efficiency scores.
3 Design and order the forward and reverse oligos for each sgRNA with no modifications, standard desalting, 25 nmol (Table 1).

▲ **CRITICAL STEP** The G immediately before the sgRNA is the first base that is transcribed, resulting in high transcription efficiency. It is not necessary to reduce the protospacer to N19 when it starts with G. The forward and reverse oligo anneals result in the formation of double-stranded DNA with 5′ overhang GTCG at one end and 5′ overhang AAAC at the other end; these are complementary to overhangs generated by Bbs1 digestion of sgRNA vectors.

Construct sgRNA vector ● Timing 2 d

4 Digest pLib7.1 and pLib7.2 with Bbs1 in NEBuffer 2.1 at 37 °C for 4 h as follows:

Component	Amount (μl)	Final concentration
10× NEBuffer 2.1	5	1×
pLib7.1 or pLib7.2 vector (100 ng/μl)	10	20 ng/μl
Bbs1 (10 U/μl)	1	0.2 U/μl
Nuclease-free water	34	
Total	50	

▲ **CRITICAL STEP** Additives in the restriction enzyme storage buffer such as glycerol can affect restriction digestion reaction. Restriction enzyme volume should not exceed 10% of the final volume.

5 Purify the digested vectors using the QIAquick PCR purification kit, according to the manufacturer's instructions.
▲ **CRITICAL STEP** After digestion of pLib7.1 and pLib7.2 with Bbs1, two fragments are generated, a 20 bp fragment and a 7.6 kb fragment. The 20 bp fragment is removed during the DNA purification process.

6 Resuspend sgRNA oligos (from Step 3) at 100 μM. For each sgRNA, mix the same amount of the forward oligo and corresponding reverse oligo in PCR tubes, as follows:

Component	Amount (μl)	Final concentration
10× NEB buffer2	1	1×
Forward oligo (100 μM)	1	10 μM
Reverse oligo (100 μM)	1	10 μM
Nuclease-free water	7	
Total	10	

▲ **CRITICAL STEP** Phosphorylation of the sgRNA oligos is not necessary because the digested vector provides phosphate groups that can form phosphodiester bonds with the hydroxyl groups in the annealed oligos.

7 Anneal forward and reverse oligos for each sgRNA using the following parameters:

Step	Condition
1	95 °C for 3 min
2	Ramp from 95 °C to 25°C, −0.1 °C/cycle, 5 s/cycle
3	Incubate at 4 °C

▲ **CRITICAL STEP** After annealing, the forward primer and reverse primer will form sticky ends that are complementary to the sticky ends of the digested vector.

8 For the two sgRNAs, one is ligated to digested pLib7.1, and the other is ligated to digested pLib7.2. Set up a ligation reaction for each sgRNA as follows, incubating the ligation reaction at 16 °C for 1 h.

Component	Amount (μl)	Final concentration
10× T4 DNA ligase buffer	1	1×
Digested pLib7.1 or pLib7.2 (20 ng/μl) from Step 5	1	2 ng/μl
Annealed oligo product from Step 7	1	
T4 DNA ligase (400 U/μl)	0.5	20 U/μl
Nuclease-free water	6.5	
Total	10	

9 Transform One Shot TOP10 competent cells with ligation product. Briefly, add 5 μl ligation product from Step 8 into 50 μl ice-cold competent cells, incubate on ice for 15 min, heat shock at 42 °C for 45 s and put back on ice for 2 min. Spread the bacteria on an LB plate containing 100 ng/μl ampicillin. Incubate the LB plates at 37 °C overnight.
▲ **CRITICAL STEP** Transform competent cells with digested pLib7.1 or pLib7.2 without ligation to determine digestion efficiency. Digested vectors without ligation normally produce very few colonies.

10 Pick four colonies for each sgRNA vector into LB medium containing 100 ng/μl ampicillin and incubate the bacterial culture at 37 °C overnight in a shaking incubator. Extract plasmid DNA using QIAprep Spin Miniprep Kit according to the manufacturer's instructions. Verify sgRNA vectors by Sanger sequencing with the primer pLib7-seq (Table 1).
■ **PAUSE POINT** Extracted plasmid DNA can be stored at −20°C for years.

PCR amplify T2A-Puro^r-SV40 polyA-U6:3-sgRNA cassette from pLib7.1 vector and T2A-Neo^r-SV40 polyA-U6:3-sgRNA cassette from pLib7.2 vector ● Timing 4 h

11 Set up the following PCR reaction to amplify T2A-Puro^r-SV40 polyA-U6:3-sgRNA cassette from pLib7.1 vector and T2A-Neo^r-SV40 polyA-U6:3-sgRNA cassette from pLib7.2 vector:

Component	Amount (μl)	Final concentration
5× Phusion GC Buffer/5× Phusion HF Buffer	10	1×
Forward primer (10 μM, see Table 1)	1	0.2 μM
Reverse primer (10 μM, see Table 1)	1	0.2 μM
dNTP (2.5 mM each)	2	100 μM each
pLib7.1 or pLib7.2 vector (5 ng/μl) from Step 10	1	100 pg/μl
Phusion High-Fidelity DNA polymerase (2 U/μl)	0.5	0.02 U/μl
Water	34.5	
Total	50	

▲ **CRITICAL STEP** To minimize PCR error, it is highly recommended to use high-fidelity polymerase. As Puro^r has high GC content, use Phusion GC Buffer for PCR. The forward primer binds to the T2A sequence; the reverse primer binds to the sequence following the sgRNA sequence. These primers can amplify T2A-Puro^r-SV40 polyA-U6:3-sgRNA cassette from pLib7.1 vector and T2A-Neo^r-SV40 polyA-U6:3-sgRNA cassette from pLib7.2 vector.

