nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-020-0361-1

Check for updates

GOTI, a method to identify genome-wide off-target effects of genome editing in mouse embryos

Erwei Zuo[1,2,14], Yidi Sun[1,3,14], Wu Wei[4,5,6,14], Tanglong Yuan[2,14], Wenqin Ying[1], Hao Sun[7], Liyun Yuan[4], Lars M. Steinmetz[6,8,9]✉, Yixue Li[4,10,11,12,13]✉ and Hui Yang[1]✉

Genome editing holds great potential for correcting pathogenic mutations. We developed a method called GOTI (genome-wide off-target analysis by two-cell embryo injection) to detect off-target mutations by editing one blastomere of two-cell mouse embryos using either CRISPR-Cas9 or base editors. GOTI directly compares edited and non-edited cells without the interference of genetic background and thus could detect potential off-target variants with high sensitivity. Notably, the GOTI method was designed to detect potential off-target variants of any genome editing tools by the combination of experimental and computational approaches, which is critical for accurate evaluation of the safety of genome editing tools. Here we provide a detailed protocol for GOTI, including mice mating, two-cell embryo injection, embryonic day 14.5 embryo digestion, fluorescence-activated cell sorting, whole-genome sequencing and data analysis. To enhance the utility of GOTI, we also include a computational workflow called GOTI-seq (https://github.com/sydaileen/GOTI-seq) for the sequencing data analysis, which can generate the final genome-wide off-target variants from raw sequencing data directly. The protocol typically takes 20 d from the mice mating to sequencing and 7 d for sequencing data analysis.

Introduction

The recent development of genome editing tools holds great promise in diverse fields, such as animal disease modeling, gene therapy, drug development, genetically modified plants and biofuel technology[1]. In addition, gene editing technology has accelerated the study of the functional organization of the genome and the causal links between genetic variations and biological phenotypes[2–5]. A booming genetic editing technology within the life sciences field, CRISPR holds great hope for the treatment of genetic diseases[6,7]. Most CRISPR–Cas9-edited products usually contain small indels at the target site due to the non-homologous end joining in response to double-strand breaks (DSBs)[8,9]. Base editors, on the other hand, induce base pair substitutions using deaminases at the target loci without generating DSBs[10–13]. There are two classes of DNA base editors. Cytosine base editors (CBEs) convert C•G base pairs to A•G base pairs[10], and adenine base editors (ABEs) conversely convert A•G base pairs to G•C base pairs[12]. However, the issue of off-target mutations, which might cause genetic instability and dysfunction, has been a major concern in the application of both these methods[14–18]. Specifically, the potential for off-target effects remains a major barrier to the applications of genome editing for human gene therapy. Several techniques have been developed to detect

[1]Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Research Center for Brain Science and Brain-Inspired Intelligence, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China. [2]Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China. [3]CAS Key Laboratory of Systems Biology, CAS Center for Excellence in Molecular Cell Science, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai, China. [4]Bio-Med Big Data Center, Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences Shanghai, Shanghai, China. [5]Center for Biomedical Informatics, Shanghai Children's Hospital, Shanghai Jiao Tong University, Shanghai, China. [6]Stanford Genome Technology Center, Stanford University, Palo Alto, CA, USA. [7]University of California, San Diego, La Jolla, CA, USA. [8]Department of Genetics, School of Medicine, Stanford University, Stanford, CA, USA. [9]Genome Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany. [10]Department of Life Sciences, Shanghai Tech University, Shanghai, China. [11]School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China. [12]Collaborative Innovation Center for Genetics and Development, Fudan University, Shanghai, China. [13]Shanghai Center for Bioinformation Technology, Shanghai Academy of Science & Technology, Shanghai, China. [14]These authors contributed equally: Erwei Zuo, Yidi Sun, Wu Wei, Tanglong Yuan. ✉e-mail: lars.steinmetz@stanford.edu; yxli@sibs.ac.cn; huiyang@ion.ac.cn

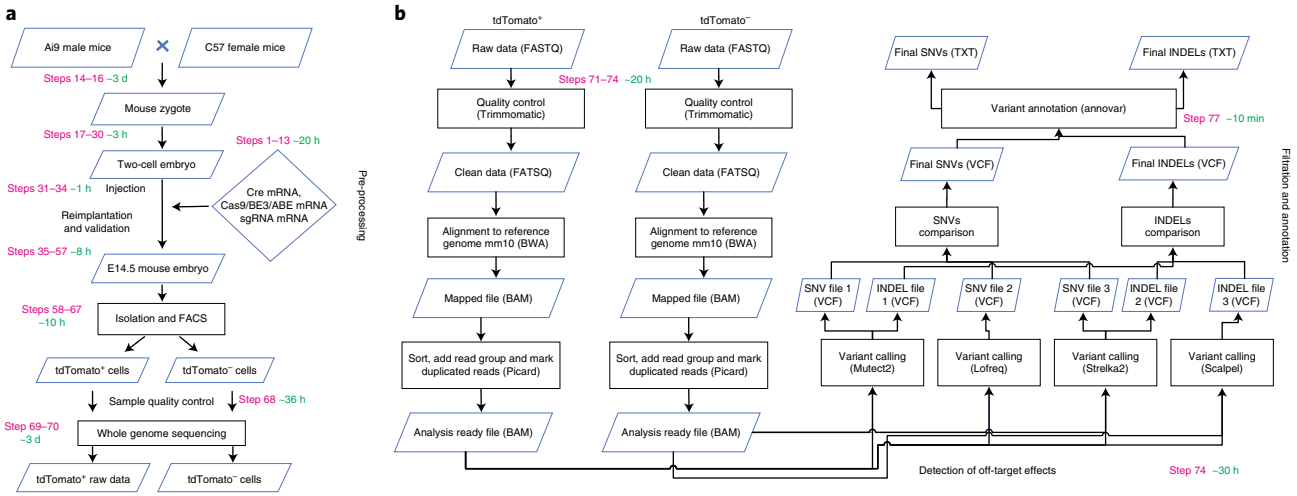

Fig. 1 | Schematics of the GOTI method. a, Overall workflow of GOTI. **b**, Flowchart of GOTI-seq data processing pipeline. File formats and software used for each step are denoted in parentheses.

genome-wide gene editing off-target activity in cells, including selective enrichment and identification of adapter-tagged DNA ends by sequencing (SITE-seq)[19], high-throughput genome-wide translocation sequencing (HTGTS)[20], genome-wide, unbiased identification of double-strand breaks (GUIDE-seq)[21] and circularization for in vitro reporting of cleavage effects by sequencing (CIRCLE-seq)[22]. However, these approaches can only detect off-targets from DSBs generated by genome editing and thus are not applicable to detect single-nucleotide variants (SNVs) in vivo. Given the rapidly increased application of base editors, it has become desirable to develop technologies for the comprehensive evaluation of off-target effects in an unbiased manner.

Recently, we developed a genome-wide off-target effects detection method called GOTI for the de novo identification of off-target mutations in mouse embryos[15]. This technique examines off-target effects in a cell population derived from a single gene edited blastomere, whereas previous studies used large pools of cells within which gene editing outcomes are variable, resulting in loss of signal for random off-target effects due to population averaging. In addition, as the edited and non-edited cells are from one single ancestor cell, GOTI can minimize any confounding effects of genetic background and somatic mutations. GOTI could thus be generalized to be used for genome editing tools that do not introduce DSBs, including triplex-forming oligonucleotides, base editors and potentially any other strategy that produces indels or SNVs.

In this article, we systemically describe the experimental procedures of GOTI and the bioinformatic pipeline GOTI-seq (https://github.com/sydaileen/GOTI-seq) in detail. We also emphasize the specific steps and small tricks that require great care to ensure the data quality generated. The GOTI method can be broadly applied to evaluate the specificity of genome editing tools in animal models.

Overview of GOTI

The overall workflow of GOTI is illustrated in Fig. 1. A mixture of Cre, gene editing messenger RNA (mRNA) (Cas9/BE3/ABE7.10) and single guide RNA (sgRNA) is injected into one blastomere of a two-cell mouse embryo, derived from Ai9 male mice mating with wild-type female mice[23]. The action of Cre, injected into only one of the two cells of the embryo, is expected to generate a chimeric embryo with half of the cells labeled with tdTomato (colored red) with the Cre-loxP system[23,24] so that the cells that received editing reagents cells can be distinguished from those that did not. When the chimeric embryo reaches embryonic day (E) 14.5, it is minced into small pieces and digested into a single-cell suspension (Fig. 2). The tdTomato$^+$ cells and tdTomato$^-$ cells are then separated by fluorescence-activated cell sorting (FACS). Next, the two populations of cells are independently processed for high-throughput whole-genome sequencing (WGS). Typically, the process from the two-cell injection to the completion of the final sequencing takes 3 weeks for an experienced operator (Fig. 1a).

