

Base editing generates substantial off-target single nucleotide variants

Erwei Zuo^{1,4†}, Yidi Sun^{2,5†}, Wu Wei^{2,3†}, Tanglong Yuan^{4†}, Wenqin Ying¹, Lars M. Steinmetz^{3,6,7*}, Yixue Li^{2,8,9*}, Hui Yang^{1,*}

¹Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

²Key Lab of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

³Stanford Genome Technology Center, Stanford University, Palo Alto, CA 94304, USA

⁴Center for Animal Genomics, Agricultural Genome Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518124, China

⁵University of Chinese Academy of Sciences, Shanghai 200031, China

⁶Department of Genetics, School of Medicine, Stanford University, Stanford, CA 94305, USA

⁷Genome Biology Unit, European Molecular Biology Laboratory (EMBL), 69117 Heidelberg, Germany

⁸Collaborative Innovation Center for Genetics and Development, Fudan University, Shanghai, P.R. China

⁹Shanghai Center for Bioinformation Technology, Shanghai Industrial Technology Institute, Shanghai, P. R. China

†These authors contributed equally to this work.

*Corresponding authors. E-mail: Lars.Steinmetz@stanford.edu, yxli@sibs.ac.cn or huiyang@ion.ac.cn

Abstract

Genome editing tools including CRISPR/Cas9 and base editors hold great promise for correcting pathogenic mutations. Unbiased genome-wide off-target effects of the editing in mammalian cells is required before clinical applications, but determination of the extent of off-target effects has been difficult due to the existence of single nucleotide polymorphisms (SNPs) in individuals. Here, we developed a method named GOTI (Genome-wide Off-target analysis by Two-cell embryo Injection) to detect off-target mutations without interference of SNPs. We applied GOTI to both the CRISPR-Cas9 and base editing (BE3) systems by editing one blastomere of the two-cell mouse embryo and then compared whole genome sequences of progeny-cell populations at E14.5 stage. Sequence analysis of edited and non-edited cell progenies showed that undesired off-target single nucleotide variants (SNVs) are rare (average 10.5) in CRISPR-edited mouse embryos, with a frequency close to the spontaneous mutation rate. By contrast, BE3 editing induced over 20-fold higher SNVs (average 283), raising the concern of using base-editing approaches for biomedical application.

CRISPR/Cas9 and base editors mediated genome editing

methods have been developed and carried great hope for the treatment of genetic diseases induced by pathogenic mutations (1) (2) (3) (4) (5) (6). Clinical applications of CRISPR/Cas9-based genome editing or base editors require a comprehensive analysis of off-target effects to minimize risk of deleterious mutations, but a well validated method robust to genetic variants is undescribed yet to detect SNVs (7). Multiple methods have been developed to detect genome-wide gene-editing off-target activity in cells, including high-throughput genome-wide translocation sequencing (HTGTS) (8), genome-wide, unbiased identification of double-strand breaks (GUIDE-seq) (9) and circularization for in vitro reporting of cleavage effects by sequencing (CIRCLE-seq) (10). However, these approaches are not applicable to detect SNVs. Thereafter, GOT1 was developed and applied to evaluate the off-target effects induced by CRISPR/Cas9 or BE3 (rAPOBEC1-nCas9-UGI), the commonly used gene editing tools (4, 11). Briefly, we performed either CRISPR/Cas9 or BE3 gene editing, together with Cre mRNA, in one blastomere of the two-cell embryos derived from Ai9 (CAG-LoxP-Stop-LoxP-tdTomato) mice (12, 13) (Fig. 1a). The progeny cells of edited and non-edited blastomeres were then sorted at E14.5 by fluorescence-activated cell sorting (FACS),

based on tdTomato expression in the gene-edited cells. Whole genome sequencing (WGS) was then performed on the tdTomato⁺ and tdTomato⁻ cells, separately (Fig. 1a).

To validate our approach, we demonstrated that edited cells treated with Cre and Cas9/BE3 system was efficiently separated with non-edited cells. Upon Cre-mediated recombination, about 50% of the embryo cells are expected to express tdTomato. This was verified at the 4- and 8-cell stages by fluorescence microscopic observation (fig. S1a, b) or at E14.5 by FACS analysis (Fig. 1b). Furthermore, we found that editing the coat-color gene tyrosinase (*Tyr*) with CRISPR/Cas9 by injection of one blastomere of 2-cell embryos with either one of the two sgRNAs (Cas9-Tyr-A, Cas9-Tyr-B) resulted in highly efficient on-target editing: Sanger sequencing of *Tyr* gene showed that 13 (for Cas9-Tyr-A) and 15 (for Cas9-Tyr-B) tdTomato⁺ cells individually collected from four dispersed 8-cell embryos, 85% and 80% of cells carrying *Tyr* mutated alleles, respectively. By contrast, none of tdTomato⁻ cells collected (16 for Cas9-Tyr-A and 15 for Cas9-Tyr-B) showed *Tyr* mutated alleles (fig. S1c).

In addition, we verified the editing efficiency of our approach on the targeted *Tyr* gene. For genome-wide sequencing studies using

the embryo injection protocol, we included four sgRNAs for CRISPR/Cas9 editing: Cas9-Tyr-A and Cas9-Tyr-B targeting at *Tyr*, one control sgRNA targeting at *LacZ* without cleavage sites in the C57 mouse genome, and a sgRNA targeting at *Pde6b*, which has one mismatch with the C57 mouse genome and was previously reported to generate thousands of SNVs (14). We validated the cleavage efficiency of these sgRNAs (~100%) *in vitro* using DNA cleavage assays (fig. S2). In addition, we examined two sgRNAs (BE3-Tyr-C and BE3-Tyr-D) that targeted *Tyr* gene via BE3-mediated editing. As control groups for gene editing, we also included three groups of embryos injected with “Cre only”, “Cre and Cas9” and “Cre and BE3”. After injection of different mixtures of CRISPR/Cas9 or BE3, and Cre mRNAs as well as sgRNAs into one blastomere of two-cell embryos, we found no impairment of embryonic development, as indicated by the normal blastocyst rate and survival rate (fig. S3a-b). By Sanger sequencing, all the examined blastocysts and E14.5 fetuses derived from Cas9-Tyr-A, Cas9-Tyr-B, BE3-Tyr-C and BE3-Tyr-D carried *Tyr* mutations (fig. S3c). To evaluate target-editing efficiencies for the *Tyr* gene, we FACS-sorted about four million tdTomato⁺ and four million tdTomato⁻ cells from one E14.5 embryo treated with either Cas9-Tyr-A or Cas9-Tyr-B, and performed TA

clone sequencing. The results from one experiment showed that tdTomato⁺ cells carried 50% and 100% mutated alleles (with 12 and 26 clones were examined) for Cas9-Tyr-A and Cas9-Tyr-B targeting, respectively (Fig. 1c). In the second repeated experiments, tdTomato⁺ cells carried 58% and 50% mutated alleles (with 31 and 14 clones examined) respectively for Cas9-Tyr-A and Cas9-Tyr-B targeting (Fig. 1c). By contrast, tdTomato⁻ cells collected from 14.5 embryos carried no *Tyr* mutated alleles (Fig. 1c). Similarly, BE3 editing demonstrated high targeting efficiencies, with 71% mutant alleles on BE3-Tyr-C and 100% on BE3-Tyr-D, and the corresponding tdTomato⁻ cells had only ~3% *Tyr* mutations (Fig. 1c). These results suggested that both CRISPR/Cas9 and BE3 editing yield high on-target editing efficiency in tdTomato⁺ cells, but essentially no on-target editing in tdTomato⁻ cells.

To further explore the on-target efficiency and potential genome-wide off-target effects, we performed WGS at an average depth of 47x on 36 samples from 18 E14.5 embryos in nine groups (with 2 embryos each): Cre only, Cre and Cas9, Cre and Cas9-LacZ, Cre and Cas9-Pde6b, Cre and Cas9-Tyr-A, Cre and Cas9-Tyr-B, Cre and BE3, Cre and BE3-Tyr-C, and Cre and BE3-Tyr-D (Omit “Cre” in the latter eight groups for short), among which only

Cas9-Tyr-A, Cas9-Tyr-B, BE3-Tyr-C and BE3-Tyr-D groups have target sites on the C57 genome (table S1). On-target analysis for Cas9-Tyr-A and Cas9-Tyr-B groups showed an average of 56% and 72% *Tyr* mutated alleles in tdTomato⁺ cells, respectively (Fig. 1d), indicating efficient on-target DNA editing at *Tyr* gene locus. Similarly, both BE3-Tyr-C and BE3-Tyr-D demonstrated high editing efficiency (averaged 75% and 92% respectively *Tyr* mutated alleles) in tdTomato⁺ cells (Fig. 1d). We also analyzed the on-target efficiency for all the other tdTomato⁺ embryos, and found zero on-target editing for all the control embryos (Fig. 1d). However, the WGS analysis also revealed a low-level of targeted editing in tdTomato⁻ cells of Cas9-Tyr-A and Cas9-Tyr-B groups in the range of 0 - 6.3% (Fig. 1d, fig. S4 and table S2), which was probably caused by false-negative FACS sorting (known to occur at a low level). Thus we only considered variants with allele frequencies more than 10% to be reliable in our following analysis.