12 Perform PCR using the following conditions:

Cycle number	Denature	Anneal	Extend
1	98 °C, 30 s		
2–31	98 °C, 8 s	59 °C, 15 s	72 °C, 1 min
32			72 °C, 10 min

13 After PCR is complete, add 10 μl DNA Gel Loading Dye (6×) to PCR product. Load 60 μl PCR product into 1.5% (wt/vol) agarose gel in TBE buffer with GelRed Nucleic Acid Gel Stain. Run the gel in TBE buffer at 15 V/cm for 30 min. A successful PCR reaction will yield a ~1.5 kb product.

14 Purify PCR products using the QIAquick gel extraction kit, according to the manufacturer's instructions.

▲ **CRITICAL STEP** When using plasmid as template, it is crucial to purify PCR product by gel extraction rather than DNA purification as the template plasmid would create a lot of background colonies.

■ **PAUSE POINT** Gel-purified PCR product can be stored at −20 °C for months.

PCR amplify homology arms from *Drosophila* genomic DNA ● Timing 4 h

15 Extract *Drosophila* genomic DNA from S2R+ cells using a Quick-DNA Microprep Kit according to the manufacturer's instructions.

16 Set up two PCR reactions as follows to PCR amplify the upstream and downstream homology arms, respectively, from *Drosophila* genomic DNA:

Component	Amount (μl)	Final concentration
5× Phusion HF Buffer	10	1×
Forward primer (10 μM, see Table 1)	1	0.2 μM
Reverse primer (10 μM, see Table 1)	1	0.2 μM
dNTP (2.5 mM each)	2	100 μM each
Genomic DNA from Step 15 (100 ng/μl)	1	2 ng/μl
Phusion High-Fidelity DNA polymerase (2 U/μl)	0.5	0.02 U/μl
Water	34.5	
Total	50	

17 Perform PCR using the following conditions:

Cycle number	Denature	Anneal	Extend
1	98 °C, 30 s		
2–31	98 °C, 8 s	59 °C, 15 s	72 °C, 1 min
32			72 °C, 10 min

18 After PCR is complete, add 10 µl DNA Gel Loading Dye (6×) to PCR product. Load 60 µl PCR product into 1.5% (wt/vol) agarose gel in TBE buffer with GelRed Nucleic Acid Gel Stain. Run the gel in TBE buffer at 15 V/cm for 30 min. A successful PCR reaction will yield a ~1 kb product.

19 Purify PCR products using the QIAquick gel extraction kit, according to the manufacturer's instructions.

PAUSE POINT Gel-purified PCR product can be stored at −20°C for months.

Ligate homology arms and T2A-Puro^r-SV40 polyA-U6:3-sgRNA or T2A-Neo^r-SV40 polyA-U6:3-sgRNA cassette to pCR-Blunt vector ● Timing 1 d

20 Set up the following reaction:

Component	Amount (µl)	Final concentration
2× Gibson Assembly master mix	10	1×
pCR-Blunt vector (25 ng/µl)	4	5 ng/µl
T2A-Puro^r or Neo^r-SV40 polyA-U6:3-sgRNA cassette (100 ng/µl, from Step 14)	1	5 ng/µl
Upstream homology arm PCR product (100 ng/µl, from Step 19)	1	5 ng/µl
Downstream homology arm PCR product (100 ng/µl, from Step 19)	1	5 ng/µl
Water	3	
Total	20	

21 Incubate the reaction for 1 h at 50 °C, then place on ice.

22 Transform One Shot TOP10 competent cells with the ligation product. Briefly, add 5 µl ligation product from Step 21 into 50 µl ice-cold competent cells, incubate on ice for 15 min, heat shock at 42 °C for 45 s and put back on ice for 2 min.

23 Allow cells to recover in fresh LB medium without antibiotics at 37 °C for 1 h.

▲ CRITICAL STEP For multiple fragment ligation, recovery after transformation is recommended because this increases colony number.

24 Spread the transformed cells on an LB plate containing 50 ng/µl kanamycin. Incubate the plate at 37 °C overnight.

25 Pick eight bacterial colonies into LB medium containing 50 ng/µl kanamycin. Incubate the bacterial culture at 37 °C overnight in a shaking incubator. Extract plasmid DNA using a QIAprep Spin Miniprep Kit according to the manufacturer's instructions. Verify correct insertions by Sanger sequencing with M13 forward primer and M13 reverse primer (see Table 1).

PAUSE POINT Plasmid can be stored at −20 °C for years.

? TROUBLESHOOTING

First round of homology-directed insertion ● Timing 20 d

26 Seed 3 × 10^6 S2R+-MT::Cas9 cells in one well of a 6-well plate in complete medium.

27 Transfect the cells with the donor vector containing the Puro^r selectable marker (from Step 25) using the Effectene transfection reagent kit, according to the manufacturer's instructions. Briefly, add 0.4 µg plasmid into 100 µl Buffer EC containing 3.2 µl Enhancer and vortex to mix. After incubating at room temperature (20–25 °C) for 3 min, add 10 µl Effectene reagent and vortex to mix. After incubating at room temperature for 15 min, add the transfection complex drop-wise onto the cells. Incubate the cells at 25 °C in cell culture incubator.

▲ CRITICAL STEP Effectene reagent enables transfection in the presence of serum, and therefore it is not necessary to use serum-free medium for transfection.