The WGS reads are subsequently processed by a customized pipeline illustrated in Fig. 1b. Raw sequencing reads are first quality checked, and the adapters are further trimmed out. Qualified reads

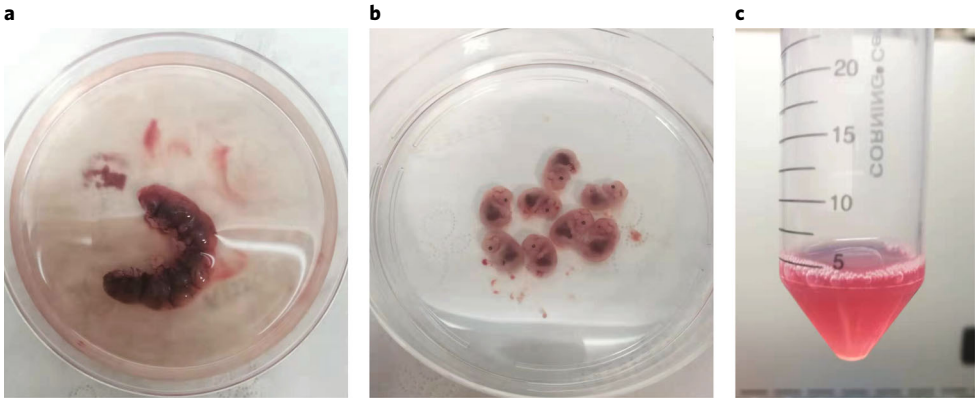

Fig. 2 | Isolation of embryonic cells. Pictures of key Steps 59–61. **a**, The embryos with the placenta. **b**, Dissect mouse embryos and wash the embryos with PBS twice briefly. **c**, Add 5 ml of pre-warmed 37 °C Trypsin-EDTA (0.05%) and transfer to a 50-ml centrifuge tube.

are then mapped to the reference genome. The mapped alignment (BAM format) files are sorted and duplicates marked before downstream analysis. The off-target SNVs and indels are identified by comparing the tdTomato$^+$ cells with tdTomato$^-$ cells using three variant calling algorithms as indicated (Fig. 1b). Several previous studies reported the discordance among variant calling methods due to different mathematical models and algorithms[25–28]. Mutect2 applies a Bayesian classifier to detect somatic mutations with even low allele fractions[29]. Strelka2 detects both SNVs and indels by a mixture-model-based estimation with high speed[30]. Lofreq is an ultra-sensitive variant caller to call somatic variants[31]. Scalpel performs well for the discovery of insertions and deletions[32]. Different algorithms apply different models and have their own preferences to variants, so, to reduce false positives, we considered the overlap of three algorithms of SNVs or indels as the true variants and annotated these for functional analysis.

Applications of GOTI

We have demonstrated that GOTI is applicable to detect the off-target effects of various genome editing tools, including CRISPR–Cas9, BE3 and ABE7.10. We have found numerous de novo SNVs induced by BE3, whereas CRISPR–Cas9 and ABE7.10 generated no unwanted mutations for the sgRNAs that were examined. Notably, the specificity of CRISPR–Cas9 depends heavily on the choice of sgRNA, so we cannot rule out the possibilities that some other sgRNAs might have significant off-target effects, even though we have examined several. GOTI is a valuable method that is expected to help improve the current gene editors and facilitate the generation of new genome editing tools with higher specificity. This method could be applied in the preclinical evaluation of genome editing reagents before their approval for clinical trials in gene therapy. GOTI depends on the Cre-loxP recombination system to distinguish edited cells from unedited cells, so it could theoretically be used in other transgenic animal models carrying the reporting system, such as rat[33,34] or pig[35,36].

Alternative methods and advantages of GOTI

Several genome-wide methods were previously developed for the detection of off-target effects. IDGV characterizes the genome-wide specificity by capturing DSBs in vivo[37]; HTGTS detects genome-wide translocations generated from DSBs in vitro[20]; and GUIDE-seq relies on the capture of double-stranded oligodeoxynucleotides integrated into the DSBs in vivo[21]. EndoV-seq[38] and Digenome-seq[39,40] are both in vitro assays to investigate the sgRNA-dependent off-target effects of genome editing tools. CIRCLE-seq[22] examines the in vitro DSBs of naked DNA for the prediction of off-target sites. In addition, DISCOVER-seq tracks the recruitment of MRE11 to DSBs in vivo[41]. These methods rely mostly on the generation of DSBs or nicks on the genome caused by the cleavage of Cas9, but base editors do not generate DSBs because they use catalytically deficient Cas9. Anderson et al.[42] and Lyer et al.[43] also applied deep sequencing for the detection of sgRNA-dependent off-target effects in genomic specific regions based on sequence similarity. The sgRNA-dependent off-target effects could be solved by using different sgRNAs. However, none of these methods was applicable to detect random off-target effects that are independent of sgRNA or Cas9.

Several previous studies also applied WGS to detect potential off-target effects of CRIPSR–Cas9 comparing edited and non-edited animals[41–47]. However, the true off-target variants could not be distinguished from the single-nucleotide polymorphisms (SNPs) by the genome comparison of two different individuals. One of the major advantages of GOTI is to directly compare edited and non-edited cells with identical genetic backgrounds. Therefore, potentially random off-target variants could be detected without bias even when they occur at low frequency.

The other main advantage of GOTI is its single-cell nature. The approach detects the off-target effects generated in a single ancestor cell and inherited by all descendant cells, which can be easily detected by sequencing, whereas previous work has been based on large pools of cells in which random off-target activity would be lost in the population average.

Limitations of GOTI

The embryo has to develop to E14.5 when the whole embryo could be readily digested to obtain enough single cells, so the duration of the whole experiment takes up to 1 month.

Also, GOTI relies on mice as a model system and requires regulatory approval with respect to animal welfare. However, GOTI can be theoretically conducted in other animal models with the Cre-loxP recombination system integrated. Another limitation is that GOTI is specific to the species in which it is performed. There are major ethical and legal considerations that prevent its use to assess off-target effects in the context of the human genome. In addition, the accomplishment of some procedures, such as two-cell embryo injection, requires a microinjection apparatus, professional training and technical skill. GOTI is much more expensive than other methods based on the enrichment of different off-target sites, which could be used to detect sgRNA-dependent off-target effects.

Experimental design

The generation of GOTI reagents (Steps 1−13)

GOTI reagents consist of sgRNA, gene editor (Cas9 or base editor) mRNA and Cre mRNA (Box 1). These components are generally transcribed in vitro from a T7 bacteriophage promoter (Supplementary Fig. 1). For sgRNA, a DNA template that contains T7 promotor, the designed sequence of sgRNA and the sgRNA scaffold is generated (as described in Steps 11−13) by performing a PCR reaction with the scaffold template (px330) and PCR primers (Table 1). The forward primer must contain a T7 promoter sequence (20 base pairs (bp)), an sgRNA target sequence (20 bp) and a scaffold template-specific sequence (19 bp). The T7 promoter sequence and the scaffold template-specific sequence are fixed. The reverse primer is a short complementary sequence targeting only the end of the guide RNA (gRNA) scaffold in the reverse direction. Then, an sgRNA containing the sgRNA target sequence is created by in vitro transcription of the DNA template with the MEGA Shortscript T7 Kit. For gene editor mRNA, a DNA template that contains the T7 promotor, gene editor coding sequence and polyA sequence is generated (as described in Steps 1–8 for Cas9 or Box 1 for BE mRNA) by performing a PCR reaction with plasmid template and PCR primers (Table 1). The forward primer includes the T7 promoter and 20–30 bp upstream sequence of start codon (ATG). The reverse primer is located downstream of the polyA sequence. The mRNA encoding the gene editor is created by in vitro transcription of the DNA template with the mMESSAGE mMACHINE T7 ULTRA Kit. For Cre mRNA, a DNA template that contains the T7 promotor, Cre coding sequence and polyA sequence is generated (as described in Steps 9 and 10) by performing a PCR reaction with pCAG-Cre plasmid template and PCR primers (Table 1). The forward primer includes the T7 promoter and a sequence upstream of the start codon (ATG). Reverse primers are located downstream of the polyA sequence. The mRNA encoding Cre is created by in vitro transcription of the DNA template with the mMESSAGE mMACHINE T7 ULTRA Kit.

Validation of on-target editing efficiency in blastocysts (Steps 44−57)

To evaluate the efficiency of on-target editing in the GOTI system (to ensure that the injected embryos were edited by genome editing tools, as, if not, downstream experiments would not be performed) in addition to transferring 80% of the injected embryos to the pseudo-pregnant mother, we culture the remaining 20% to the blastocyst stage in vitro (Step 34). The single blastocyst is then lysed, and PCR amplification products of the targeting sites are ligated into T-Vectors. The ligation mixture is transformed into *Escherichia coli* DH5α and cloned into *E. coli* (TA clone). Thirty TA clones are picked for Sanger sequencing to estimate the on-target editing efficiency. If at least six TA clones (20%) show on-target editing, further in vivo experiments (Steps 35−43) are performed.

Box 1 | Generation of editor mRNA (Cas9 or BE) for vectors with a T7 promoter

For vectors including T7 promoters ~20 bp upstream of coding sequence of gene editing protein, plasmid is linearized by restriction enzyme cutting. Restriction site is selected downstream of polyA. Here we provide a brief protocol for linearization of plasmid pCMV-BE3 and in vitro transcription of BE3 mRNA as an example.

1 Purify plasmid pCMV-BE3, using the TIANprep Rapid Mini Plasmid Kit according to the manufacturer's instructions.

2 Prepare the following reaction mixture at room temperature:

Component	Volume	Final concentration
10× FastDigest Green buffer	5 μl	1×
Plasmid pCMV-BE3 (0.5 μg/μl)	10 μl	0.1 μg/μl
FastDigest BbsI	5 μl	5 μl/50 μl
FastDigest PvuI	5 μl	5 μl/50 μl
FastDigest NotI	5 μl	5 μl/50 μl
ddH_2O	20 μl	—
Total	50 μl	—

3 Mix gently and spin down for a few seconds.

4 Incubate at 37 °C in water thermostat for 30 min.

5 Prepare 1× TAE by diluting appropriate amount of 50× TAE with deionized water.

6 Mix 0.5 g of agarose powder with 50 ml of 1× xTAE in a 250-ml conical flask.

7 Microwave for 1–3 min until the agarose is completely dissolved. If not, boil the solution again.

8 Cool down the agarose solution to ~60 °C.

9 Add 5 μl of nucleic acid dye to the agarose solution and mix gently.

10 Pour the agarose into a gel tray at room temperature and rest for 30 min, until it has completely solidified.

11 Remove gel dams and place agarose gel into the gel box. Add 1× TAE into the gel box until the gel is covered.

12 Load 5 μl of DNA Marker into the first lane of the gel.

13 Load 50 μl of digested plasmid DNA into the additional well of the gel.

14 Run the gel at 110 V for 30 min. Three digested DNA fragments are supposed to have 5,412, 1,170 and 1,410 bp in length, respectively, as shown in Supplementary Fig. 3.

! CAUTION Examine the linearized template DNA on a gel to confirm that cleavage is complete.