To assess the off-target editing effects, we analyzed the genome-wide *de novo* variants by comparing the tdTomato⁺ cells with the tdTomato⁻ cells in each embryo with three different variant calling algorithms simultaneously (14, 15), with variants defined by all three algorithms to be the true variants (Fig. 2a, fig. S5, table S3 and Methods). We found only 0 - 4 indels in embryos from all

nine groups, a result further validated by Sanger sequencing (Table 1, fig. S6a and table S3). Meanwhile, in the “Cre only” embryos, we observed an average of 14 SNVs (Table 1 and table S4). For the CRISPR/Cas9-treated embryos (Cas9, Cas9-LacZ, Cas9-Pde6b, Cas9-Tyr-A, and Cas9-Tyr-B), there were an average of 12.5, 5, 0, 16 and 19 SNVs, respectively in each group (Table 1 and table S4), and no significant difference among them (Fig. 2a) or in comparison with the “Cre-only” group (Fig. 2a-b). In addition, we observed no increased SNVs in the Cas9-Pde6b edited embryo, consistent with many previous studies (16-19). All off-target SNVs detected in the CRISPR/Cas9 edited embryos were further confirmed by Sanger sequencing (fig. S6b). The SNV variants detected in the Cre- or CRISPR/Cas9- treated samples were likely caused by spontaneous mutations during genome replication in the development, since the number of variants was within the range of the simulated spontaneous mutations (fig. S7, Methods) (20). Besides, none of the samples in our study shared identical variants (fig. S8a), and no overlap was found for SNV sites predicted by Cas-OFFinder, CRISPOR and Digenome-seq putative off-target sites (fig. S8b and tables S5-7, Methods) (21, 22). We also observed no sequence similarity between the adjacent sequences of the identified SNVs with the target sites (fig.

S8c). In addition, by conversely calling variants through comparing the tdTomato⁻ sample with tdTomato⁺ sample in each embryo, we found similar numbers of SNVs (fig. S9a and table S8), suggesting that CRISPR/Cas9 editing induces no off-target edits and the SNVs we observed are from spontaneous mutations.

Surprisingly, in contrast with the CRISPR/Cas9 editing, we found an average of 283 SNVs in all BE3-treated embryos with or without sgRNA (BE3, BE3-Tyr-C, BE3-Tyr-D), a level of SNVs at least 20 times more than that observed in Cre-injected or CRISPR/Cas9-treated embryos (Figure 2a, b, Table 1 and fig. S5b). By conversely calling variants in each embryo (tdTomato⁻ vs tdTomato⁺), only 20 SNVs on average were detected (fig. S9b and table S9), which was presumed to be the spontaneous mutations considering the analysis on the CRISPR/Cas9-treated samples. These results indicated that the increased SNVs identified were caused by the injection of BE3.

Notably, for the off-targets detected in BE3-treated samples, none of them were shared by any of the embryos in each group (Fig. 2c), and they were randomly distributed across the genome (Fig. 2d, e, table S10). We then compared these off-target mutations with all the potential off-target sites predicted by both

Cas-OFFinder and CRISPROR softwares (tables S11 and S12) (21, 22). Not surprisingly, the two prediction tools shared a large amount of off-target sites, but none of them were found in our detected SNV sites (Fig. 2f). Besides, no sequence similarity was observed between adjacent sequences of the identified SNVs and the BE3 sgRNA targeting sites, whereas the top predicted off-target sites showed sequence similarity with BE3 on-target loci (Fig. 2g). Remarkably, in spite of the uniqueness of the SNVs generated by BE3-editing, the mutation types were the same as those preferred by APOBEC1 (23, 24) (Fig. 2h and fig. S10). In fact, more than 90% of the SNVs identified in the BE3-edited cells were mutated from G to A or C to T, which was not observed in Cre- or CRISPR/Cas9- treated cells (Fig. 2h and fig. S10), indicating that mutations were not spontaneous but induced by BE3-editing. Statistical significance was also observed in the proportion of G>A and C>T among all mutations between BE3-treated cells and Cas9- ($P = 6.7 \times 10^{-4}$) or Cre- ($P = 3.6 \times 10^{-2}$) treated cells (Fig. 2i). It has been reported that DNA accessibility is related to gene editing efficiency (25). Thus we evaluated whether the identified off-target sites enriched in open chromatin regions by checking the ATAC-seq datasets derived from embryonic mouse cells in the Cistrome database (26-28). In fact, off-target sites were

significantly enriched in the regions with higher accessibilities in a dataset from E8.5 embryos with a mixed C57BL6/DBA2 background (Fig. 2j) and other four high quality datasets in Cistrome database. (fig. S11).

Among 1698 SNVs induced by BE3 editing in six BE3-treated embryos, 26 were located on exons (tables S13). We successfully amplified 20 of them by PCR and confirmed their presence by Sanger sequencing (Fig. 3a, fig. S12 and tables S13). Among the 26 SNVs, 14 of them lead to nonsynonymous changes in the coded proteins, and two SNVs result in early stop-codons in *Trim23* and *Aim2* genes. The *Trim23* encodes an E3 ubiquitin-protein ligase, and its dysfunction could cause muscular dystrophy (29). Previous studies reported an essential role of *Aim2* gene in innate immunity underlying the defense against viral infection (30, 31) (Fig. 3b). We also found 1 SNV in a proto-oncogene and 13 SNVs in tumor suppressors (Fig. 3c), raising serious concern about the oncogenic risk of BE3-editing.

One of the big advantages of our approach is that rules out genetic background differences by comparing edited to non-edited cells in a single mouse. Previous studies were confounded by genetic background variations when comparing edited and non-

edited mice. In fact, we also applied our analysis method to a published dataset (15) and found an average of ~1000 SNVs and ~100 indels between CRISPR/Cas9-edited mice and their non-edited siblings (tables S14). Based on our finding, we suggest that these differences among siblings are due to genetic variations rather than the results of CRISPR/Cas9-editing. Furthermore, when the sequences between any two different embryos in our study were compared, we observed more SNVs (3706 ± 5232) and indels (583 ± 762) ($n = 18$ pairs), as the embryos used in our study were not all derived from the same parents (table S15). These results illustrate that it is difficult to find a robust control in off-target analysis by comparing an edited to non-edited mouse as there is a significant amount of genetic variations between mice even when they have the same parents.

In summary, our studies demonstrate the advantages of GOT1 approach using whole-genome sequencing analysis of progeny cells of sister blastomeres in studying off-target effects induced by gene-editing. We also showed that undesired off-target mutations induced by CRISPR/Cas9-mediated gene editing are rare in mouse embryos, supporting previous findings that CRISPR/Cas9-based editing *in vivo* does not cause significant SNVs and indels (15, 32-34). However, we could not exclude the big deletions or

large chromosomal translocations reported by other studies (35, 36). By contrast, we found numerous *de novo* SNVs induced by BE3-editing, raising the safety concern of base editing approaches in therapeutic applications.

AUTHOR CONTRIBUTIONS

EZ designed and performed experiments. YS and WW performed WGS analysis. TY performed PCR analysis. WY performed mouse embryo transfer. HY, YL and LS supervised the project and designed experiments. All the authors wrote the paper.