28 Four days after transfection, gently aspirate culture medium with a pipette and add puromycin selection medium. Incubate for 15 d to select for successfully transformed puromycin-resistant colonies.
▲ CRITICAL STEP Do not allow cells to become confluent during the puromycin selection process because this decreases their sensitivity to puromycin. As S2R+ cells are not firmly attached, gently aspirate the culture medium to avoid cells floating when changing medium. Change puromycin selection medium every 5 d. Ideally, cell clusters are typically observed after 15 d of selection.
? TROUBLESHOOTING

Second round of homology-directed insertion ● Timing 20 d

29 After 15 d of puromycin selection, resuspend the resistant cells by pipetting and determine the cell concentration with a hemocytometer. Seed 5×10^5 cells in one well of a 6-well plate in puromycin selection medium, and then culture the cells for 3–4 d to 80–90% confluency in a 25 °C cell culture incubator.

30 Resuspend cells from Step 29 by pipetting and determine the cell concentration with a hemocytometer. Seed 3×10^6 cells in one well of a 6-well plate with complete medium.

31 Transfect the cells with the donor vector containing the Neo^r selectable marker (from Step 25) as described in Step 27.

32 Four days after transfection, gently aspirate the culture medium with a pipette and add neomycin selection medium. Incubate for 15 d to select for successfully transformed neomycin-resistant colonies.
▲ CRITICAL STEP Do not allow cells to become confluent during the neomycin selection process because this decreases sensitivity to G418. As S2R+ cells are not firmly attached, gently aspirate culture medium to avoid cells floating when changing medium. Change neomycin selection medium every 5 d. Ideally, cell clusters are typically observed after 15 d of selection.
? TROUBLESHOOTING

Single-cell cloning of the selected population ● Timing 20 d

33 Add 150 µl of conditioned medium (see 'Reagent setup') into each well of five 96-well plates.

34 Resuspend the resistant cells from Step 32 by pipetting and determine the cell concentration with a hemocytometer. Make a cell suspension at 1×10^5–10^6/ml with complete medium, and filter 0.5 ml of this suspension with a 40 µm cell strainer.

35 Sort a single cell into each well of the five 96-well plates from Step 33 by flow cytometry. Set a forward scatter (FSC) versus a side scatter (SSC) gate to select the live cell population. Then set a forward scatter height (FSC-H) versus a forward scatter area (FSC-A) gate to select the single-cell population (Supplementary Fig. 1).
▲ CRITICAL STEP To guarantee cell viability, use a 100 µm nozzle for low-pressure sorting. Sort 100 cells into the A1 well of each 96-well plate to facilitate focusing for microscopy observations.

36 Seal the plates with Parafilm and incubate in a humidity box with moist paper towels in a 25°C incubator for 2–3 weeks.
▲ CRITICAL STEP After 2–3 weeks, ~20–50% wells will contain cell colonies with hundreds of cells. It is not necessary to change the medium during single-cell cloning.
? TROUBLESHOOTING

Genotype analysis of cell clones ● Timing 7 d

37 Split each cell clone into corresponding wells of two 96-well plates with complete medium. Incubate the cells in a humidity box with moist paper towels in a 25 °C cell culture incubator. Grow the cells for 3–5 d to reach 100% confluency.
▲ CRITICAL STEP One plate is for cell culture. Once the knockout colonies have identified, they can be amplified from this plate. The other plate is for genomic DNA analysis as described in Step 38.

38 Aspirate culture medium from one 96-well plate from Step 37. Add 30 µl QuickExtract DNA Extraction Solution per well, pipette up and down for 15 s to lyse cells, transfer the solution into tubes and prepare DNA samples according to the manufacturer's instructions. Briefly, incubate at 65 °C for 6 min, then incubate at 98 °C for 2 min.
▲ CRITICAL STEP QuickExtract DNA Extraction Solution enables simple DNA preparation for subsequent PCR analysis.

Fig. 3 | Representative sequencing results of knockout loci in knockout cells after two rounds of homology-directed insertion (procedure A). Note that only indel alleles are amplified; cassette insertion alleles do not generate a PCR product. **a**, Sequences and frequency (% of total reads) of *Tnks* indel alleles identified by DNA sequencing in *Tnks* knockout cells. Black dashes, deleted bases. **b**, Protein sequence alignment of *CG8786* in *CG8786* knockout cells. Premature stop codons are marked with an asterisk (*).

39 Prepare PCR reaction as follows:

Component	Amount (μl)	Final concentration
5× Phusion HF Buffer	10	1×
Forward primer (10 μM, see Table 1)	1	0.2 μM
Reverse primer (10 μM, see Table 1)	1	0.2 μM
dNTP (2.5 mM each)	2	100 μM each
DNA solution from Step 38	3	
Phusion High-Fidelity DNA polymerase (2 U/μl)	0.5	0.02 U/μl
Water	32.5	
Total	50	

40 Perform PCR using the following conditions:

Cycle number	Denature	Anneal	Extend
1	98 °C, 30 s		
2–31	98 °C, 8 s	59 °C, 15 s	72 °C, 30 s
32			72 °C, 10 min

41 Purify PCR products with the QIAquick PCR purification kit according to the manufacturer's instructions.

42 Test frameshifts by DNA sequencing of PCR products (Fig. 3).

Procedure B: Homology-directed insertion with multiple sgRNAs

Design sgRNAs for target gene ● Timing 1 d

1 Identify sgRNAs using the DRSC online design tool (http://www.flyrnai.org/crispr/). Input target gene ID or symbol, select target CDS, 3 mismatches and submit.