15 Excise the 5,412-bp DNA fragment band under long-wavelength ultraviolet light.

16 Purify DNA fragment from agarose gels using the Universal DNA Purification Kit according to the manufacturer's instructions.

17 Elute the DNA fragment with 40 μl of nuclease-free water.

18 Use 1 μg of purified DNA fragment as the template for in vitro transcription of BE3 mRNA using the mMESSAGE mMACHINE T7 Kit according to the protocol.

19 After purification, dilute the BE3 mRNA to 500 ng/μl with nuclease-free water and check its quality on a 1% (wt/vol) agarose gel in TAE buffer. The in vitro transcribed BE mRNA should have a band at ~1,000 bp (Supplementary Fig. 3). Failed in vitro transcription products show no clear bands at these positions as also shown in Supplementary Fig. 3.

20 Dispense 1 μl of the purified mRNA into 0.2-ml PCR tubes.

■ PAUSE POINT The samples could be stored at −80 °C for up to 1 year.

FACS gating strategy for isolation of embryonic cells (Steps 58–67)

We describe a step-by-step protocol to digest embryos and isolate embryonic cells by flow cytometry. The procedure is simple, and a cell suspension can be prepared from a single mouse embryo using enzymatic digestion and mechanical disaggregation in less than 35 min. To maximize cell viability, FACS must be performed as soon as possible after preparation of the sample. We aim to finish sorting within 4 h after digesting embryos.

We show the FACS gating strategy for isolation of embryo cells in Supplementary Fig. 2.

Sorting accuracy validation (Step 68)

After FACS, DNA of tdTomato$^+$ and tdTomato$^-$ cells are extracted separately and PCR amplified. These PCR products are ligated into T-Vectors, and 20 TA clones are picked for Sanger sequencing. Further experiments (Steps 69–76) will be conducted only if less than 10% (2/20) of clones show on-target editing in tdTomato$^-$ cells and more than 20% (4/20) of clones show on-target editing in tdTomato$^+$ cells. The sensitivity and specificity are used to validate the sorting accuracy of FACS. Some tdTomato$^+$ cells might not be edited, and tdTomato$^-$ cells might be edited, as tdTomato$^+$ and tdTomato$^-$ cells might be mis-sorted sometimes. So, to maximize the sorting accuracy, we set sensitivity and specificity thresholds here to guarantee that tdTomato$^+$ and tdTomato$^-$ cells for downstream analyses are well separated. Sensitivity is indicated by the percentage of tdTomato$^+$ cells edited, and specificity is represented by the percentage of tdTomato$^-$ cells edited (which did not receive editing reagents at the one-cell stage).

Table 1 | Primers for in vitro transcription and genotyping

Primer	Sequence (5′–3′)	Purpose	Step
Cas9 IVT F	TAATACGACTCACTATAGGGAGATTTCAGGTTGGACCGGTG	Generation of Cas9 mRNA	1
Cas9 IVT R	GACGTCAGCGTTCGAATTGC	Generation of Cas9 mRNA	1
BE3 IVT F	TCCGCGGCCGCTAATACGACT	Generation of BE3 mRNA	1
BE3 IVT R	TGGTTCTTTCCGCCTCAGAAGCC	Generation of BE3 mRNA	1
Cre IVT F	TAATACGACTCACTATAGGGAGACAGATCACCTTTCCTATCAACC	Generation of Cre mRNA	2
Cre IVT R	TCGGTATTTCCAGCACACTGGA	Generation of Cre mRNA	2
sgRNA IVT F	TAATACGACTCACTATAGGGNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAG	Generation of sgRNA	9
sgRNA IVT R	AAAAGCACCGACTCGGTGCC	Generation of sgRNA	9
Tyr-OF	GTTATCCTCACACTACTTCTG	On-target genotyping	44
Tyr-OR	GTAATCCTACCAAGAGTCTCA	On-target genotyping	44
Tyr-IF	TCCTCACACTACTTCTGATG	On-target genotyping	44
Tyr-IR	GTCTCAAGATGGAAGATCAC	On-target genotyping	44

Controls

Two kinds of controls are used in GOTI. The tdTomato$^-$ cells are sequenced and analyzed together with tdTomato$^+$ cells to eliminate the influence of genetic background. In addition, a control group with only the injection of Cre (Cre-only group) is necessary for the experiment to control the background noise in GOTI. In addition, the survival rate of embryos to blastocysts in the Cre-only group (usually more than 90%) acts as a control for evaluating the toxicity of genome editing tools. Genome editing tools with less than 80% blastocyst rate are potentially toxic to embryos, as the normal blastocyst rates are above 80%[15]. Groups with different sgRNAs and no sgRNAs are optional to control for the sgRNA-dependent and -independent off-target effects.

Materials

Biological materials

- C57BL/6 female mice; Shanghai Laboratory Animal Center (SLAC) Laboratory Animal Company (Shanghai SLAC Laboratory)
- Ai9 male mice[23] (full name: B6.Cg-Gt (ROSA) 26Sortm9 (CAG-td-Tomato) Hze/J), JAX strain 007909; The Jackson Laboratory)
- ICR female mice; Shanghai SLAC Laboratory
- Vasectomized male mice (ICR strain mice); Shanghai SLAC Laboratory **!CAUTION** Experimental procedures involving animals must be carried out according to all relevant institutional and governmental regulations.
- DH5α competent cells (Tiangen, cat. no. CB101)

Reagents

- KOD-Plus-Neo (Toyobo, cat. no. KOD-401)
- Premix Taq (Ex Taq Version 2.0 plus dye; Takara, cat. no. RR902A)
- DL10,000 DNA Marker (Takara, cat. no. 3584A)
- DL15,000 DNA Marker (Takara, cat. no. 3582B)
- 6× loading buffer (Takara, cat. no. 9156)
- EDTA (Sangon Biotech, cat. no. A500838-0500)
- Tris (Sangon Biotech, cat. no. A100826-0500)
- Glacial acetic acid (Sangon Biotech, cat. no. A501931-0500)
- PCR DNA primers oligomers (Shanghai HuaGen Biotech)
- mMESSAGE mMACHINE T7 ULTRA Kit (Life Technologies, cat. no. AM1345)
- MEGA Shortscript T7 Kit (Life Technologies, cat. no. AM1354)
- MEGA Clear Kit (Life Technologies, cat. no. AM1908)
- HEPES-CZB medium (see 'Reagent setup' section)
- Pregnant mare serum gonadotropin (PMSG; Sigma, cat. no. G4527)
- Human chorionic gonadotropin (hCG; Sigma, cat. no. C8554)
- KSOM medium (Millipore, cat. no. MR-106-D) **!CAUTION** Store the medium at −20°C. After thawing, keep it at 4 °C and use it within 2 weeks.

- Hyaluronidase (Sigma, cat. no. H-3884)
- M2 medium (Millipore, cat. no. MR-015-D)
- Cytochalasin B (Sigma, cat. no. C6762)
- DMSO (Sigma, cat. no. D2650) **! CAUTION** It is flammable, harmful if swallowed and toxic when in contact with skin and eye. Use protective gloves and safety glasses when handling.
- Mineral oil (Sigma, cat. no. M8410)
- Nuclease-free water (Life Technologies, cat. no. AM9932)
- Trypsin-EDTA (0.05% (wt/vol); Gibco, cat. no. 25300062)
- Dulbecco's Modified Eagle Medium (DMEM; Gibco, cat. no.11965092)
- Fetal bovine serum (FBS; Gibco, cat. no. 16000044)
- Tween 20 (Sangon Biotech, cat. no. A100777-0500)
- Triton X-100 (Sangon Biotech, cat. no. A110694-0500)
- Proteinase K (Tiangen, cat. no. RT403-02)
- DNeasy Blood and Tissue Kit (Qiagen, cat. no. 69504)
- Universal DNA Purification Kit (Tiangen, cat. no. DP214)
- TIANprep Rapid Mini Plasmid Kit (Tiangen, cat. no. DP105)
- px330 (Addgene plasmid no. 42230)
- px260 (Addgene plasmid no. 42229)
- pCMV-BE3 (Addgene plasmid no. 73021)
- pCAG-Cre (Yang lab no. ZP156)
- FastDigest BbsI (Thermo Fisher Scientific, cat. no. FD1014)
- FastDigest PvuI (Thermo Fisher Scientific, cat. no. FD0624)
- FastDigest NotI (Thermo Fisher Scientific, cat. no. FD0595)
- Nucleic acid dye (Tiangen, cat. no. RT210)
- NaCl (Sigma, cat. no. S-5886)
- KCl (Sigma, cat. no. P-5405)
- $CaCl_2$ $2H_2O$ (Sigma, cat. no. C-7902)
- $MgSO_4$ $7H_2O$ (Sigma, cat. no. M-5921)
- KH_2PO_4 (Sigma, cat. no. P-5655)
- EDTA·2Na·$2H_2O$ (Sigma, cat. no. E-4884)
- $NaHCO_3$ (Sigma, cat. no. S-5761)
- L-glutamine (Sigma, cat. no. G-8540)
- Na-lactate (Sigma, cat. no. L7900)
- Sodium pyouvate (Sigma, cat. no. P-8574)
- Sodium penicillin (Sigma, cat. no. P-3032)
- Streptomycin (Sigma, cat. no. S1277)
- BSA (Sigma, cat. no. A-3311)
- HEPES (Sigma, cat. no. H-7006)
- H2O (Merck Millipore, cat. no. TMS-006-B)
- 2,2,2-tribromoethanol (Sigma, cat. no. T48402)
- 2-Methyl-2-butanol (Sigma, cat. no. 240486)

Reagent setup

Ampicillin

Prepare a stock solution of 100 mg/ml in H_2O and filter sterilize. Store at −20°C for up to 12 months.