ACKNOWLEDGMENTS

We thank Dr. Mu-ming Poo, Dang-sheng Li, Guo-liang Xu and Kevin Roy for helpful discussions and insightful comments on this manuscript. This work was supported by National Science and Technology Major Project (2017YFC1001302), CAS Strategic Priority Research Program (XDB02050007, XDA01010409), the National High-tech R&D Program (863 Program; 2015AA020307), the National Natural Science Foundation of China grants (31522037, 31500825), China Youth Thousand Talents Program (to HY), Break-through project of Chinese Academy of Sciences, Shanghai City Committee of science and technology project (16JC1420202, HY) and National Key Research and Development Program of China

(2017YFC0908405) to WW.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

REFERENCES

1. K. Musunuru, The Hope and Hype of CRISPR-Cas9 Genome Editing A Review. *Jama Cardiol* **2**, 914-919 (2017).
2. C. Fellmann, B. C. Cowen, P. C. Lin, J. A. Doudna, J. E. Corn, Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nat Rev Drug Discov* **16**, 89-100 (2017).
3. T. Koo, J. S. Kim, Therapeutic applications of CRISPR RNA-guided genome editing. *Brief Funct Genomics* **16**, 38-45 (2017).
4. T. I. Cornu, C. Mussolino, T. Cathomen, Refining strategies to translate genome editing to the clinic. *Nature Medicine* **23**, 415-423 (2017).
5. C. E. Dunbar *et al.*, Gene therapy comes of age. *Science* **359**, 175-+ (2018).
6. H. Seo, J. S. Kim, Towards therapeutic base editing. *Nat Med* **24**, 1493-1495 (2018).
7. S. Q. Tsai, J. K. Joung, Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat Rev Genet* **17**, 300-312 (2016).
8. R. L. Frock *et al.*, Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol* **33**, 179-186 (2015).
9. S. Q. Tsai *et al.*, GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* **33**, 187-197 (2015).
10. S. Q. Tsai *et al.*, CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Methods* **14**, 607-614 (2017).
11. H. A. Rees, D. R. Liu, Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet*, (2018).
12. L. Madisen *et al.*, A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* **13**, 133-140 (2010).
13. L. Wang *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).
14. K. A. Schaefer *et al.*, Unexpected mutations after CRISPR-Cas9 editing in vivo. *Nat Methods* **14**, 547-548 (2017).
15. V. Iyer *et al.*, Off-target mutations are rare in Cas9-modified mice. *Nat Methods* **12**, 479 (2015).
16. S.-T. Kim *et al.*, Response to "Unexpected mutations after CRISPR-Cas9 editing in vivo" . *Nature Methods* **15**, 239 (2018).
17. C. A. Lareau *et al.*, Response to "Unexpected mutations after CRISPR-Cas9 editing in vivo" . *Nature Methods* **15**, 238 (2018).

18. R. M. Lescarbeau, B. Murray, T. M. Barnes, N. Bermingham, Response to “Unexpected mutations after CRISPR–Cas9 editing in vivo” . *Nature Methods* **15**, 237 (2018).
19. L. M. J. Nutter *et al.*, Response to “Unexpected mutations after CRISPR–Cas9 editing in vivo” . *Nature Methods* **15**, 235 (2018).
20. J. W. Drake, B. Charlesworth, D. Charlesworth, J. F. Crow, Rates of spontaneous mutation. *Genetics* **148**, 1667–1686 (1998).
21. S. Bae, J. Park, J. S. Kim, Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).
22. M. Haeussler *et al.*, Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* **17**, 148 (2016).
23. B. Teng, C. F. Burant, N. O. Davidson, Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science* **260**, 1816–1819 (1993).
24. A. Chester, J. Scott, S. Anant, N. Navaratnam, RNA editing: cytidine to uridine conversion in apolipoprotein B mRNA. *Biochim Biophys Acta* **1494**, 1–13 (2000).
25. R. M. Yarrington, S. Verma, S. Schwartz, J. K. Trautman, D. Carroll, Nucleosomes inhibit target cleavage by CRISPR–Cas9 in vivo. *Proceedings of the National Academy of Sciences* **115**, 9351–9358 (2018).
26. M. Radman-Livaja, O. J. Rando, Nucleosome positioning: how is it established, and why does it matter? *Dev Biol* **339**, 258–266 (2010).
27. M. Tsompana, M. J. Buck, Chromatin accessibility: a window into the genome. *Epigenetics Chromatin* **7**, 33 (2014).
28. S. Mei *et al.*, Cistrome Data Browser: a data portal for ChIP–Seq and chromatin accessibility data in human and mouse. *Nucleic Acids Res* **45**, D658–D662 (2017).
29. E. Kudryashova, J. Wu, L. A. Havton, M. J. Spencer, Deficiency of the E3 ubiquitin ligase TRIM32 in mice leads to a myopathy with a neurogenic component. *Hum Mol Genet* **18**, 1353–1367 (2009).
30. V. A. Rathinam *et al.*, The AIM2 inflammasome is essential for host defense against cytosolic bacteria and DNA viruses. *Nat Immunol* **11**, 395–402 (2010).
31. A. M. Krieg, AIMing 2 defend against intracellular pathogens. *Nat Immunol* **11**, 367–369 (2010).
32. E. Zuo *et al.*, One-step generation of complete gene knockout mice and monkeys by CRISPR/Cas9-mediated gene editing with multiple sgRNAs. *Cell Res*, (2017).
33. P. Akcakaya *et al.*, In vivo CRISPR editing with no detectable genome-wide off-target mutations. *Nature* **561**, 416–419 (2018).
34. M. Willi, H. E. Smith, C. Wang, C. Liu, L. Hennighausen, Mutation frequency is not increased in CRISPR–Cas9-edited mice. *Nat Methods* **15**, 756–758 (2018).
35. M. Kosicki, K. Tomberg, A. Bradley, Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol* **36**, 765–771 (2018).
36. H. Y. Shin *et al.*, CRISPR/Cas9 targeting events cause complex deletions and insertions at 17 sites in the mouse genome. *Nat Commun* **8**, 15464 (2017).
37. K. Cibulskis *et al.*, Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**, 213–219 (2013).

38. A. Wilm *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).
39. C. T. Saunders *et al.*, Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**, 1811-1817 (2012).
40. H. Fang *et al.*, Indel variant analysis of short-read sequencing data with Scalpel. *Nat Protoc* **11**, 2529-2548 (2016).
41. K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).