2 Select four sgRNAs with high efficiency within a 300–500 bp region, and name them sgRNA1, sgRNA2, sgRNA3, sgRNA4.

3 Design and order sgRNA PCR primers for each pCFD4 vector with no modifications, standard desalting, 25 nmol (Fig. 2a, Table 1). Each pCFD4 vector will contain two different sgRNAs; name the sgRNA1- and sgRNA2-containing vectors pCFD4-1, and the sgRNA3- and sgRNA4-containing vectors pCFD4-2.

▲ CRITICAL STEP The G immediately before sgRNA is the first base that is transcribed, resulting in high transcription efficiency. It is not necessary to reduce the protospacer to N19 when it starts with G.

Construct a four-sgRNA cassette ● Timing 4 d

4 Repeat Steps 11–14 of procedure A using pCFD4 vector as a PCR template and primers from Step 3 of procedure B for pCFD4-1 and pCFD4-2, individually.

5 Digest pCFD4 vector with Bbs1 in NEBuffer 2.1 at 37 °C for 4 h:

Component	Amount (μl)	Final concentration
10× NEBuffer 2.1	5	1×
pCFD4 vector (100 ng/μl)	10	20 ng/μl
Bbs1 (10 U/μl)	1	0.2 U/μl
Nuclease-free water	34	
Total	50	

▲**CRITICAL STEP** Additives in the restriction enzyme storage buffer such as glycerol can affect restriction digestion reaction. Restriction enzyme volume should not exceed 10% of the final volume.

6 After vector digestion is complete, add 10 μl DNA Gel Loading Dye (6X) to product. Load 60 μl product into 1% (wt/vol) agarose gel in TBE buffer with GelRed Nucleic Acid Gel Stain. Run the gel in TBE buffer at 15 V/cm for 30–45 min. A successful digestion reaction should yield one 600 bp band and one 6.4 kb band.

7 Purify the 6.4 kb band using the QIAquick gel extraction kit, according to the manufacturer's instructions.

8 Set up Gibson Assembly reactions for each PCR product from Step 4 as follows:

Component	Amount (μl)	Final concentration
2× Gibson Assembly Master Mix	5	1×
Linearized pCFD4 from Step 7 (20 ng/μl)	4	8 ng/μl
PCR product from Step 4 (20 ng/μl)	1	2 ng/μl
Total	10	

9 Incubate at 50 °C for 1 h.

10 Transform One Shot TOP10 competent cells with ligation products. Briefly, add 5 μl ligation product from Step 9 into 50 μl ice-cold competent cells, incubate on ice for 15 min, heat shock at 42 °C for 45 s and put back on ice for 2 min. Spread the bacteria on an LB plate containing 100 ng/μl ampicillin and incubate at 37 °C overnight.

11 Pick 4–8 colonies for each sgRNA vector and verify sgRNA vectors by Sanger sequencing with primer pCFD4-seq (Table 1).

12 After sequencing, repeat Step 5 of Procedure B to digest pCFD4-1 with Spe1 and Bgl2 using the corresponding NEBuffer (Fig. 2b).

13 Meanwhile, PCR amplify sgRNA cassette from pCFD4-2. Spe1 and Bgl2 digestion sites are introduced into the PCR product by the primers.

Component	Amount (μl)	Final concentration
5× Phusion HF Buffer	10	1×
pCFD4-2 forward primer (10 μM, Table 1)	1	0.2 μM
pCFD4-2 reverse primer (10 μM, Table 1)	1	0.2 μM
dNTP (2.5 mM each)	2	100 μM each
pCFD4-2 from Step 11 (10 ng/μl)	1	0.2ng/μl
Phusion High-Fidelity DNA polymerase (2 U/μl)	0.5	0.02 U/μl
Water	34.5	
Total	50	

14 Perform PCRs using the following conditions:

Cycle number	Denature	Anneal	Extend
1	98 °C, 30 s		
2–31	98 °C, 8 s	59 °C, 15 s	72 °C, 1 min
32			72 °C, 10 min

15 Repeat Step 5 of procedure B to digest pCFD4-2 PCR product from Step 14 with Spe1 and Bgl2 using corresponding NEBuffer.

16 Ligate digested pCFD4-2 PCR product from Step 15 to linearized pCFD4-1 vector from Step 12 with T4 ligase. Set up a ligation reaction as follows, incubate the ligation reaction at 16 °C for 1 h:

Component	Amount (μl)	Final concentration
10× T4 DNA ligase buffer	2	1×
Linearized pCFD4-1 vector from Step 12 (20 ng/μl)	2	2 ng/μl
Digested pCFD4-2 PCR product from Step 15 (10 ng/μl)	1	0.5 ng/μl
T4 DNA ligase (400 U/μl)	1	20 U/μl
Nuclease-free water	6.5	
Total	20	

17 Transform One Shot TOP10 competent cells with ligation product. Briefly, add 5 μl ligation product from Step 16 into 50 μl ice-cold competent cells, incubate on ice for 15 min, heat shock at 42 °C for 45 s and put back on ice for 2 min. Spread the bacteria on an LB plate containing 100 ng/μl ampicillin. Incubate the LB plates at 37 °C overnight.

18 Pick eight bacteria colonies from LB plate to PCR tubes containing 5 μl water. Set up PCR reactions to amplify four-sgRNA cassettes from bacteria colonies as follows.