Cell lysis buffer

Component	Volume	Final concentration
Tween (1% (vol/vol) Tween in ddH_2O)	1 ml	0.1% (vol/vol)
Triton X-100 (1% (vol/vol) Triton X-100 in ddH_2O)	1 ml	0.1% (vol/vol)
Proteinase K (20 mg/ml)	2 ml	4 mg/ml
ddH_2O	6 ml	
Total	10 ml	

Store at 4 °C for up to 3 months.

LB medium
Add 10 g of bacto-tryptone, 5 g of yeast extract and 5 g of NaCl to 1 L of ddH_2O and sterilize by autoclaving. Store at 4 °C for up to 1 month.

LB agar plates
Add 12 g of agar to 1 L of LB medium before autoclaving. To prepare plates, allow medium to cool to 50 °C and then add antibiotic stock to achieve a final concentration of 100 mg/L, mix by gentle swirling and pour or pipette ~10 ml into each sterile Petri dish (60-mm diameter). Note that plates should be covered to prevent evaporation, and store agar-side up at 4 °C for up to 1 month until use.

50× TAE buffer
242 g of Tris base, 18.61 g of EDTA and 57.1 ml of glacial acetic acid (100%) per liter of water (final pH 8.5). The stock can be stored at room temperature (25–30 °C) for up to 3 months.

HEPES-CZB medium
(1 L) Store at 4 °C for up to 2 months.

Component	Final concentration (g/L)	Final concentration (mM)
NaCl	4.76	81.62
KCl	0.36	4.83
$CaCl_2$ $2H_2O$	0.25	1.70
$MgSO_4$ $7H_2O$	0.29	1.18
KH_2PO_4	0.16	1.18
EDTA·2Na·$2H_2O$	0.04	0.11
$NaHCO_3$	0.42	5.00
L-glutamine	0.15	1.00
Na-lactate	0.35	31.30
Sodium pyouvate	0.03	0.27
Sodium penicillin	3.56	100.00
Streptomycin	0.51	0.70
BSA	5	—
HEPES	0.125	0.48
H_2O	1	—

Hyaluronidase (Hy), 10× stock, 10 mg/ml
Add 100 mg of Hy to 10 ml of M2 medium, and then divide it into 100 × 100-µl tubes. Store at −20 °C for up to 12 months.

M2 + Hy, 10 mg/ml
Add 100 µl of Hy stock solution to 900 µl of M2 medium. Mix it just before use.

Cytochalasin B (CB), 50× stock, 500 µg/ml
Add 2 mg of CB to 4 ml of DMSO, and then divide it into 100 × 20-µl tubes. Store at −20 °C for up to 12 months.

HEPES-CZB + CB, 10 µg/ml
Add 20 µl of CB stock solution to 1 ml of HEPES-CZB. Mix it just before use.

Avertin A stock solution
Prepare by dissolving 10 g of 2,2,2-tribromoethanol in 10 ml of 2-methyl-2-butanol in a 50 °C water bath until it is fully dissolved. Prepare a working solution of 2.5% avertin by mixing 2.5 ml of stock solution with 97.5 ml of PBS. Sterilize by passing solution through a 0.22-mm bottle top vacuum filter. Store the solution at 4 °C in the dark for up to 3 months.

DMEM medium
Add 50 ml of FBS to 450 ml of DMEM. Store at 4 °C for up to 2 months.

Equipment

- Cell strainer (40 μm; Falcon, cat. no. 352340)
- Polystyrene round-bottom tube (5 ml, with 35-μm cell strainer cap, 12 × 75 mm^2; BD Biosciences, cat. no. 352235) and Falcon 10-cm (100 × 20 mm) dishes; bottoms are suitable for oocyte/embryo collection, and lids that are suitable for micromanipulation (Becton Dickinson, cat. no. 353003)
- Falcon 6-cm (60 × 15-mm) dishes (Becton Dickinson, cat. no. 351007)
- Microloader tips (Eppendorf, 5242 956.003)
- Thin Wall Borosilicate Glass with Filament (Borosilicate, BF100-78-10)
- Thin Wall Borosilicate Glass without Filament (Borosilicate, B100-75-10)
- 0.2-ml PCR tubes (Axygen, cat. no.14-222-262)
- FemtoJet microinjector (Eppendorf)
- Inverted microscope with Hoffman optics (Olympus, IX73)
- Micromanipulator set (Narishige, MMO-202ND)
- CO_2 incubator (Thermo, BB15)
- Stereo microscope (Olympus, SZ61)
- Cell sorter (Beckman, MoFlo XDP)
- Centrifuge (Eppendorf, 5424R)
- Micropipette puller (Sutter Instrument, P97)
- Micropipette microforge (Narishige, MF-900)
- Fluorescence microscopy (Olympus, BX51)

Procedure

Generation of Cas9 or BE3 mRNA ● Timing ~10 h

1 For vectors without T7 promoter, first provide a template for in vitro transcription of Cas9 or BE3 mRNA and add the T7 promoter sequence to the Cas9 or BE3 coding region by PCR amplification using the appropriate primer pair listed in Table 1 and the following reaction mix, using Cas9 as an example (see Box 1 instead for details of how to generate Cas9 or BE3 mRNA with T7 promoters):

Component	Final concentration	Volume
10× PCR Buffer for KOD -Plus- Neo	5 μl	1×
2 mM dNTPs	5 μl	0.2 mM each
Cas9 IVT F (10 μM) for Cas9	2 μl	0.4 μM
Cas9 IVT R (10 μM) for Cas9	2 μl	0.4 μM
KOD -Plus-Neo (1 U/μl)	1 μl	1.0 U/50 μl
px260 (100 ng/μl) for Cas9	0.5 μl	1 ng/μl
25 mM $MgSO_4$	3 μl	1.5 mM
ddH_2O	31.5 μl	
Total	50 μl	

2 Perform PCR using the following cycling conditions:

Cycle number	Denature	Extend
1	95 °C, 5 min	
2–34	95 °C, 30s	68 °C, 5 min
35		72 °C, 15 min

3 Run 5 μl of PCR products on a 1% (wt/vol) agarose gel at 110 V for 30 min with 6× loading buffer in TAE buffer to validate that the DNA fragment is unique and of the expected size (~4.5 kb for Cas9 and ~5.4 kb for BE3; Supplementary Fig. 3).

4 Purify PCR products, using the Universal DNA Purification Kit according to the manufacturer's instructions.

5 Use 1 μg of purified PCR product as the template for in vitro transcription of Cas9 mRNA using the mMESSAGE mMACHINE T7 Kit according to the kit protocol.

6 Purify the mRNA using the MEGAclear Kit following the manufacturer's instructions and elute the RNA in 100 µl of TB buffer. Determine the RNA concentration using the NanoDrop 2000 spectrophotometer following the manufacturer's instructions.

7 Dilute the purified mRNA to 500 ng/µl in 0.1 mM RNase-free ddH_2O and check its quality on a 1% (wt/vol) agarose gel in TAE buffer. The in vitro transcribed Cas9 mRNA should have a band at ~1,000 bp (Supplementary Fig. 3). Failed in vitro transcription products show no clear bands at these positions as also shown in Supplementary Fig. 3.

8 Dispense 1 µl of the purified mRNA into 0.2-ml PCR tubes.

PAUSE POINT The samples can be stored at −80 °C for up to 1 year.

Generation of Cre mRNA ● Timing ~10 h

9 To provide a template for in vitro transcription of Cre mRNA, add the T7 promoter sequence to the Cre coding region by PCR amplification using the appropriate primer pair listed in Table 1 and the following reaction mix:

Component	Volume	Final Concentration
10× PCR Buffer for KOD -Plus- Neo	5 µl	1×
2 mM dNTPs	5 µl	0.2 mM each
Cre IVT F (10 µM) for Cre	2 µl	0.4 µM
Cre IVT F (10 µM) for Cre	2 µl	0.4 µM
KOD -Plus- Neo (1 U/µl)	1 µl	1.0 U/50 µl
pCAG-Cre (100 ng/µl) for Cre	0.5 µl	1 ng/µl
25 mM $MgSO_4$	3 µl	1.5 mM
ddH_2O	31.5 µl	
Total	50 µl	

10 Repeat Steps 2−8, diluting the purified sgRNA to 500 ng/µl in 0.1 mM EDTA. The band in Step 3 should be of the expected size (~2 kb; Supplementary Fig. 3). The in vitro transcribed Cre mRNA should have a band at ~400 bp (Supplementary Fig. 3).

PAUSE POINT The samples can be stored at −80 °C for up to 1 year.

Generation of sgRNA ● Timing ~8 h

11 For sgRNA preparation, prepare the following PCR reaction mix to add the T7 promoter sequence to the sgRNA template by PCR amplification using the appropriate primer pair listed in Table 1:

Component	Volume	Final concentration
10× PCR buffer for KOD -Plus- Neo	5 µl	1×
2 mM dNTPs	5 µl	0.2 mM each
sgRNA IVT F (10 µM)	2 µl	0.4 µM
sgRNA IVT R (10 µM)	2 µl	0.4 µM
KOD -Plus- Neo (1.0 U/µl)	1 µl	1.0 U/50 µl
px330 (100 ng/µl)	0.5 µl	1 ng/µl
25 mM $MgSO_4$	3 µl	1.5 mM
ddH_2O	31.5 µl	
Total	50 µl	

12 Perform PCR using the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	95 °C, 5 min		
2−34	95 °C, 30 s	60 °C, 30 s	72 °C, 30 s
35			72 °C, 5 min

13 Repeat Step 10 (the band in Step 3 should be of the expected size (120 bp), and the in vitro transcribed sgRNA should have a clear band at ~100 bp (Supplementary Fig. 3)).
PAUSE POINT The samples can be stored at −80 °C for up to 1 year.

Superovulation and mating ● Timing ~3 d and ~2 h hands-on

14 Inject 10 female C57BL/6 mice (3−4 weeks old) with 5 IU of PMSG through intraperitoneal (i.p.) injection at 14:00 on day 1.
? TROUBLESHOOTING

15 Approximately 47−49 h after PMSG injection (13:00–15:00 on day 3), inject the female mice i.p. with 5 IU of hCG to induce ovulation.