Figures and figure legends

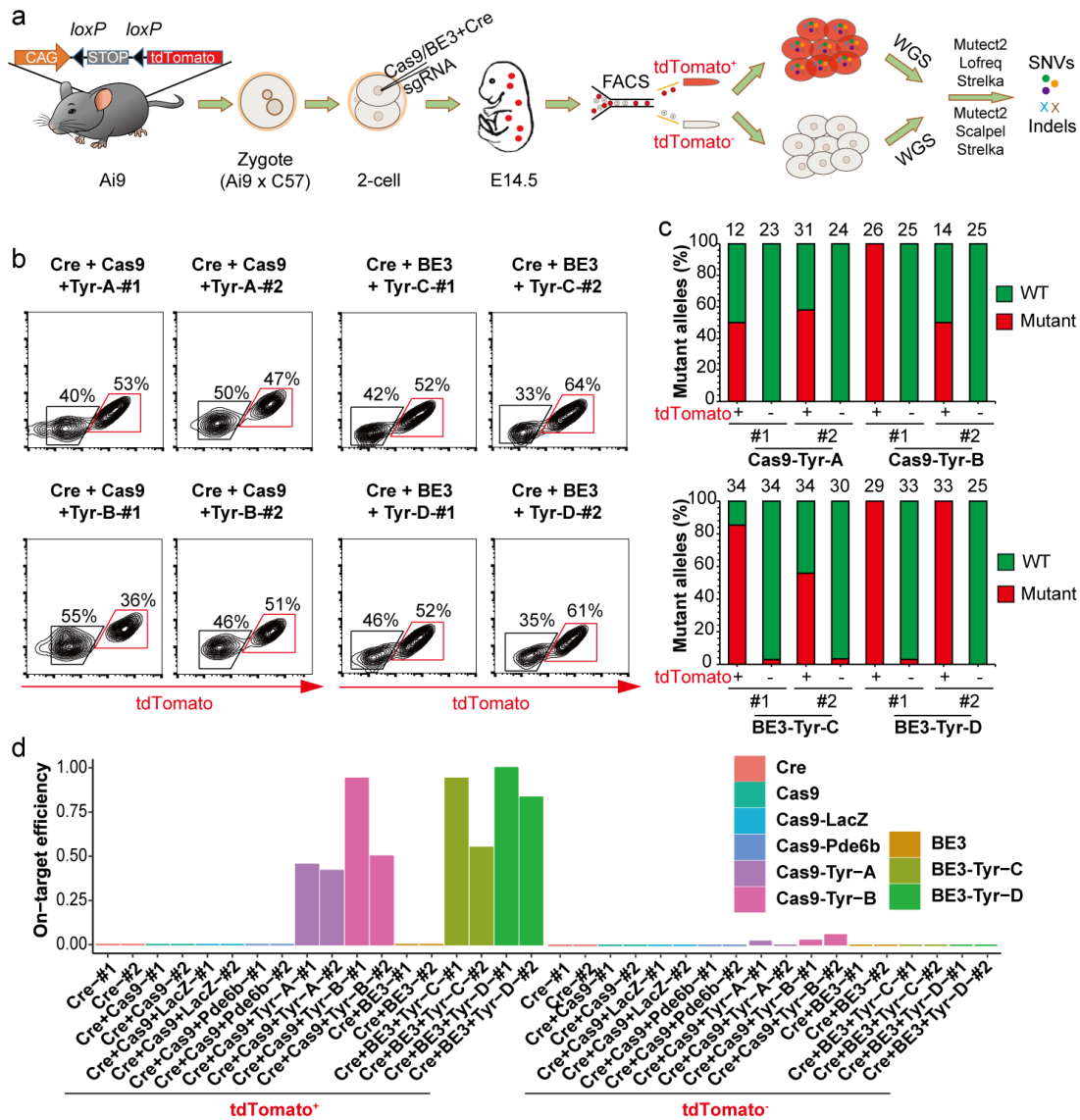


Figure 1. CRISPR/Cas9- or BE3-mediated gene editing in one blastomere of 2-cell embryos. **a**, Experimental design. The mixture of Cre, Cas9/BE3 and sgRNA was injected into one blastomere of a 2-cell embryo, derived from Ai9 male mice mating with wild-type female mice. The action of Cre is expected to generate a chimeric embryo with half of cells labeled with tdTomato (colored red). The tdTomato⁺ cells and tdTomato⁻ cells were isolated from the chimeric embryo at E14.5 by FACS and used for WGS respectively. The off-target SNVs and indels were identified by comparing the tdTomato⁺ cells with tdTomato⁻ cells using three variant calling algorithms as indicated (Mutect2, Lofreq and Strelka for SNVs, and Mutect2, Scalpel and Strelka for indels). SNVs and indels were indicated as colored dots and crosses. **b**, FACS analysis on E14.5 embryos treated with Cas9-Tyr-A, Cas9-Tyr-B, BE3-Tyr-C or BE3-Tyr-D. FACS analysis of uninjected embryo was shown in fig. S1b. **c**, On-target analysis of TA clone sequencing on E14.5 embryos treated with Cas9-Tyr-A, Cas9-Tyr-B, BE3-Tyr-C or BE3-Tyr-D. Number above each column, total clones analyzed. **d**, On-target efficiency for the tdTomato⁺ cells (left) and tdTomato⁻ cells (right) included in the study based

on WGS. Embryos with the same treatment were represented in the same color.

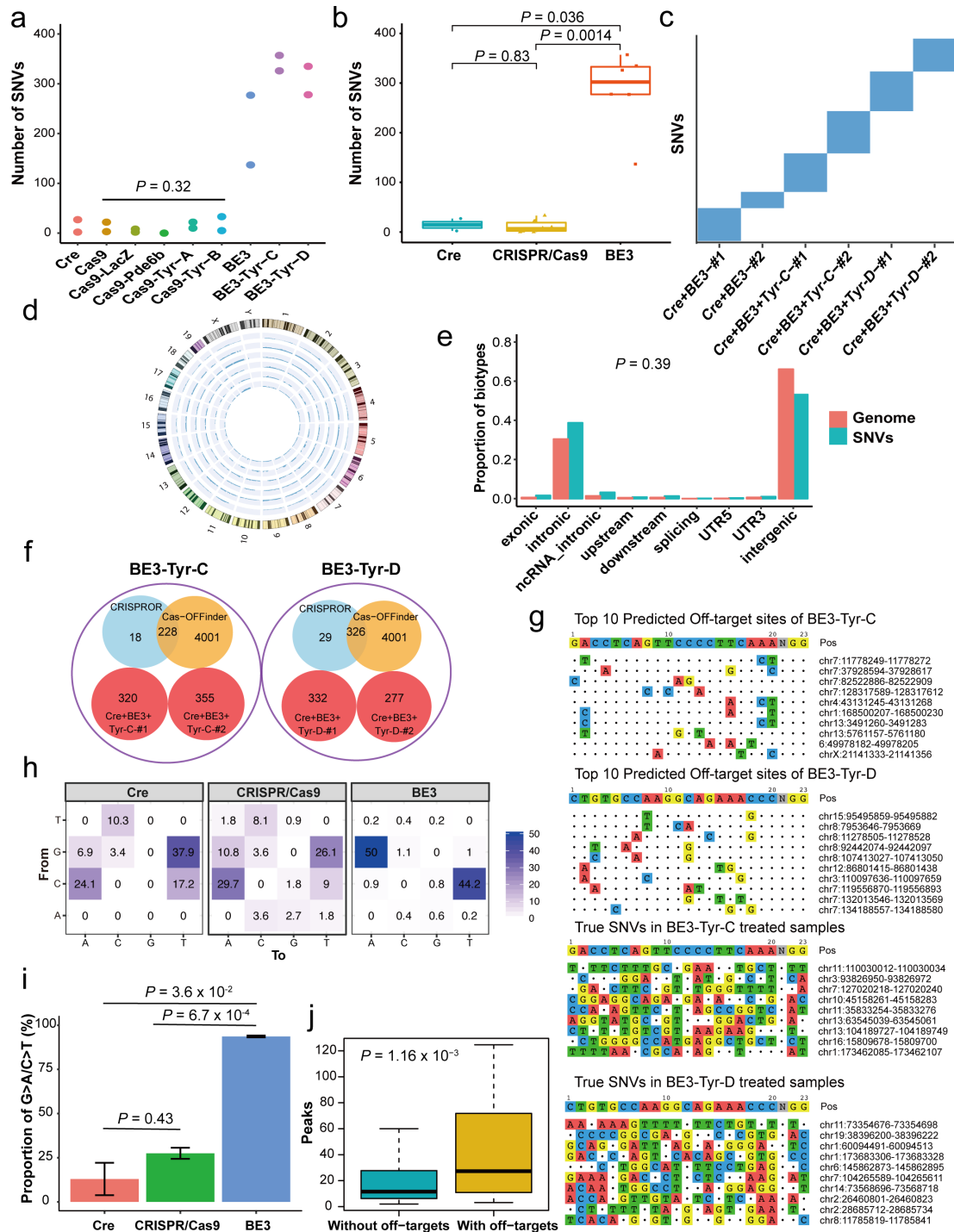


Figure 2. Genome-wide SNVs identified in Cre-, CRISPR/Cas9- and BE3-treated mouse embryos. **a**, The number of SNVs identified in Cre-treated, CRISPR/Cas9-treated (Cas9, Cas9-LacZ, Cas9-Pde6b, Cas9-Tyr-A, and Cas9-Tyr-B) and BE3-treated (BE3, BE3-Tyr-C, BE3-Tyr-D) groups. Each point indicated one embryo. Colors indicated different treatment conditions. *P*-value was calculated by Kruskal-Wallis test among CRISPR/Cas9-treated embryos. **b**, The comparison of the total number of detected SNVs in Cre-treated only, CRISPR/Cas9-treated and BE3-treated samples. **c**, The SNVs identified from each embryo with BE3 injected were mutually exclusive from each other. **d**, The detected SNVs were evenly distributed in the mouse genome in all the six BE3-treated samples. Embryos from inner circle

to outer circle were Cre+BE3-#1, Cre+BE3-#2, Cre+BE3-Tyr-C-#1, Cre+BE3+Tyr-C-#2, Cre+BE3+Tyr-D-#1 and Cre+BE3+Tyr-D-#2, respectively. **e**, The biotypes of our detected SNVs were similar to the structure of mm10 genomes ($P = 0.39$ by Wilcoxon rank sum test). **f**, The overlap among SNVs detected from our results with predicted off-targets sites by Cas-OFFinder and CRISPOR. **g**, The sequence similarities between the targeted sequence with predicted off-target sites or our identified mutation sequences. **h**, The distribution of mutation types in Cre-, CRISPR/Cas9- and BE3-treated groups. The number in each cell indicated the proportion of a certain type of mutation among all the mutations. Deeper colors represented higher proportion of mutation types. **i**, The proportion of C>T and G>A mutations in the Cre only, CRISPR/Cas9- and BE3-treated embryos. P values were calculated by two-sided Wilcoxon rank sum test and $P < 0.05$ was considered as significant. **j**, The comparison of peak regions with or without the identified off-targets from a ATAC-seq dataset with GEO accession GSM2551664. P -values were calculated with Wilcoxon rank sum test.