Component	Amount (μl)	Final concentration
5× Phusion HF Buffer	10	1×
4sgRNA forward primer (10 μM, Table 1)	1	0.2 μM
4sgRNA reverse primer (10 μM, Table 1)	1	0.2 μM
dNTP (2.5 mM each)	2	100 μM each
Bacteria colony	5	
Phusion High-Fidelity DNA polymerase (2 U/μl)	0.5	0.02 U/μl
Water	30.5	
Total	50	

19 Perform PCRs using the following conditions:

Cycle number	Denature	Anneal	Extend
1	98 °C, 30 s		
2–31	98 °C, 8 s	59 °C, 15 s	72 °C, 2 min
32			72 °C, 10 min

20 After PCR is complete, add 10 μl DNA Gel Loading Dye (6×) to PCR product. Load 60 μl PCR product into 1% (wt/vol) agarose gel in TBE buffer with GelRed Nucleic Acid Gel Stain. Run the gel in TBE buffer at 15 V/cm for 30–45 min. A successful PCR reaction will yield a ~3 kb product.

21 Gel purify the ~3 kb PCR product using the QIAquick gel extraction kit, according to the manufacturer's instructions.

PCR amplify T2A-Puro^r-SV40 polyA cassette from pLib6.4 vector ● Timing 4 h

22 Repeat Steps 11–14 of procedure A using pLib6.4 vector as PCR template, pLib6.4-forward primer and pLib6.4-reverse primer (Table 1).

PCR amplify homology arms from *Drosophila* genomic DNA ● Timing 4 h

23 Repeat Steps 16–19 of procedure A using the primers listed in Table 1.

Ligate homology arms, T2A-Puro^r -SV40 polyA cassette and four-sgRNA cassette to pCR-Blunt vector ● Timing 1 d

24 Set up the following reaction:

Component	Amount (μl)	Final concentration
2× Gibson Assembly master mix	10	1x
pCR-Blunt vector (25 ng/μl)	4	5 ng/μl
T2A cassette PCR product from Step 22 (100 ng/μl)	1	5 ng/μl
sgRNA cassette PCR product from Step 21(100 ng/μl)	1	5 ng/μl
Upstream homologous arm PCR product from Step 23 (100 ng/μl)	1	5 ng/μl
Downstream homologous arm PCR product from Step 23 (100 ng/μl)	1	5 ng/μl
Water	2	
Total	20	

▲ **CRITICAL STEP** In general, good cloning efficiency is achieved when using 50–200 ng of vector and inserts, respectively.

25 Repeat Steps 21–25 of procedure A

■ **PAUSE POINT** Plasmid can be stored at −20 °C for years.

? TROUBLESHOOTING

Homology-directed insertion ● Timing 20 d

26 Repeat Steps 26–28 of procedure A to transfect vector from Step 25 of procedure B into Cas9-expressing cells and select for successful transformants.

? TROUBLESHOOTING

Single-cell cloning of the selected population ● Timing 20 d

27 Repeat Steps 33–36 of procedure A.

? TROUBLESHOOTING

Genotype analysis of cell clones ● Timing 7 d

28 Repeat Steps 37 and 38 of procedure A.

29 Perform first round of nested PCR amplification as follows:

Component	Amount (μl)	Final concentration
5× Phusion HF Buffer	10	1×
First round PCR-forward primer (10 μM, Table 1)	1	0.2 μM
First round PCR-reverse primer (10 μM, Table 1)	1	0.2 μM
dNTP (2.5 mM each)	2	100 μM each
DNA solution from Step 28	3	
Phusion High-Fidelity DNA polymerase (2 U/μl)	0.5	0.02 U/μl
Water	32.5	
Total	50	

30 Perform PCR using the following conditions:

Cycle number	Denature	Anneal	Extend
1	98 °C, 30 s		
2–31	98 °C, 8 s	59 °C, 15 s	72 °C, 30 s
32			72 °C, 10 min

31 Purify PCR products with the QIAquick PCR purification kit according to the manufacturer's instructions, elute with 50 μl elution buffer.

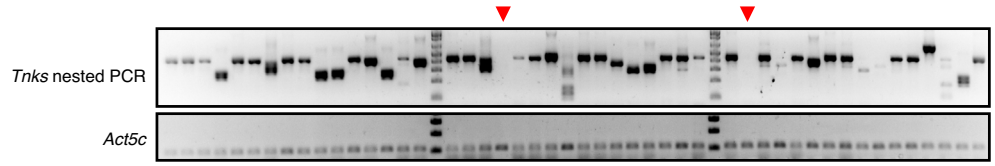

Fig. 4 | *Tnks* nested PCR results for single-cell clones using procedure B. Using procedure B for gene knockout by homology-directed insertion with multiple sgRNAs, we knocked out two genes, *Tnks* and *Apc*. For *Tnks* knockout experiment, we screened 96 colonies, three of which are homozygous for the insertion mutation. For *Apc* knockout experiment (not shown), we screened 96 colonies, two of which are homozygous for the insertion mutation. Some clones (red triangles) lost amplifiable *Tnks* target sites, suggesting that *Tnks* was knocked out in these clones by the insertion cassette. Some clones showed obvious band shifts, which might indicate deletion alleles generated by multiple sgRNAs. PCR products amplified from *Act5c* loci are used as a positive control for the PCR analysis. Please note that the efficiency of homozygous insertion mutation is less than the actual knockout efficiency, as clones with heterozygous insertions could also be mutant due to NHEJ events. The source data for this figure can be found in Supplementary Fig. 2.