16 Put each hormone-stimulated female together with Ai9 males in a 1:1 ratio in a mating cage overnight.
? TROUBLESHOOTING

Zygote collection and processing ● Timing ~3 h

17 Prepare two 35-mm dishes each containing 14 drops (30 μl for each drop) of KSOM medium covered with mineral oil for embryo culture. Transfer dishes into a 37 °C incubator for at least 20 min before use. Prepare one 100-mm dish with eight drops (200 μl for each drop) of M2 medium.

18 Euthanize females by CO_2 asphyxiation at 22–24 h after hCG injection.

19 Isolate oviducts and place all the oviducts into one 200-μl drop of M2 medium in the 100-mm dishes (Supplementary Fig. 4).

20 Under the stereoscopic microscope, transfer one oviduct at a time into the second 200-μl drop of M2 medium (prepared in Step 17). Tear the oviduct where it is most swollen using a 1-ml syringe attached to a 26-gauge needle, releasing the zygote-cumulus complexes (ZCCs). Repeat this step for every oviduct until all the zygotes are released.

21 Add 200 μl of hyaluronidase to the ZCCs in the droplet and then place the dish into a 37 °C incubator for 3 min.
▲CRITICAL STEP Incubate for no longer than 3 min.

22 Pipette the droplet up and down several times with a yellow tip until the cumulus cells are completely removed from the zygotes. Successively transfer the zygote clockwise through five wash drops of M2 medium (prepared in Step 17).

23 Finally, transfer the zygotes to one drop of KSOM medium in the pre-equilibrated 35-mm dish from Step 17 using a hand pipette; pass the zygotes through 6–10 additional KSOM droplets to wash; and place the dish at 37 °C in a 5% CO_2 incubator up to 24 h until ready for injection.

Microinjection preparation ● Timing ~20 min

24 Approximately 48−50 h after hCG injection, the zygotes develop to the end of the two-cell stage, and the cytoplasms of the two blastomeres should be completely separated.

25 Prepare the appropriate injection mix depending on the aim of the experiment, as outlined in the table below. Cas9 or BE3 mRNA (from Step 8 or Box 1) and sgRNA stock from Step 13 are all used at concentrations of 500 ng/μl. Cre mRNA stock (from Step 10) should be diluted down to 20 ng/μl with nuclease-free water. Combine the following reagents at room temperature in a sterile, RNase-free eight-well PCR strip:

Component	Volume	Final concentration
Cas9/BE3 mRNA (500 ng/μl)	1 μl	50 ng/μl
sgRNA (500 ng/μl)	1 μl	50 ng/μl
Cre mRNA (20 ng/μl)	1 μl	2 ng/μl
Nuclease-free water	7 μl	—
Total	10 μl	—

Mix all the injection components just before use.

26 Centrifuge components at 15,000*g* for 5 min at 4 °C and and store on ice ready for microinjection.

27 Load 3 μl of injection mix into the injection needle using a microloader tip. Keep the injection needle vertically for 5 min by sticking it to a piece of plasticine to remove small air bubbles in the needle.

28 Place a droplet of HEPES-CZB + CB on top of a 10-cm dish and then cover it with mineral oil. Place the dish under a stereoscopic microscope.
29 Attach the loaded injection needle to the instrument holder connected to the FemtoJet; lower the capillary into the medium drop on the stage such that it is positioned in the center just below the embryos; and switch on the FemtoJet and allow it to reach pressure.
30 Attach the holding pipette to the other side of the micromanipulator.

Injection of embryos ● Timing ~1 h

31 Transfer 100 two-cell stage embryos into a large drop of HEPES-CZB + CB (prepared in Step 28) using a hand pipette (Supplementary Video 1). The number of zygotes to be moved into the microinjection drop should be determined by the skills of the injector and quality of the setup.
▲ CRITICAL STEP Do not attempt to work with more zygotes than can be injected within 30 min.
32 Ensure that the microinjection capillary is open and not clogged by placing the tip of the microinjection capillary close to a zygote in the same horizontal plane under a continuous flow stream. If the microinjection capillary is open, a stream of DNA will move the zygote away from the tip of the microinjection capillary. If the injection pipette is not open, tip it carefully on the holding pipette until it breaks at a larger dimension.
33 Hold an embryo using a holding pipette. Insert the injection tip into one blastomere and pause briefly halfway inside the blastomere to see the formation of a small droplet around the injection tip.
34 When all the embryos in the chamber have been injected, they should immediately be moved back into KSOM medium from Step 17, washed 6–8 times and incubated at 37 °C in 5% CO_2 for 30 min. At this point, 80% of the embryos will be transferred to a foster mother oviduct for development to E14.5 (Steps 35–43), whereas the other 20% will be cultured in vitro to blastocyst stage for genotyping (Steps 44–57).

Reimplantation of injected embryos ● Timing ~2 h

35 Prepare four pseudo-pregnant foster mothers by mating estrous ICR female mice with vasectomized male mice on the same day as zygote collection.
36 On the morning of embryo transfer, identify 0.5 days-post-coitum foster mothers with visible copulatory plugs.
37 Anesthetize foster mothers by peritoneal injection of Avertin using a 1-ml syringe attached to a 26-gauge needle, after having weighed foster mothers to calculate the injection dose of Avertin (2.5% Avertin solution at a dose of 0.01 ml/g of body weight).
38 After 1 min, check that the mouse is fully anesthetized by lightly pinching the most medial toe. When the anesthetized mouse is unresponsive, place it under the stereomicroscope and disinfect with 75% (vol/vol) ethanol.
39 Make a small longitudinal incision (≤1 cm) parallel to the midline at the level of the last rib and expose the body wall by sliding aside the skin. Use forceps to pick up the body wall and make a small incision through the body wall over the site of the ovary. Expose the ovary, oviduct and part of the uterus through this incision.
40 Load the glass transfer pipette by drawing KSOM medium ~1 cm up the pipette. Next, draw up one small air bubble and then 18–22 embryos from Step 34 in a minimal amount of medium, followed by another small air bubble.
41 Make a small hole in the upper ampulla using a 30-gauge needle. Insert a glass transfer pipette into the hole and transfer the embryos with air bubbles.
42 Remove the transfer pipette; gently return the ovary, oviduct and uterus back inside the body; and seal the incision with absorbable sutures.
43 Keep the mice warm on a 37 °C warming plate until the mouse recovers from the effects of the anesthesia.
? TROUBLESHOOTING

Nested PCR detection of targeted embryos ● Timing ~6 h

44 Incubate the remaining 20% of injected embryos from Step 34 at 37 °C and 5% CO_2 for 3 d until blastocyst stage.
45 Wash single blastocysts 3–6 times with KSOM.

46 Transfer blastocysts directly into PCR tubes (one blastocyst per tube) and add 1.5 μl of embryo lysis buffer. ▲CRITICAL STEP Use the stereo microscope to confirm that the blastocysts are indeed put into PCR tubes.

47 Incubate at 56 °C for 30 min and heat inactivate at 95 °C for 10 min.

48 Prepare nested PCR amplification reactions to amplify the crude DNA solutions, using the following reaction mix:

Component	Volume	Final concentration
Premix Taq (Ex Taq Version 2.0 plus dye)	25 μl	1×
Tyr-OF/IF (10 μM)	2 μl	0.4 μM
Tyr-OR/IR (10 μM)	2 μl	0.4 μM
Crude DNA solution	1.5 μl	
Nuclease-free water	Up to 50 μl	

49 Perform PCR using the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	95 °C, 5 min		
2–34	95 °C, 30 s	60 °C, 30 s	72 °C, 30 s
35			72 °C, 5 min

50 Purify PCR products from Step 49, using the Universal DNA Purification Kit according to the manufacturer's instructions.

51 Ligate purified PCR products into pMD-19T using the pMD-19T Cloning Vector Kit, according to the manufacturer's instructions, as follows:

Component	Volume	Final concentration
pMD-19T (50 μg/μl)	0.5 μl	2.5 μg/μl
PCR product (50 μg/μl)	1 μl	50 μg/μl
Solution I	5 μl	1×
Nuclease-free water	Up to 10 μl	

Mix all the components and incubate at 56 °C for 30 min.

52 To clone into *E. coli* (TA clone), transform the above reaction mix into *E. coli* DH5α by adding 10 μl of each ligation reaction directly into a tube of 50 μl of DH5α competent cells, mix by tapping gently and incubate the cells on ice for 30 min. Heat shock the cells for 45 s in a 42°C water bath and incubate the cells on ice for 2 min. Add 500 μl of room temperature LB medium and then shake the cells at 37 °C for 1 h at 225 r.p.m. in a shaking incubator.

53 Plate all of the transformation onto a 10-cm LB agar plate with 100 μg/ml of ampicillin and incubate plates at 37 °C overnight.

54 Pick a single colony from an agar plate using a pipette tip and drop the pipette tip in a tube containing 5 ml of LB medium with 100 μg/ml of ampicillin.

55 Incubate the bacteria in a shaking incubator for 15 h at 37 °C at 200 r.p.m.

56 Extract plasmid from *E. coli* using the TIANprep Rapid Mini Plasmid Kit according to the manufacturer's instructions.

57 Determine the number of clones with indels or SNVs by Sanger sequencing. Pick 20 TA clones for Sanger sequencing, and only conduct further experiments if the on-target editing was more than 20% (4/20 clones).

Isolation of embryonic cells and FACS ● Timing ~10 h

58 Euthanize 14.5-d pregnant mice from Step 43 by cervical dislocation or anesthesia and separate the embryos with the placenta, and then aseptically dissect mouse embryos. Wash the embryos with PBS twice briefly.