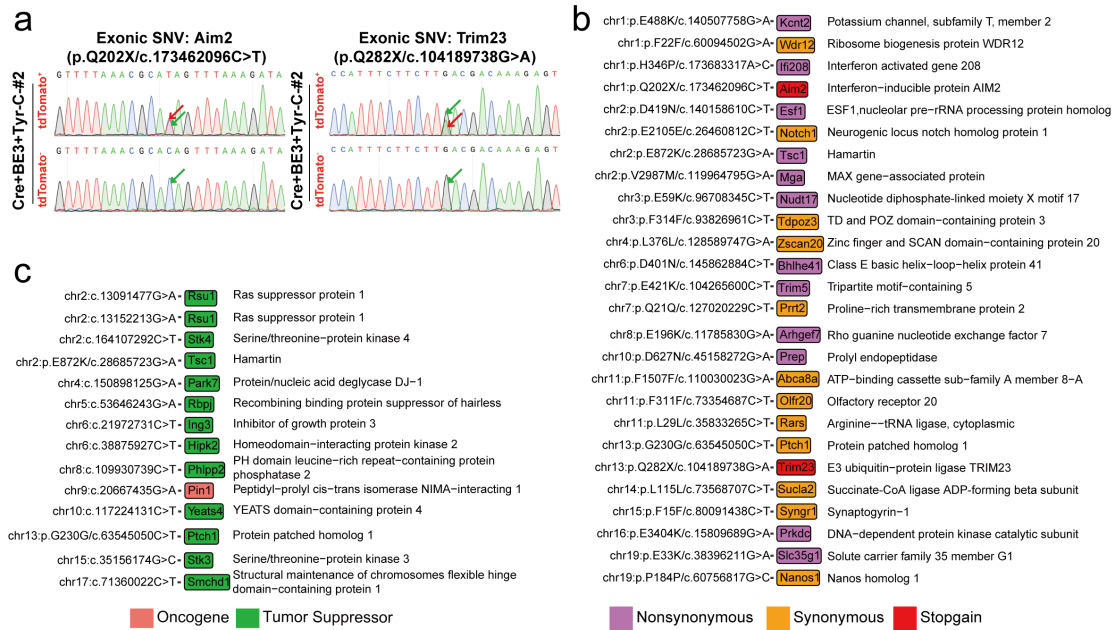


Figure 3. The BE3 induced off-targets could be detrimental to gene functions. a, Sanger sequencing validation on exonic SNVs located on *Aim2* and *Trim23* genes. The mutated and WT base sequence were pointed out by red and green arrows, respectively. **b**, SNVs located on the protein coding exons. Nonsynonymous, synonymous and stop-gain SNVs were indicated by different colors. **c**, The SNVs located within the cancer associated genes. Red and green colors indicated oncogene and tumor suppressor, respectively. Primers were listed in table S16.

Table 1. Summary of SNVs and indels identified from WGS in each embryo.

Variants	Cre +																	
	Cas9+								BE3+									
	#1	#2	#1	#2	LacZ- #1	LacZ- #2	Pde6b- #1	Pde6b- #2	Tyr-A #1	Tyr-A #2	Tyr-B #1	Tyr-B #2	#1	#2	Tyr-C #1	Tyr-C #2	Tyr-D #1	Tyr-D #2
WGS SNVs	2	26	22	3	8	2	0	0	22	10	5	33	277	137	320	355	332	277
WGS Indels	0	3	0	1	0	1	0	0	0	0	2	0	1	0	1	4	1	0
Exon SNVs	0	0	0	0	0	0	0	0	1	0	0	4	3	4	3	6	6	4
Exon Indels	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nonsynonymous SNVs	0	0	0	0	0	0	0	0	0	0	0	2	2	2	0	4	4	2
Frameshift Indels	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Predicted off-target sites							0	0	0	0	0	0						
Putative Digenome sites							0	0	0	0	0	0						

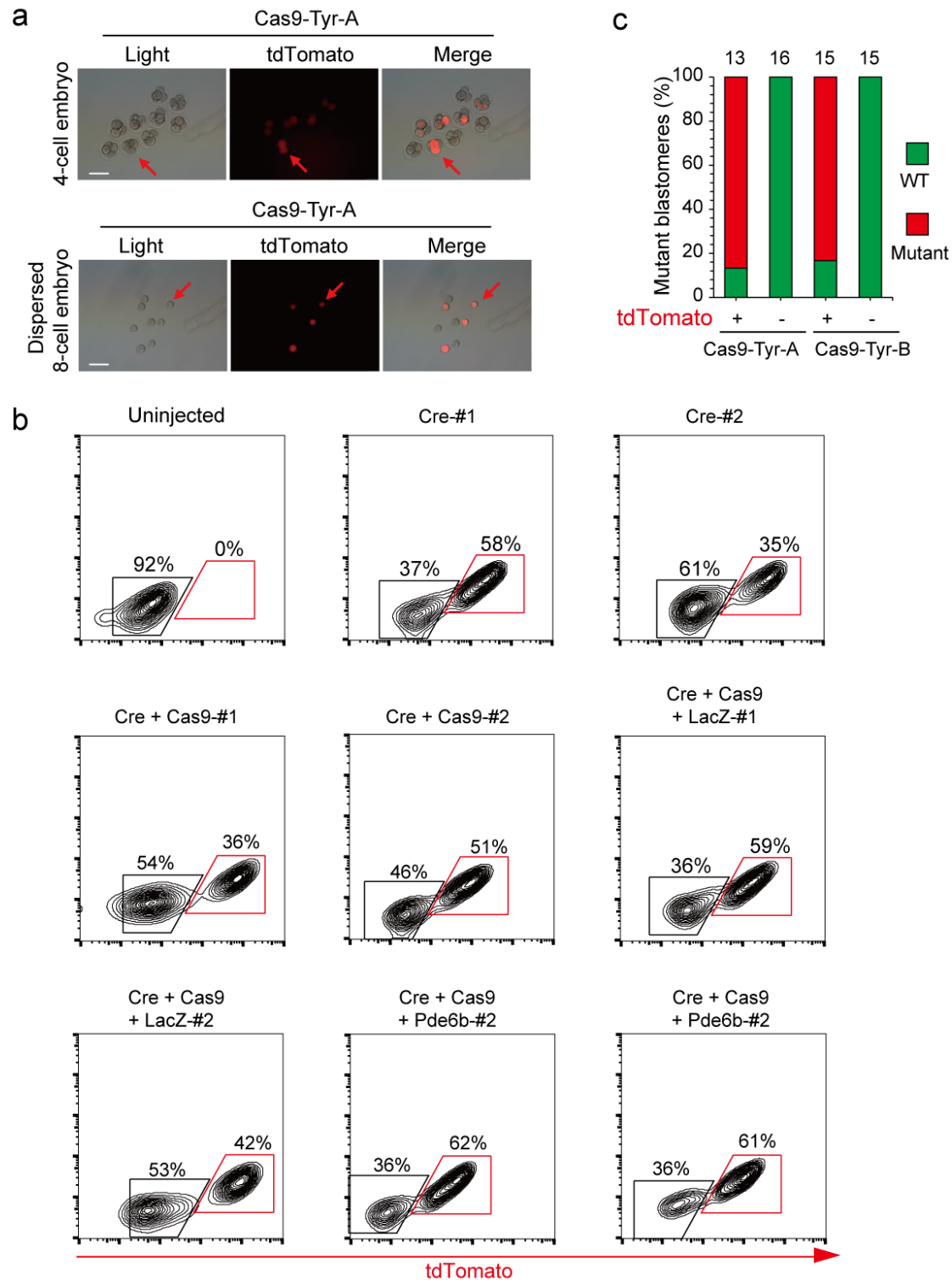


Figure S1. FACS analysis on E14.5 embryos treated with different mixtures. a, Representative images of Cas9-Tyr-A gene-targeted embryos. Top, 4-cell embryos; bottom, a dispersed 8-cell embryo. Red arrows indicate tdTomato⁺ blastomeres. Scale bar, 100µm. **b,** From left to right and top to the bottom successively: uninjected, Cre-#1, Cre-#2, Cre+Cas9-#1, Cre+Cas9-#2, Cre+Cas9+LacZ-#1, Cre+Cas9+LacZ-#2, Cre+Cas9+Pde6b-#1 and Cre+Cas9+Pde6b-#2. **c,** Genotypes of *Tyr* gene-targeted 8 cell embryos. Individual tdTomato⁺ and tdTomato⁻ blastomeres were isolated from four Cas9-Tyr-A and four Cas9-Tyr-B gene-

targeted 8-cell embryos. Number, total blastomeres analyzed. WT, wild-type allele. Mutant, *Tyr* mutated allele.