32 Take 1 µl eluted PCR product for the second round of nested PCR amplification:

Component	Amount (µl)	Final concentration
5× Phusion HF Buffer	10	1×
Second round PCR-forward primer (10 µM, Table 1)	1	0.2 µM
Second round PCR-reverse primer (10 µM, Table 1)	1	0.2 µM
dNTP (2.5 mM each)	2	100 µM each
Elution from Step 31	1	
Phusion High-Fidelity DNA polymerase (2 U/µl)	0.5	0.02 U/µl
Water	34.5	
Total	50	

33 Perform PCR using the following conditions:

Cycle number	Denature	Anneal	Extend
1	98 °C, 30 s		
2–31	98 °C, 8 s	59 °C, 15 s	72 °C, 30 s
32			72 °C, 10 min

34 After the second round of nested PCR, add 10 µl DNA Gel Loading Dye (6×) to PCR product. Load 5 µl PCR product into 1.5% (wt/vol) agarose gel in TBE buffer with GelRed Nucleic Acid Gel Stain. Run the gel in TBE buffer at 15 V/cm for 30 min. Colonies with homozygous insertion mutations do not have non-insertional alleles, and thus would lose PCR products from non-insertional alleles (Fig. 4).

35 Repeat Steps 15–18 of procedure A with PCR primers HDR-forward and HDR-reverse (Table 1) to verify the insertion cassette in the knockout cell lines (Fig. 5a,b).

36 Verify the knockout cell lines by western blotting[29] if possible (Fig. 5c,d).

Troubleshooting

Troubleshooting advice can be found in Table 2.

Table 2 | Troubleshooting table

Step	Problem	Possible reason	Solution
25 of procedure A and 25 of procedure B	No bacterial colonies observed after transformation	Vector design mistake	Carefully check vector design, especially check overlapping sequences for Gibson Assembly
		Low ligation efficiency	Increase DNA amount for ligation
		Low transformation efficiency	Use competent cells with high transformation efficiency

Table continued

Table 2 (continued)

Step	Problem	Possible reason	Solution
	Too many background bacterial colonies observed after transformation	Plasmids used as PCR templates remained in the samples	Be sure to perform gel purification rather than DNA purification after PCRs using plasmid as PCR templates
28 and 32 of procedure A and 26 of procedure B	No cell survival after selection	Vector design mistake	Check whether resistance gene is in frame with target gene
		PCR mutation	Check T2A-Puro^r coding region mutation
		Puromycin/G418 concentration too high	Decrease puromycin/G418 concentration
		Low sgRNA efficiency	Check sgRNA efficiency with T7EI assay
36 of procedure A and 27 of procedure B	No cell clones	Old conditioned medium	Prepare new conditioned medium according to our protocol

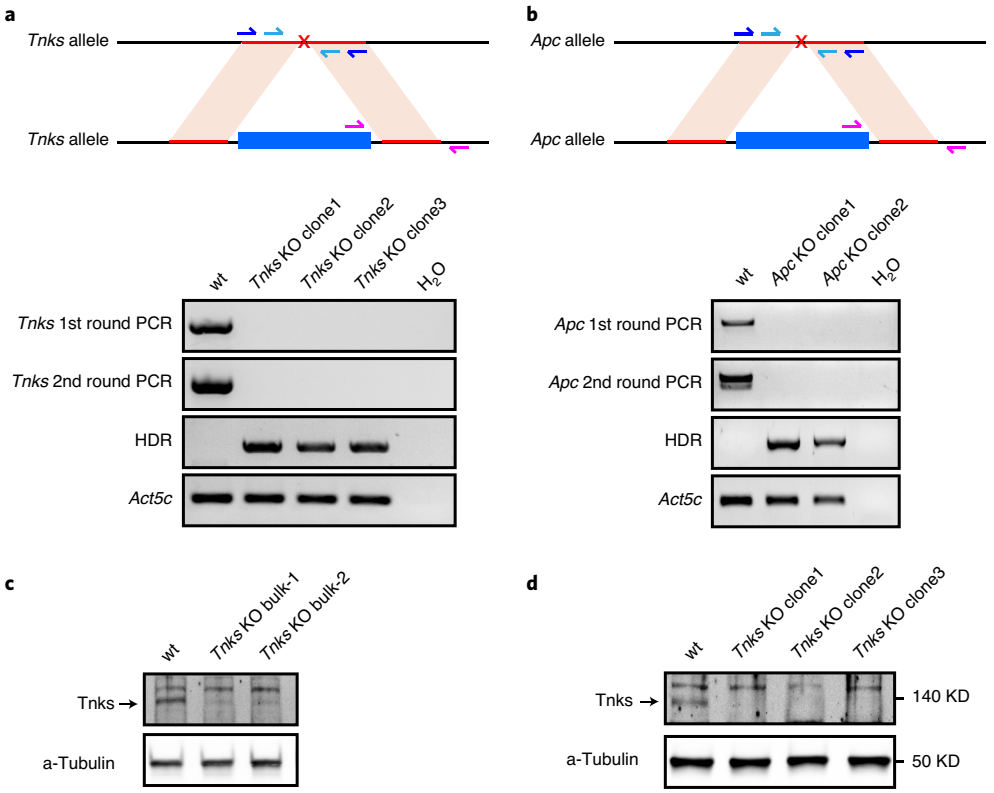

Fig. 5 | *Tnks* knockout cell lines and *Apc* knockout cell lines established using procedure B. a, Loss of non-insertion alleles of *Tnks* loci was tested by PCR using nested sets of primers shown in dark and light blue. Cassettes inserted into *Tnks* loci were tested by PCR using primers shown in pink. PCR products amplified from *Act5c* loci are used as a positive control for the PCR analysis. **b**, Loss of non-insertion alleles of *Apc* loci were tested by PCR using nested primers shown in dark and light blue. Cassettes inserted into *Apc* loci were tested by PCR using primers shown in pink. **c**, Tnks protein levels in the bulk population before single cell cloning. Results from two replicates, bulk-1 and bulk-2, are shown. **d**, Protein level of Tnks in *Tnks* knockout cell lines. The source data for this figure can be found in Supplementary Fig. 3.