59 Place each embryo into a 100-mm Petri dish and mince them into the smallest possible pieces less than 1 mm.
60 Add 5 ml of pre-warmed 37 °C Trypsin-EDTA (0.05%) and transfer to a 500-ml centrifuge tube and incubate at 37 °C for 30 min.
61 Add 5 ml of DMEM and homogenize fetal tissues by passing 30–40 times through a 1-ml pipette tip.
62 Centrifuge the cell suspension at 200*g* for 6 min at room tempetature and resuspend the pellet in 5 ml of DMEM.
63 Filter the cell suspension through a 40-μm cell strainer.
64 Centrifuge the cell suspension at 200*g* for 6 min at room temperature, aspirate the medium gently without disturbing the pellet and resuspend the pellet in 5 ml of DMEM.
65 Pass cell suspension through a 35-μm cell strainer cap into a 5-ml polystyrene round-bottom tube.
66 Sort the tdTomato$^+$ cells and tdTomato$^-$ cells into two separate tubes (~4 million cells for each) using a Moflo XDP Fluorescence-Activated Cell Sorter (Beckman Coulter), respectively. The FSC/SSC gates of the starting cell population were set to include all cells. Then, doublet cells were excluded by SSC-H versus SSC-A. Positive and negative boundaries were defined by control progeny cells of non-edited blastomeres. (Supplementary Fig. 2).
67 Extract genomic DNA from sorted cells using the DNeasy Blood and Tissue Kit according to the manufacturer's instructions.

Sample quality control ● Timing ~36 h

68 Perform nested PCR amplification reactions on the extracted purified genomic DNA from tdTomato$^+$ cells and tdTomato$^-$ cells, respectively, as described in Steps 48–57. Conduct further experiments (Steps 69–78) only if fewer than 10% (2/20) of clones show on-target editing in tdTomato$^-$ cells and more than 20% (4/20) of clones show on-target editing in tdTomato$^+$ cells.

WGS ● Timing ~3 d

69 Perform WGS on the extracted genomic DNA from Step 67 at mean coverage of 50× by Illumina HiSeq X Ten.

Processing of raw reads ● Timing ~20 h

70 Quality control raw sequencing reads by fastQC, using the following commands:

```
fastqc raw/cre_neg_R1.fastq.gz raw/cre_neg_R2.fastq.gz -o fastQC/pretrim
fastqc raw/cre_pos_R1.fastq.gz raw/cre_pos_R2.fastq.gz -o fastQC/pretrim
```

If the quality of sequencing reads is good (quality score per base >20; Supplementary Fig. 5), skip to Step 73. Otherwise, apply Step 72 to trim low-quality reads and remove adapter sequences (Supplementary Fig. 5).

71 (Optional) Use Trimmomatic to trim the low-quality reads and adapter sequences in the FASTQ files, using the following commands:

```
java -jar trimmomatic-0.36.jar PE -threads 8 cre_neg_R1.fastq.gz cre_neg_R2.
fastq.gz
fastQC/trim/cre_neg_R1_paired.fastq.gz  fastQC/trim/cre_neg_R1_unpaired.
fastq.gz
fastQC/trim/cre_neg_R2_paired.fastq.gz  fastQC/trim/cre_neg_R2_unpaired.
fastq.gz ILLUMINACLIP:TruSeq3-
PE.fa:2:30:10 LEADING:5 TRAILING:5 MINLEN:70
java -jar trimmomatic-0.36.jar PE -threads 8 cre_neg_R1.fastq.gz cre_
neg_R2.fastq.gz
fastQC/trim/cre_pos_R1_paired.fastq.gz fastQC/trim/cre_pos_R1_unpaired.
fastq.gz
fastQC/trim/cre_pos_R2_paired.fastq.gz  fastQC/trim/cre_pos_R2_unpaired.
fastq.gz ILLUMINACLIP:TruSeq3-
PE.fa:2:30:10 LEADING:5 TRAILING:5 MINLEN:70
```

```
fastqc fastQC/trim/cre_neg_R1_paired.fastq.gz fastQC/trim/cre_neg_
R2_paired.fastq.gz -o fastQC/posttrim
fastqc fastQC/trim/cre_pos_R1_paired.fastq.gz fastQC/trim/cre_pos_
R2_paired.fastq.gz -o fastQC/posttrim
```

72 Construct the BWA index file of the reference genome mm10.fa, using the following command:

```
bwa index mm10.fa
```

73 Align clean reads to the reference genome using BWA mem and use Picard-tools to reorder, sort, add read groups and mark duplicates of the aligned BAM files as follows:

```
bwa mem -t 8 -M mm10.fa fastQC/trim/cre_neg_R1_paired.fastq.gz fastQC/
trim/cre_neg_R2_paired.fastq.gz | samtools view -bS -o cre_neg.bam
bwa mem -t 8 -M mm10.fa fastQC/trim/cre_pos_R1_paired.fastq.gz fastQC/
trim/cre_pos_R2_paired.fastq.gz | samtools view -bS -o cre_pos.bam
java -Xmx20g -jar picard.jar ReorderSam INPUT=cre_neg.bam OUTPUT=
cre_neg.reorder.bam
REFERENCE=mm10.fa
java -Xmx20g -jar picard.jar ReorderSam INPUT=cre_pos.bam OUTPUT=
cre_pos.reorder.bam
REFERENCE=mm10.fa
java -Xmx20g -jar picard.jar SortSam INPUT=cre_neg.reorder.bam OUTPUT=
cre_neg.reorder.sort.bam
SORT_ORDER=coordinate
java -Xmx20g -jar picard.jar SortSam INPUT=cre_pos.reorder.bam OUTPUT=
cre_pos.reorder.sort.bam
SORT_ORDER=coordinate
java -Xmx20g -jar picard.jar AddOrReplaceReadGroups VALIDATION_STRIN-
GENCY=SILENT
INPUT=cre_neg.reorder.sort.bam OUTPUT=cre_neg.reorder.sort.add.bam
RGLB=WES
RGPL=Illumina RGPU=HiSeq RGSM=cre_neg
java -Xmx20g -jar picard.jar AddOrReplaceReadGroups VALIDATION_
STRINGENCY=SILENT
INPUT=cre_pos.reorder.sort.bam OUTPUT=cre_pos.reorder.sort.add.bam
RGLB=WES RGPL=Illumina RGPU=HiSeq RGSM=cre_neg
java -Xmx20g -jar picard.jar MarkDuplicates VALIDATION_STRINGENCY=
SILENT
INPUT=cre_neg.reorder.sort.add.bam OUTPUT=cre_neg.reorder.sort.
add.mkdup.bam
METRICS_FILE=cre_neg.reorder.sort.add.mkdup.metrics
java -Xmx20g -jar picard.jar MarkDuplicates VALIDATION_STRINGENCY=
SILENT
INPUT=cre_pos.reorder.sort.add.bam OUTPUT=cre_pos.reorder.sort.
add.mkdup.bam
METRICS_FILE=cre_pos.reorder.sort.add.mkdup.metrics
java -Xmx20g -jar picard.jar BuildBamIndex VALIDATION_STRINGENCY=
SILENT
INPUT=cre_neg.reorder.sort.add.mkdup.bam
java -Xmx20g -jar picard.jar BuildBamIndex VALIDATION_STRINGENCY=
SILENT
INPUT=cre_pos.reorder.sort.add.mkdup.bam
```

Detection of off-target effects ● Timing ~30 h

74 To detect whether off-target variants specifically existed in tdTomato$^+$ cells, apply three variant calling tools to identify SNVs and indels by comparing tdTomato$^+$ cell with tdTomato$^-$ cells from the same embryo, using the following commands:

```
java -Xmx20g -jar GenomeAnalysisTK.jar -R mm10.fa -T MuTect2 -I:tumor
cre_pos.reorder.sort.add.mkdup.bam -I:normal cre_neg.reorder.sort.
add.mkdup.bam -o
mutect/cre_pos/output.vcf lofreq somatic -n cre_neg.reorder.sort.add.
mkdup.bam -t
cre_pos.reorder.sort.add.mkdup.bam -f mm10.fa -threads 8 -o lofreq/
cre_pos strelka-
2.7.1.centos5_x86_64/bin/configureStrelkaSomaticWorkflow.py --normalBam
cre_neg.reorder.sort.add.mkdup.bam --tumorBam cre_pos.reorder.sort.
add.mkdup.bam --
referenceFasta mm10.fa --runDir strelka/cre_pos
scalpel-0.5.3/scalpel-discovery --somatic -normal cre_neg.reorder.
sort.add.mkdup.bam --
tumor cre_pos.reorder.sort.add.mkdup.bam -bed mm10.bed --window 600
--numprocs 8 --ref
mm10.fa --dir scalpel/cre_pos
awk '$7=="PASS" {print $0}' strelka/cre_pos/results/variants/
somatic.snvs.vcf >
strelka/cre_pos/results/variants/somatic.snvs.pass.vcf
awk '$7=="PASS" {print $0}' strelka/cre_pos/results/variants/
somatic.indels.vcf >
strelka/cre_pos/results/variants/somatic.indels.pass.vcf
awk '$7=="PASS" {print $0}' mutect/cre_pos/output.vcf > mutect/
cre_pos/output.pass.vcf
awk 'length($4)==1 && length($5)==1 {print $0}' mutect/cre_pos/
output.pass.vcf >
mutect/cre_pos/output.pass.snv.vcf
awk 'length($4)>1 || length($5)>1 {print $0}' mutect/cre_pos/output.
pass.vcf >
mutect/cre_pos/output.pass.indel.vcf
```

The SNVs identified by all three algorithms—Mutect2, Strelka2 and Lofreq—are considered true SNVs (Supplementary Fig. 6), and the overlap of Mutect2, Strelka2 and Scalpel identifies true indels:

```
perl filter_overlap.pl mutect/cre_pos/output.pass.snv.vcf lofreq/
cre_pos/somatic_final.snvs.vcf
strelka/cre_pos/results/variants/somatic.snvs.pass.vcf cre_pos.snv.
overlap.vcf
perl filter_overlap.pl mutect/cre_pos/output.pass.indel.vcf scalpel/
cre_pos/somatic_final.indels.vcf
strelka/cre_pos/results/variants/somatic.indels.pass.vcf cre_pos.indel.
overlap.vcf
```