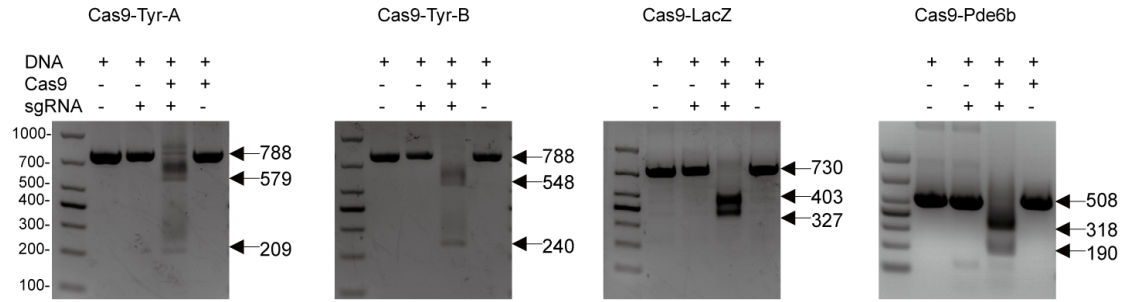


Figure S2. The cleavage efficiencies of sgRNAs using in vitro cleavage of DNA assay.

Agarose gel electrophoresis from left to right showed the result of Cas9-Tyr-A, Cas9-Tyr-B, Cas9-LacZ and Cas9-Pde6b, respectively. The genomic region or construct flanking the sgRNA target site for each gene was PCR amplified and PCR products were incubated with Cas9 ribonucleoproteins and sgRNA for 3 h.

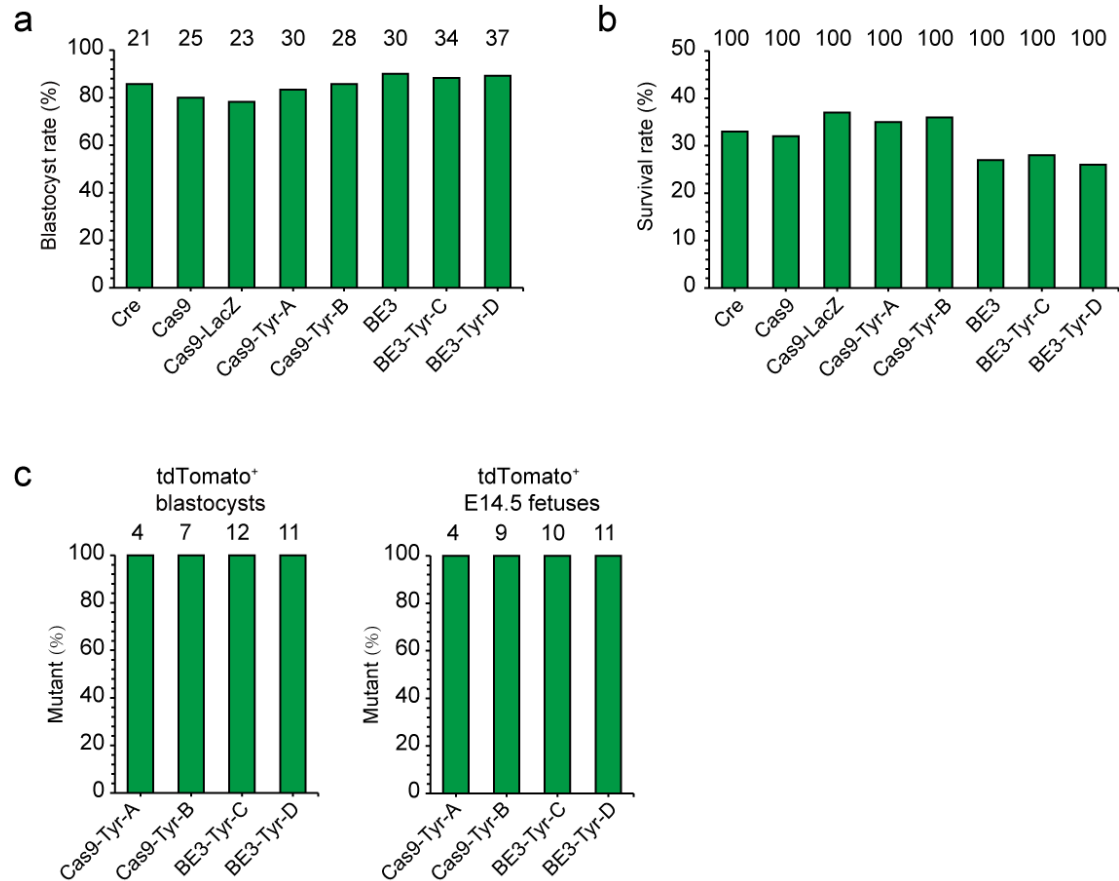


Figure S3. Development and genotyping of CRISPR/Cas9 and BE3-treated chimeric embryos. **a**, Percentages of tdTomato⁺ blastocysts from embryos injected with different mixtures. Number, total blastocysts or embryos counted. **b**, Survival rate at E14.5 of embryos injected with different mixtures. Number, total embryos counted or transferred. **c**, The percentage of tdTomato⁺ blastocysts and E14.5 embryos with *Tyr* mutations. Number, total blastocysts or embryos analyzed.

			tdTomato ⁺					tdTomato ⁻	
Cas9-Tyr-A	#1	GCGAAGGCACCGCCCTCTTTTGG	WT (18/33)		Cas9-Tyr-A	#1	GCGAAGGCACCGCCCTCTTTTGG	WT (37/38)	
		GCGAAGGCACCG-----	-14bp (15/33)				GCGAAGGCACCG-----	-14bp (1/38)	
Cas9-Tyr-B	#2	GCGAAGGCACCGCCCTC---TTTTGG	WT (18/31)		Cas9-Tyr-B	#2	GCGAAGGCACCGCCCTCTTTTGG	WT (19/19)	
		GCGAAGGCACCGCCCTCATGATTTTGG	+4bp (13/31)						
Cas9-Tyr-B	#1	CCCAGAAGCCAATGCACCTATCGG	WT (2/34)		Cas9-Tyr-B	#1	CCCAGAAGCCAATGCACCTATCGG	WT (31/32)	
		CCCAGAAGCCAATGCA-----CGG	-5bp (12/34)				CCCAGAAGCCAATGCA-----CGG	-5bp (1/32)	
Cas9-Tyr-B	#2	CCCAGAAGCCAATGCACCTATCGG	WT (11/22)		Cas9-Tyr-B	#2	CCCAGAAGCCAATGCACCTATCGG	WT (30/32)	
		CCCAGAAGCCAAT---CCTATCGG	-3bp (11/22)				CCCAGAAGCCAAT---CCTATCGG	-3bp (2/32)	
BE3-Tyr-C	#1	GACCTCAGTTCCCTTCAAAGGG	WT (2/34)		BE3-Tyr-C	#1	GACCTCAGTTCCCTTCAAAGGG	WT (40/40)	
		GACCTTAGTTCCCTTCAAAGGG	MUT (32/34)						
BE3-Tyr-C	#2	GACCTCAGTTCCCTTCAAAGGG	WT (18/40)		BE3-Tyr-C	#2	GACCTCAGTTCCCTTCAAAGGG	WT (50/50)	
		GACCTTAGTTCCCTTCAAAGGG	MUT (22/40)						
BE3-Tyr-D	#1	CTGTGCCAAGGCAGAAACCC TGG	WT (0/34)		BE3-Tyr-D	#1	CTGTGCCAAGGCAGAAACCC TGG	WT (47/47)	
		CTGTGCTAAGGCAGAAACCC TGG	MUT (19/34)						
BE3-Tyr-D		CTGTGTTAAGGCAGAAACCC TGG	MUT (15/34)		BE3-Tyr-D	#2	CTGTGCCAAGGCAGAAACCC TGG	WT (32/32)	
	#2	CTGTGCCAAGGCAGAAACCC TGG	WT (2/12)						
		CTGTGCTAAGGCAGAAACCC TGG	MUT (10/12)						

Figure S4. On-target sequences from WGS in tdTomato⁺ and tdTomato⁻ cells. WT and mutant sequences of the WGS results were shown as “WT” and “MUT”, the number before the slash shows the WT or mutant sequence reads, and the total reads were shown after the slash. Mutated sequences were marked in red color, and PAM was marked with green.