Timing

Procedure A

Steps 1–3, design sgRNAs for target gene: 1 d

Steps 4–10, construct sgRNA vector: 2d

Steps 11–14, PCR amplify T2A-Puro^r-SV40 polyA-U6:3-sgRNA cassette from pLib7.1 vector and T2A-Neo^r-SV40 polyA-U6:3-sgRNA cassette from pLib7.2 vector: 4 h
Steps 15–19, PCR amplify homology arms from *Drosophila* genomic DNA: 4 h
Steps 20–25, ligate homology arms, T2A-Puro^r/Neo^r-SV40 polyA-U6:3-sgRNA cassette to pCR-Blunt vector: 1 d
Steps 26–28, first round of homology-directed insertion: 20 d
Steps 29–32, second round of homology-directed insertion: 20 d
Steps 33–36, single-cell cloning of the selected population: 20 d
Steps 37–42, genotype analysis of cell clones: 7 d
Procedure B
Steps 1–3, design sgRNAs for target gene: 1 d
Steps 4–21, construct a four-sgRNA cassette: 4 d
Step 22, PCR amplify T2A-Puro^r-SV40 polyA cassette from pLib6.4 vector: 4 h
Step 23, PCR amplify homology arms from *Drosophila* genomic DNA: 4 h
Steps 24 and 25, ligate homology arms, T2A-Puro^r-SV40 polyA cassette and 4 sgRNAs cassette to pCR-Blunt vector: 1d
Step 26, homology-directed insertion: 20 d
Step 27, single-cell cloning of the selected population: 20 d
Steps 28–36, genotype analysis of cell clones: 7 d

Anticipated results

Drosophila S2R+ cells are pseudotetraploid cells, with four or more copies of most of its genes, and therefore knocking out all alleles of a gene is a challenge. We thus designed an approach to generate gene knockout cell lines based on homology-directed insertion. In procedure A, we use two sequential rounds of homology-directed insertion. After two rounds of homology-directed insertion, non-insertion alleles remained in the polyploid *Drosophila* S2R+ cells (Fig. 3). In procedure B, we use a homology-directed insertion cassette containing four sgRNAs. The insertion cassette with multiple sgRNAs is able to replace all alleles. Using procedure B, we established *Tnks* and *Apc* knockout cell lines. For the *Tnks* knockout experiment, we screened 96 colonies, and found that three of them are homozygous for the insertion mutation (Fig. 4). For the *Apc* knockout experiment, we screened 96 colonies, and found that two of them are homozygous for the insertion mutation. Figure 5 shows that we observed the expected results by PCR and western blot after establishing the gene knockout cell lines. This method should be applicable to other polyploid cell lines such as cancer cell lines and should be useful to target any gene present in high copy number. Moreover, procedure B enables identification of homozygous insertion cassette knockout cell lines using a PCR-based strategy without any need for sequencing.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw data used to generate the figures are included as supplementary information. There are no restrictions on data availability.

References

1. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
2. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
3. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
4. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
5. Grav, L. M. et al. One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. J.* **10**, 1446–1456 (2015).
6. Wang, F. et al. A comparison of CRISPR/Cas9 and siRNA-mediated ALDH2 gene silencing in human cell lines. *Mol. Genet. Genomics* **293**, 769–783 (2018).
7. Li, C. et al. Generation of FOS gene knockout lines from a human embryonic stem cell line using CRISPR/Cas9. *Stem Cell Res.* **39**, 101479 (2019).

8. Gantz, V. M. & Bier, E. Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science* **348**, 442–444 (2015).
9. Housden, B. E. et al. Identification of potential drug targets for tuberous sclerosis complex by synthetic screens combining CRISPR-based knockouts with RNAi. *Sci. Signal* **8**, rs9 (2015).
10. Housden, B. E., Nicholson, H. E. & Perrimon, N. Synthetic lethality screens using RNAi in combination with CRISPR-based knockout in Drosophila cells. *Bio. Protoc.* https://doi.org/10.21769/BioProtoc.2119 (2017).
11. Housden, B. E. & Perrimon, N. Detection of indel mutations in Drosophila by high-resolution melt analysis (HRMA). *Cold Spring Harb. Protoc.* https://doi.org/10.1101/pdb.prot090795 (2016).
12. Boettcher, M. & McManus, M. T. Choosing the right tool for the job: RNAi, TALEN, or CRISPR. *Mol. Cell* **58**, 575–585 (2015).
13. Scharf, I. et al. Dynamics of CRISPR/Cas9-mediated genomic editing of the AXL locus in hepatocellular carcinoma cells. *Oncol. Lett.* **15**, 2441–2450 (2018).
14. Lee, H. et al. DNA copy number evolution in Drosophila cell lines. *Genome Biol.* **15**, R70 (2014).
15. Friedel, R. H. et al. Gene targeting using a promoterless gene trap vector ("targeted trapping") is an efficient method to mutate a large fraction of genes. *Proc. Natl Acad. Sci. USA* **102**, 13188–13193 (2005).
16. Ye, L. et al. Programmable DNA repair with CRISPRa/i enhanced homology-directed repair efficiency with a single Cas9. *Cell Discov.* **4**, 46 (2018).
17. Maruyama, T. et al. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538–542 (2015).
18. Chu, V. T. et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**, 543–548 (2015).
19. Yu, C. et al. Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* **16**, 142–147 (2015).
20. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).
21. Ma, Z. et al. PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components. *Nature* **568**, 259–263 (2019).
22. El-Brolosy, M. A. et al. Genetic compensation triggered by mutant mRNA degradation. *Nature* **568**, 193–197 (2019).
23. Hu, Y., Comjean, A., Perrimon, N. & Mohr, S. E. The Drosophila gene expression tool (DGET) for expression analyses. *BMC Bioinforma.* **18**, 98 (2017).
24. Cherbas, L. et al. The transcriptional diversity of 25 Drosophila cell lines. *Genome Res.* **21**, 301–314 (2011).
25. Sakurai, T., Watanabe, S., Kamiyoshi, A., Sato, M. & Shindo, T. A single blastocyst assay optimized for detecting CRISPR/Cas9 system-induced indel mutations in mice. *BMC Biotechnol.* **14**, 69 (2014).
26. Port, F., Chen, H. M., Lee, T. & Bullock, S. L. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proc. Natl Acad. Sci. USA* **111**, E2967–E2976 (2014).
27. Port, F. & Bullock, S. L. Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. *Nat. Methods* **13**, 852–854 (2016).
28. Viswanatha, R., Li, Z., Hu, Y. & Perrimon, N. Pooled genome-wide CRISPR screening for basal and context-specific fitness gene essentiality in Drosophila cells. *eLife* https://doi.org/10.7554/eLife.36333 (2018).
29. Caudron-Herger, M. et al. Identification, quantification and bioinformatic analysis of RNA-dependent proteins by RNase treatment and density gradient ultracentrifugation using R-DeeP. *Nat. Protoc.* **15**, 1338–1370 (2020).