Validation of the off-target variants using Sanger sequencing ● Timing ~2 d

75 Design primers for a dozen of the detected SNVs for Sanger sequencing. The primers should have ~50% GC content.

76 Perform PCR amplification as described in Steps 48 and 49 and Sanger sequence the purified PCR products.

Filtration and functional annotation ● Timing ~10 min

77 Use Annovar as follows to annotate the identified variants from Step 74 and use only those with allele frequency more than 10% for the final results:

```
annotate_variation.pl -buildver mm10 -downdb -webfrom annovar refGene
mousedb/
table_annovar.pl cre_pos.snv.overlap.vcf mousedb -buildver mm10 -out
cre_pos.snv.overlap.vcf.anno -remove -
protocol refGene -operation g -nastring. -vcfinput
perl  awk_anno.pl  cre_pos.snv.overlap.vcf.anno.mm10_multianno.txt
cre_pos.anno.tsv
awk '$10>0.1 {print $0}' cre_pos.anno.tsv > cre_pos.anno.0.1.tsv
```

Sequence comparison between off-target and on-target variants ● Timing ~20 min

78 Retrieve the adjacent 23-bp sequences of the off-target variants (5 bp upstream and 17 bp downstream as the target base of on-target site in this study is the 6th nucleotide of the sgRNA sequence) from the marked BAM file (output from Step 74) and then blast with the on-target sequence sgRNA.fasta (23 bp; 20 bp sgRNA target sequence + 3 bp PAM). Meanwhile, blast the predicted off-target sites from Cas-OFFinder[48] based on sequence similarity with the on-target sequences using National Center for Biotechnology Information Basic Local Alignment Search Tool as follows:

```
cat cre_pos.anno.0.1.tsv | while read line
do
chr=$(echo $line | cut -d" " -f 1)
pos=$(echo $line | cut -d" " -f 2)
start=`expr $pos - 5`
end=`expr $pos + 17`
samtools faidx mm10.fa $chr:${start}-${end} » cre_pos.anno.0.1.tsv.
fasta
done
makeblastdb -in sgRNA.fasta -dbtype nucl -parse_seqids
blastn -db sgRNA.fasta -query cre_pos.anno.0.1.tsv.fasta -dust no
-outfmt 0 -word_size=7 -out
cre_pos.sgRNA.blast.out
```

This step reveals the sequence similarity between off-target variants and on-target edits. High sequence similarities indicate the off-target variants are associated with sgRNAs, whereas low sequence similarities suggest that the off-target effects are sgRNA independent.

Troubleshooting

Troubleshooting advice can be found in Table 2.

Table 2 | Troubleshooting table

Step	Problem	Possible reason	Solution
15	Too many or too few eggs	Incorrect age of female mice	Select female mice that are 3–4 weeks old
16	Low fertilization ratio	Frequent use of male mice	Mate at most twice a week for the male mice, with at least 2 d in between
44–46	Low development rate	The injected mRNA was not pure and is toxic	Re-generate the mRNAs from Steps 1–16
58–66	Too low a proportion of tdTomato$^+$ cells after FACS	Gene editing tools might be toxic for the development of tdTomato$^+$ cells	Choose high-fidelity version of gene editing tools
		The targeting gene is essential or important to the development of cells	Select non-essential genes as target genes
		The quality of mRNA is inadequate	Re-prepare mRNA for injection
74	Mutect2 takes too much time to complete	Mutect2 takes a lot of memory when running WGS data	Split the WGS into separate chromosomes to run each time

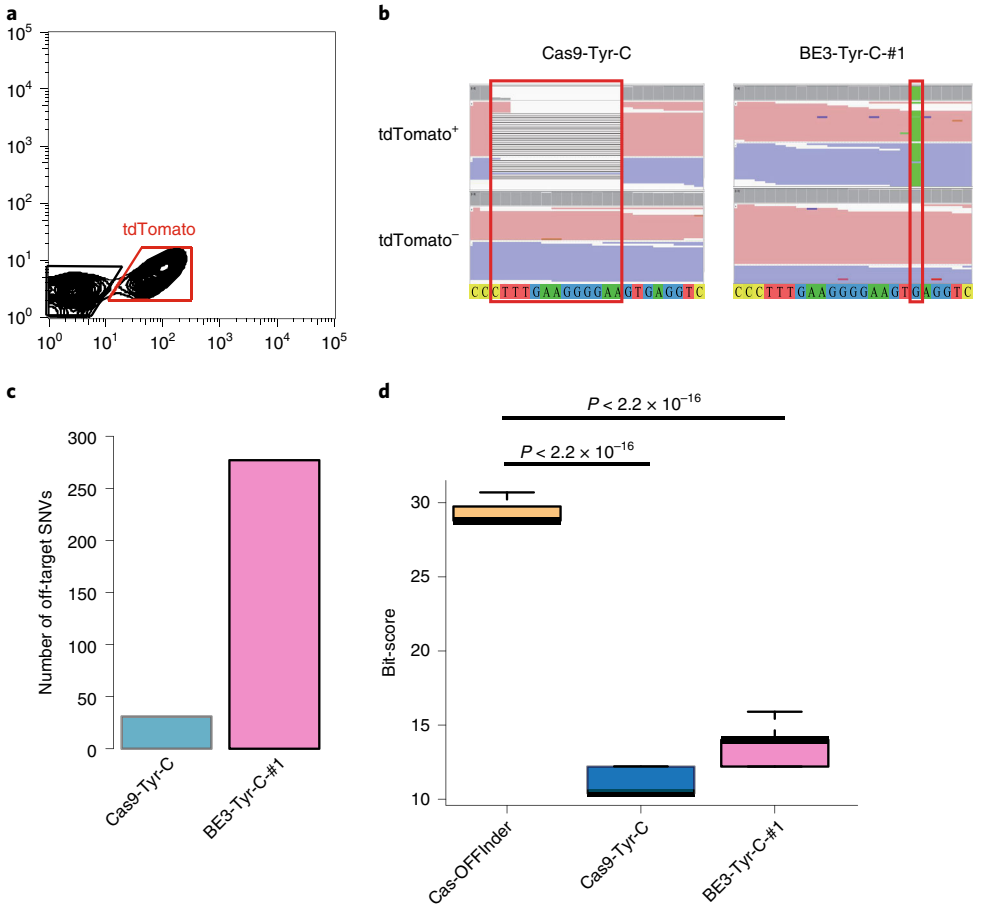

Fig. 3 | Anticipated results from GOTI. a, The separation of tdTomato$^+$ and tdTomato$^-$ cells. FACS analysis for E14.5 embryo of Cas9-Tyr-C[15]. **b**, On-target efficiency for tdTomato$^+$ and tdTomato$^-$ cells based on WGS for Cas9-Tyr-C[15] and BE3-Tyr-C-#1[15]-treated embryos. The target site was edited with high efficiency in tdTomato$^+$ cells but not in tdTomato$^-$ cells. The deletions induced by Cas9-Tyr-C and base substitutions in BE3-Tyr-C-#1 are highlighted in red rectangles. Dark lines represent deletions in the region, and green rectangles represent substitutions. **c**, The number of off-target SNVs detected in Cas9-Tyr-C and BE3-Tyr-C-#1-treated embryos. BE3-Tyr-C-#1 induced much higher number of off-target variants than Cas9-Tyr-C-#1. **d**, The sequence similarity (Bit-score) between on- and off-target sequences identified by GOTI in Cas9-Tyr-C[15] and BE3-Tyr-C-#1[15]-treated embryos or predicted by Cas-OFFinder. The off-target SNVs are sgRNA independent. $n = 1{,}809$ for Cas-OFFinder, $n = 18$ for Cas9-Tyr-C and $n = 247$ for BE3-Tyr-C-#1. Box and whisker plots: center line indicates the median, and the bottom and top lines of the box represent the first quartile and third quartile of the values, respectively. The bottom and top lines represent the minimum and maximum value. *P* values are calculated with two-sided Wilcoxon rank-sum tests.

Timing

Steps 1–8, generation of Cas9 mRNA: ~10 h
Steps 9–11, generation of Cre mRNA: ~10 h
Steps 12–13, generation of sgRNA: ~8 h
Steps 14–16, superovulation and mating: ~3 d
Steps 17–23, zygote collection and processing: ~3 h
Steps 24–30, microinjection preparation: ~20 min
Steps 31–34, injection of embryos: ~1 h
Steps 35–43, reimplantation of injected embryos: ~2 h
Steps 44–57, nested PCR detection of targeted embryos: ~6h
Steps 58–67, isolation of embryonic cells and FACS: ~10 h
Step 68, sample quality control: ~36 h
Step 69, WGS: ~3 d
Steps 70–73, processing of raw reads: ~20 h
Step 74, detection of off-target effects: ~30 h

Steps 75 and 76, validation of the off-target variations using Sanger sequencing: ~2 d
Step 77, filtration and functional annotation: ~10 min
Step 78, sequence comparison between off-target and on-target sequences ~20 min

Anticipated results

A successful application of the experimental procedure will generate two equal populations of cells with tdTomato or not (Fig. 3a). Embryos with obviously unequal proportions of tdTomato$^+$ cells should not be sequenced for further analysis (Supplementary Fig. 7). Examples of the on-target efficiency for CRISPR–Cas9 (Cas9-Tyr-C[15]) and BE3 (BE3-Tyr-C-#1[15]) editing after injection of sgRNA are shown in Fig. 3b. The bioinformatic pipeline will reveal the number of off-target variants in each embryo (Supplementary Fig. 6). The off-target mutations induced by CRISPR–Cas9 or BE3 editing were compared, and BE3 induced many more off-target edits than CRISPR–Cas9 (Fig. 3c). The off-target variants can be used for downstream analysis by group comparison, mutation bias and functional enrichment[15]. The adjacent sequences of identified variants are compared with the on-target sites (Fig. 3d), but poor sequence similarity is observed (mean Bit-score = 10.4 for Cas9-Tyr-C and mean Bit-score = 12.2 for BE3-Tyr-C-#1). By contrast, potential off-target sequences predicted by Cas-OFFinder show high similarity with the on-target sequence (mean Bit-score = 29.4; Fig. 3d). These results suggest that the off-target variants identified by GOTI in CRISPR–Cas9 or BE3 editing are sgRNA independent.