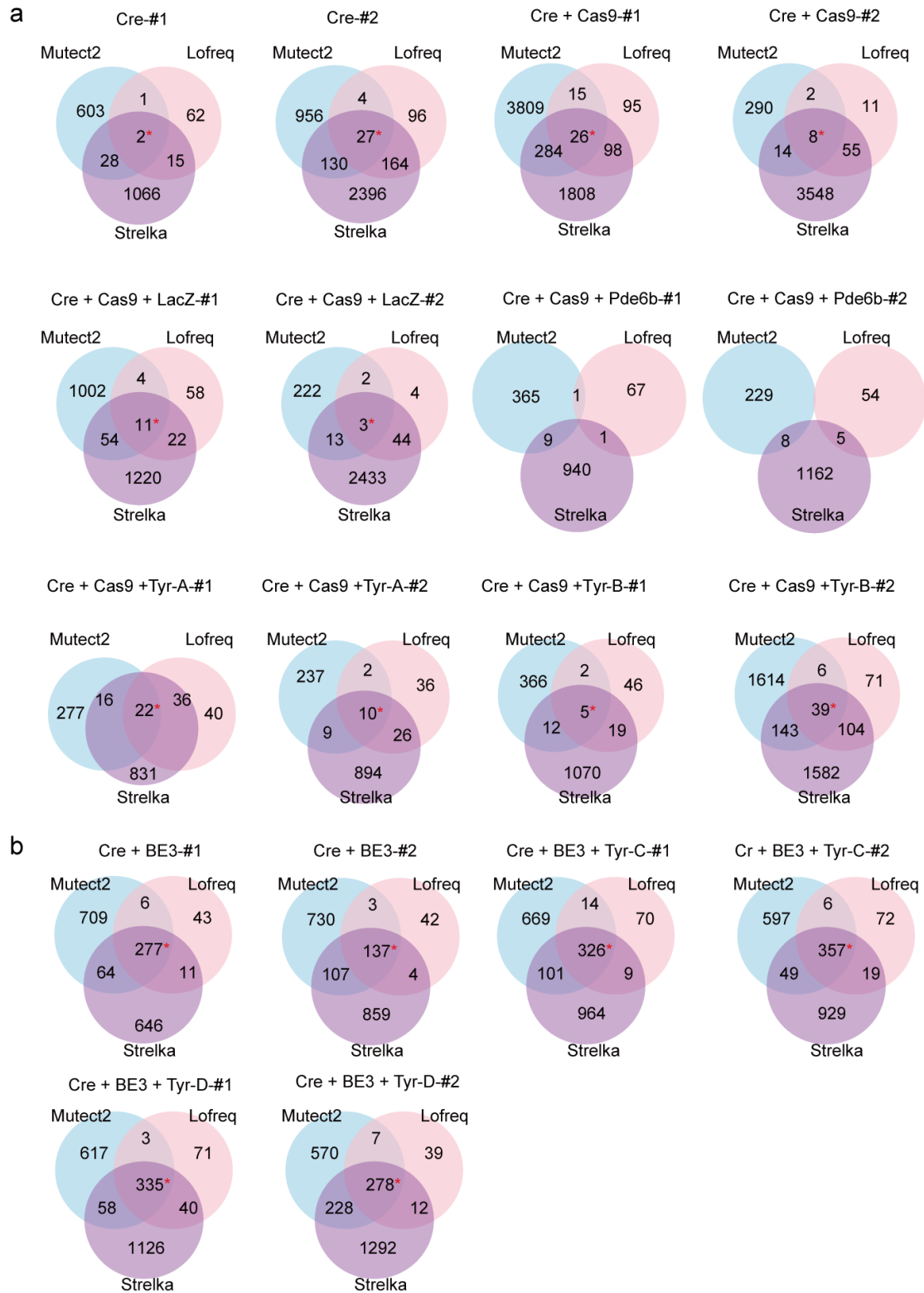


Figure S5. Venn diagrams of SNVs detected in each embryo by WGS data using the indicated software tools. a, SNVs detected in Cre or CRISPR/Cas9-treated samples. b, SNVs identified in BE3-treated embryos. *The overlap SNVs with allele frequencies less than 10% were removed from the following analysis.

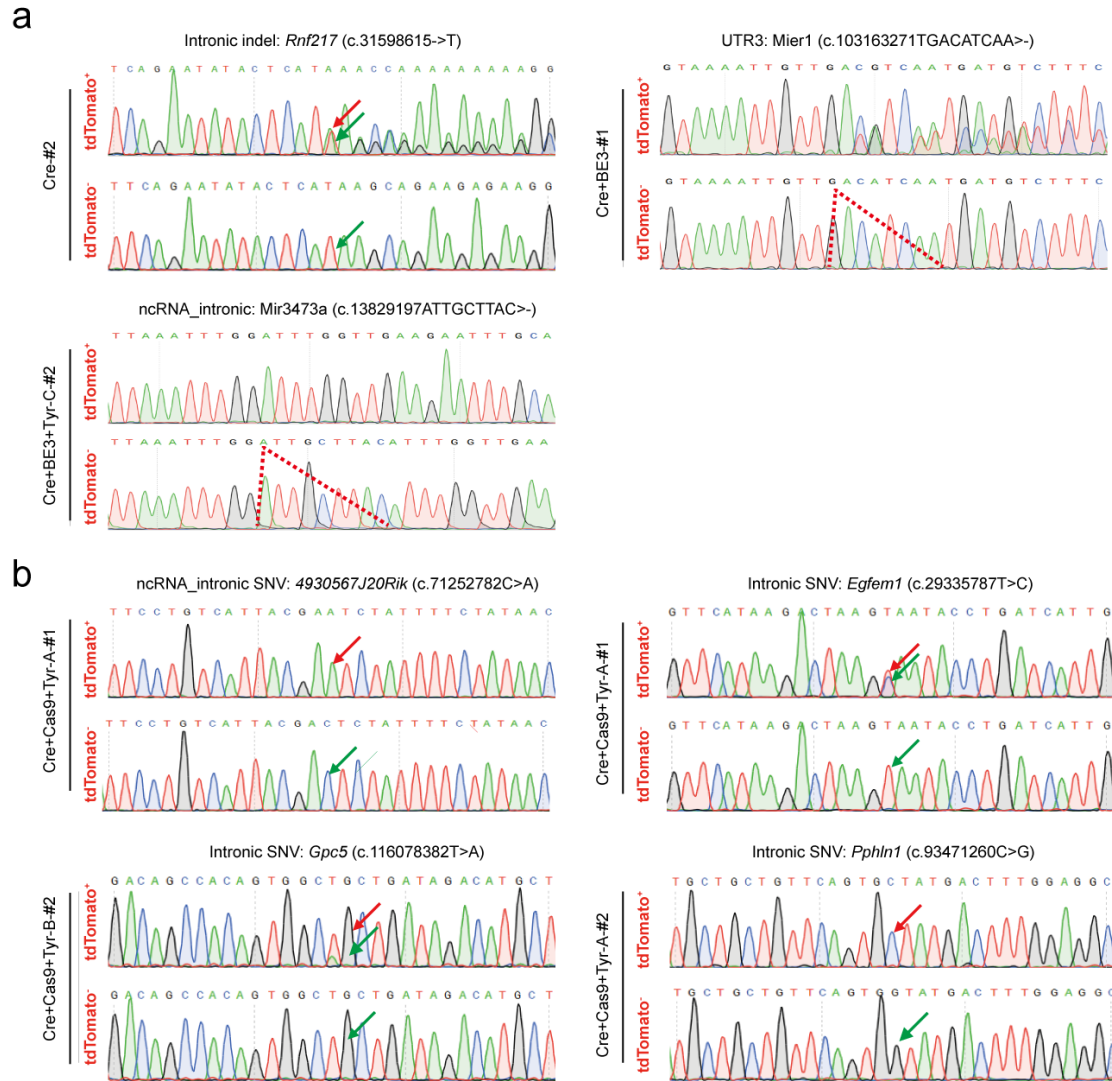


Figure S6. Representative Sanger sequencing chromatograms of mutations detected by WGS in Cre- or CRISPR/Cas9-treated embryos. **a**, Indels from indicated sample were PCR amplified and Sanger sequenced. Green arrows, wild type; red arrows, inserted nucleotide. Red dot line, deleted nucleotides. **b**, SNVs from indicated samples were validated by Sanger sequencing. Green arrows, wild-type nucleotide; red arrows, mutated nucleotide. Primers were listed in table S16.