Acknowledgements

We thank the Harvard Medical School Immunology Flow Cytometry Facility for cell sorting and Professor Y. Ahmed (Department of Molecular and Systems Biology, Dartmouth Geisel School of Medicine) for the anti-Tnks antibody. This study was supported by NIH NIGMS R01 GM067761, NIH NIGMS P41 GM132087 and NIH ORIP R24 OD019847. S.E.M. is additionally supported in part by the Dana Farber/Harvard Cancer Center, which is supported in part by NCI Cancer Center Support grant number NIH 5 P30 CA06516. N.P. is an investigator of the Howard Hughes Medical Institute.

Author contributions

B.X. and N.P. designed the experiments. B.X. developed the protocol with G.A., R.V., J.Z. and S.E.M. B.X. and G.A. performed the experiments. B.X., S.E.M. and N.P. wrote the protocol.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41596-020-0383-8.

Correspondence and requests for materials should be addressed to N.P.

Peer review information *Nature Protocols* thanks E. Bier, J-L. Liu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 17 December 2019; Accepted: 9 July 2020;
Published online: 21 September 2020

Related links
Key references using this protocol:
Housden, B. E. et al. *Sci. Signal.* **8**, rs9 (2015): https://doi.org/10.1126/scisignal.aab3729
Housden, B. E. et al. *Bio. Protoc.* **7**, e2119 (2017): https://doi.org/10.21769/BioProtoc.2119
Nicholson, H. E. et al. *Sci. Signal.* **12**, eaay0482 (2019): https://doi.org/10.1126/scisignal.aay0482

nature research

Corresponding author(s): Norbert Perrimon

Last updated by author(s): Mar 10, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	*Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.*
Data analysis	*Provide a description of all commercial, open source and custom code used to analyse the data in this study, specifying the version used OR state that no software was used.*

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated or analysed during this study are included in the supplementary information files.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	In this study, for gene knockout by homology-directed insertion with multiple sgRNAs, we knocked out two genes, Tnks and Apc. For Tnks knockout experiment, we screened 96 colonies, 3 of them are homozygous insertion mutation. For Apc knockout experiment, we screened 96 colonies, 2 of them are homozygous insertion mutation.
Data exclusions	No data were excluded from the analysis.
Replication	We used this protocol to knockout different genes.
Randomization	For single cell cloning, individual cells are sorted randomly into each wells by FACS.
Blinding	The investigators were not blinded during data collection. We chose the colonies with homozygous insertion mutation for further validation.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*
Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for*

Timing and spatial scale	*these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access and import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	anti-Tnks antibody (from Prof. Yashi Ahmed, Dept. of Molecular and Systems Biology, Dartmouth Geisel School of Medicine)
Validation	Yashi Ahmed's Lab generated Tnks antibody by immunizing guinea pigs with purified Ankyrin repeat region of Tnks. The antibody validation data is provided in this study: Wang, Z., et al. (2016). "The ADP-ribose polymerase Tankyrase regulates adult intestinal stem cell proliferation during homeostasis in Drosophila." Development 143(10): 1710-1720.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Drosophila S2R+ cells
Authentication	S2R+-MT::Cas9 cell line is available at Drosophila Genomics Resource Center.
Mycoplasma contamination	The cell lines were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	none

Palaeontology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	*For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*

Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*
Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	cell suspension for single cell cloning
Instrument	BD FACSAria
Software	BD FACSAria
Cell population abundance	sort live single cell
Gating strategy	gating live single cell

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s) — *Specify: functional, structural, diffusion, perfusion.*

Field strength — *Specify in Tesla*

Sequence & imaging parameters — *Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*

Area of acquisition — *State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software — *Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

Normalization — *If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*

Normalization template — *Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*

Noise and artifact removal — *Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring — *Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings — *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested — *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016) — *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

Correction — *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity — *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis — *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis — *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*