Reporting Summary

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

Data availability

The sequencing data were deposited in the National Center for Biotechnology Information's Sequence Read Archive under project accession SRP119022 and http://www.biosino.org/node/project/detail/OEP000195.

Code availability

The GOTI-seq pipeline is publicly available in GitHub at https://github.com/sydaileen/GOTI-seq. The code in this protocol has been peer reviewed.

References

1. Knott, G. J. & Doudna, J. A. CRISPR–Cas guides the future of genetic engineering. *Science* **361**, 866–869 (2018).
2. de la Fuente-Nunez, C. & Lu, T. K. CRISPR–Cas9 technology: applications in genome engineering, development of sequence-specific antimicrobials, and future prospects. *Integr. Biol.* **9**, 109–122 (2017).
3. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR–Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
4. Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR–Cas9. *Science* **346**, 1258096 (2014).
5. Cox, D. B., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nat. Med.* **21**, 121–131 (2015).
6. Musunuru, K. The hope and hype of CRISPR–Cas9 genome editing: a review. *JAMA Cardiol.* **2**, 914–919 (2017).
7. Jiang, F. & Doudna, J. A. CRISPR–Cas9 structures and mechanisms. *Annu. Rev. Biophys.* **46**, 505–529 (2017).
8. Wyman, C. & Kanaar, R. DNA double-strand break repair: all's well that ends well. *Annu. Rev. Genet.* **40**, 363–383 (2006).
9. Paquet, D. et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**, 125–129 (2016).
10. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
11. Nishida, K. et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**, aaf8729 (2016).
12. Gaudelli, N. M. et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).
13. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* **19**, 770–788 (2018).

14. Pattanayak, V. et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).
15. Zuo, E. et al. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* **364**, 289–292 (2019).
16. Jin, S. et al. Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* **364**, 292–295 (2019).
17. Grunewald, J. et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* **569**, 433–437 (2019).
18. Zhou, C. et al. Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* **571**, 275–278 (2019).
19. Cameron, P. et al. Mapping the genomic landscape of CRISPR–Cas9 cleavage. *Nat. Methods* **14**, 600–606 (2017).
20. Frock, R. L. et al. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* **33**, 179–186 (2015).
21. Tsai, S. Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR–Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015).
22. Tsai, S. Q. et al. CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets. *Nat. Methods* **14**, 607–614 (2017).
23. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
24. Wang, L. et al. CRISPR–Cas9-mediated genome editing in one blastomere of two-cell embryos reveals a novel Tet3 function in regulating neocortical development. *Cell Res.* **27**, 815–829 (2017).
25. O'Rawe, J. et al. Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. *Genome Med.* **5**, 28 (2013).
26. Bao, R. et al. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Inform.* **13**, 67–82 (2014).
27. Cameron, D. L., Di Stefano, L. & Papenfuss, A. T. Comprehensive evaluation and characterisation of short read general-purpose structural variant calling software. *Nat. Commun.* **10**, 3240 (2019).
28. Field, M. A. et al. Recurrent miscalling of missense variation from short-read genome sequence data. *BMC Genomics* **20**, 546 (2019).
29. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
30. Kim, S. et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods* **15**, 591–594 (2018).
31. Wilm, A. et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res.* **40**, 11189–11201 (2012).
32. Narzisi, G. et al. Accurate de novo and transmitted indel detection in exome-capture data using micro-assembly. *Nat. Methods* **11**, 1033–1036 (2014).
33. Schonig, K. et al. Conditional gene expression systems in the transgenic rat brain. *BMC Biol.* **10**, 77 (2012).
34. Bryda, E. C. et al. A novel conditional ZsGreen-expressing transgenic reporter rat strain for validating Cre recombinase expression. *Sci. Rep.* **9**, 13330 (2019).
35. Wang, K. et al. Cre-dependent Cas9-expressing pigs enable efficient in vivo genome editing. *Genome Res.* **27**, 2061–2071 (2017).
36. Li, L. et al. Production of a reporter transgenic pig for monitoring Cre recombinase activity. *Biochem. Biophys. Res. Commun.* **382**, 232–235 (2009).
37. Gabriel, R. et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* **29**, 816–823 (2011).
38. Liang, P. et al. Genome-wide profiling of adenine base editor specificity by EndoV-seq. *Nat. Commun.* **10**, 67 (2019).
39. Kim, D. et al. Digenome-seq: genome-wide profiling of CRISPR–Cas9 off-target effects in human cells. *Nat. Methods* **12**, 237–243 (2015).
40. Kim, D. et al. Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat. Biotechnol.* **35**, 475–480 (2017).
41. Wienert, B. et al. Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science* **364**, 286–289 (2019).
42. Anderson, K. R. et al. CRISPR off-target analysis in genetically engineered rats and mice. *Nat. Methods* **15**, 512–514 (2018).
43. Iyer, V. et al. Off-target mutations are rare in Cas9-modified mice. *Nat. Methods* **12**, 479 (2015).
44. Smith, C. et al. Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell* **15**, 12–13 (2014).
45. Tang, X. et al. A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol.* **19**, 84 (2018).
46. Iyer, V. et al. No unexpected CRISPR–Cas9 off-target activity revealed by trio sequencing of gene-edited mice. *PLoS Genet.* **14**, e1007503 (2018).
47. Willi, M., Smith, H. E., Wang, C., Liu, C. & Hennighausen, L. Mutation frequency is not increased in CRISPR–Cas9-edited mice. *Nat. Methods* **15**, 756–758 (2018).

48. Bae, S., Park, J. & Kim, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).

Acknowledgements

We thank the FACS facility in ION. This work was supported by the R&D Program of China (2018YFC2000100 and 2017YFC1001302 to H.Y. and 2017YFC0908405 to W.W.), the CAS Strategic Priority Research Program (XDB32060000), the National Natural Science Foundation of China (31871502 and 31522037), the Shanghai Municipal Science and Technology Major Project (2018SHZDZX05), the Shanghai City Committee of Science and Technology Project (18411953700, 18JC1410100) and a National Institutes of Health P01 Center grant (P01HG00020527 to L.M.S.).

Author contributions

E.Z. designed and performed experiments. Y.S., W.W., H.S. and L.Y. performed data analysis. T.Y. performed PCR analysis. W.Y. performed mouse embryo transfer. H.Y., Y.L. and L.M.S. supervised the project and designed experiments. E.Z., Y.S. and W.W. wrote the paper.

Competing interests

L.M.S. has consulted for companies on CRISPR editing. The remaining authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41596-020-0361-1.

Correspondence and requests for materials should be addressed to L.M.S., Y.L. or H.Y.

Peer review information *Nature Protocols* thanks T. J. Cradick, Caixia Gao 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: 31 July 2019; Accepted: 21 May 2020;
Published online: 14 August 2020

Related links
Key references using this protocol
Zuo et al. *Science* **364**, 289–292 (2019): https://doi.org/10.1126/science.aav9973
Zuo et al. *bioRxiv* (2020): https://doi.org/10.1101/2020.02.07.939074

nature research

Corresponding author(s): Lars M. Steinmetz

Last updated by author(s): Apr 8, 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	Micropipette Puller (Sutter Instrumen, P97), Micropipette Microforge (Narishige, MF-900),FemtoJet microinjector (Eppendorf), Stereomicroscope (Olympus, cat. no. SZ61)
Data analysis	fastqQC (v0.11.3), Trimmomatic (v0.36), BWA (v0.7.12), Picard-tools (v2.3.0), GATK (v3.5), Lofreq (v2.1.2), Strelka (v2.7.1), Scalpel (v0.5.3), Annovar (version 2016-02-01), NCBI BLAST+ (v2.2.29)

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

The sequencing data were deposited in NCBI Sequence Read Archive (SRA) under project accession SRP119022 and http://www.biosino.org/node/project/detail/NODEP00371807OEP000195.

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

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	No statistical method was used to predetermine sample size. Sample size was chosen based on our previous works.
Data exclusions	No data was excluded.
Replication	All statistic data are presented as mean ± standard error of the mean. All samples represent a minimum of two replicates.
Randomization	The samples were not randomized. Different cell passages were used for different biological replicates.
Blinding	Blinding was not used.

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

Animals and other organisms

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

Laboratory animals	he following mouse strains were used in the manuscript: C57BL/6J mice: female, 3-4 week-old; Ai9 mice: male , 8-15week-old; ICR mice: females, 8 week-old;
Wild animals	No wild animals were involved in this study.
Field-collected samples	The study did not involve samples collected from field.
Ethics oversight	The use and care of animals complied with the guideline of the Biomedical Research Ethics Committee of Shanghai Institutes for Biological Science, Chinese Academy of Sciences.

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

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	To isolate cells, the prepared tissues were dissociated enzymatically in an incubation solution of 5 mL Trypsin-EDTA (0.05%) at 37°C for 30min. The digestion was stopped by adding 5 ml of DMEM medium with 10% Fetal Bovine Serum (FBS). Fetal tissues were then homogenized by passing 30-40 times through a 1ml pipette tips. The cell suspension was centrifuged for 6 min (800 rpm), and the pellet was resuspended in DMEM medium with 10% FBS. Finally, the cell suspension was filtered through a 40-µm cell strainer, and tdtomato+/tdtomato- cells were isolated by FACS. Samples were found to be >95% pure when assessed with a second round of flow cytometry and fluorescence microscopy analysis.
Instrument	Cell Sorter (Beckman, MoFlo XDP)
Software	FlowJoVX
Cell population abundance	Samples were found to be >95% pure when assessed with a second round of flow cytometry and fluorescence microscopy analysis.
Gating strategy	Positive boundaries were determined by E14.5 tdTomato+ embryos injected with Cre at zygote stage, and negative boundaries were determined by control E14.5 tdTomato- embryos that were not injected with Cre.

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