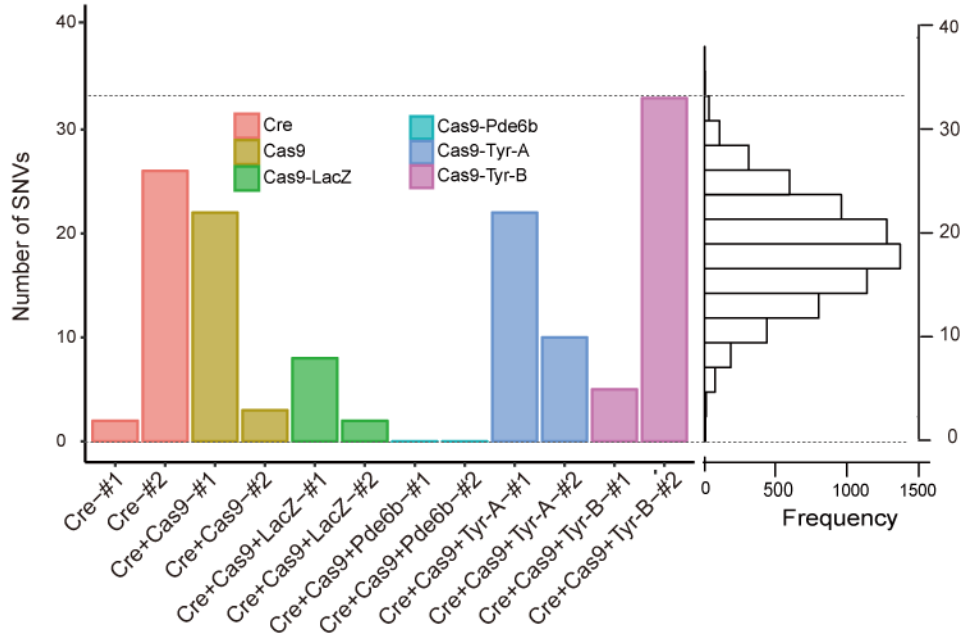


Figure S7. The number of SNVs detected from WGS data in embryos treated with Cre and CRISPR/Cas9. Embryos in the same group were indicated in the same color. The right panel is the distribution of simulated number of spontaneous mutations.

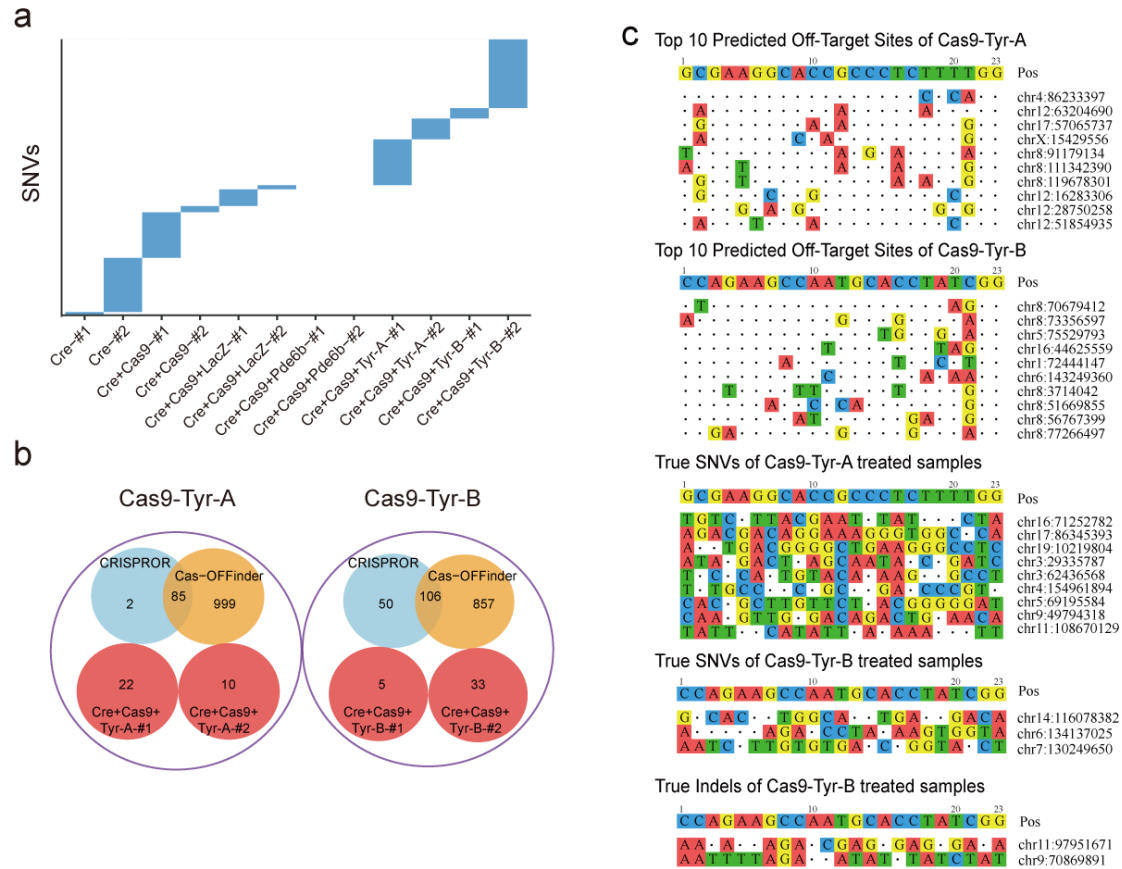


Figure S8. Off-target SNVs and indels identified from Cre- and CRISPR/Cas9-treated embryos. **a**, The SNVs identified from each embryo with Cre or CRISPR/Cas9 injected was mutually exclusive from the others. **b**, The overlap among SNVs detected from CRISPR/Cas9-treated embryos with predicted off-targets sites by Cas-OFFinder and CRISPROR. **c**, Sequence alignment of top 10 predicted off-target sites with Cas9-Tyr-A and Cas9-Tyr-B and the detected mutations from the WGS data with Cas9-Tyr-A and Cas9-Tyr-B.

a

Variants	Cre + Cas9 +											
	Cre-#1	Cre-#2	-#1	-#2	LacZ-#1	LacZ-#2	Pde6b-#1	Pde6b-#2	Tyr-A-#1	Tyr-A-#2	Tyr-B-#1	Tyr-B-#2
WGS SNVs	3	12	7	8	1	6	0	0	36	8	13	14
WGS Indels	3	0	1	2	0	1	0	0	4	1	2	1
Exon SNVs	0	0	0	0	0	0	0	0	1	0	2	2
Exon Indels	0	0	0	0	0	0	0	0	0	0	0	0
Nonsynonymous SNVs	0	0	0	0	0	0	0	0	1	0	1	2
Frameshift Indels	0	0	0	0	0	0	0	0	0	0	0	0
Predicted offtarget sites							0	0	0	0	0	0
Digenome-seq putative offtarget sites							0	0	0	0	0	0

b

Variants	Cre+ BE3 +					
	Cre+BE3-#1	Cre+BE3-#2	Tyr-C-#1	Tyr-C-#2	Tyr-D-#1	Tyr-D-#2
WGS SNVs	15	19	12	19	30	7
WGS Indels	1	1	1	0	1	1
Exon SNVs	0	0	0	0	0	0
Exon Indels	0	0	0	0	0	0
Nonsynonymous SNVs	0	0	0	0	0	0
Frameshift Indels	0	0	0	0	0	0

Figure S9. Summary of variants calling inversely from WGS data by comparing tdTomato⁻ with tdTomato⁺ cells from the same embryos. a, Converse variants calling from Cre- and CRISPR/Cas9-treated samples. **b,** Converse calling results from BE3-treated samples.

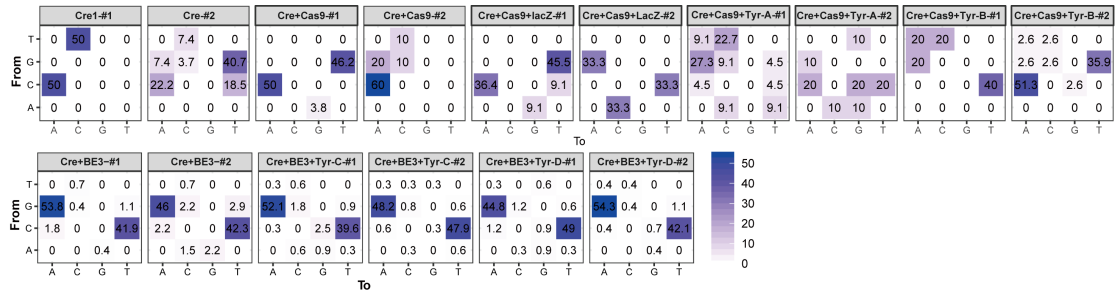


Figure S10. Mutation types of identified SNVs in each embryo included in the study. The number in each cell indicated the proportion of a certain mutation type, and deeper colors represented higher proportions.

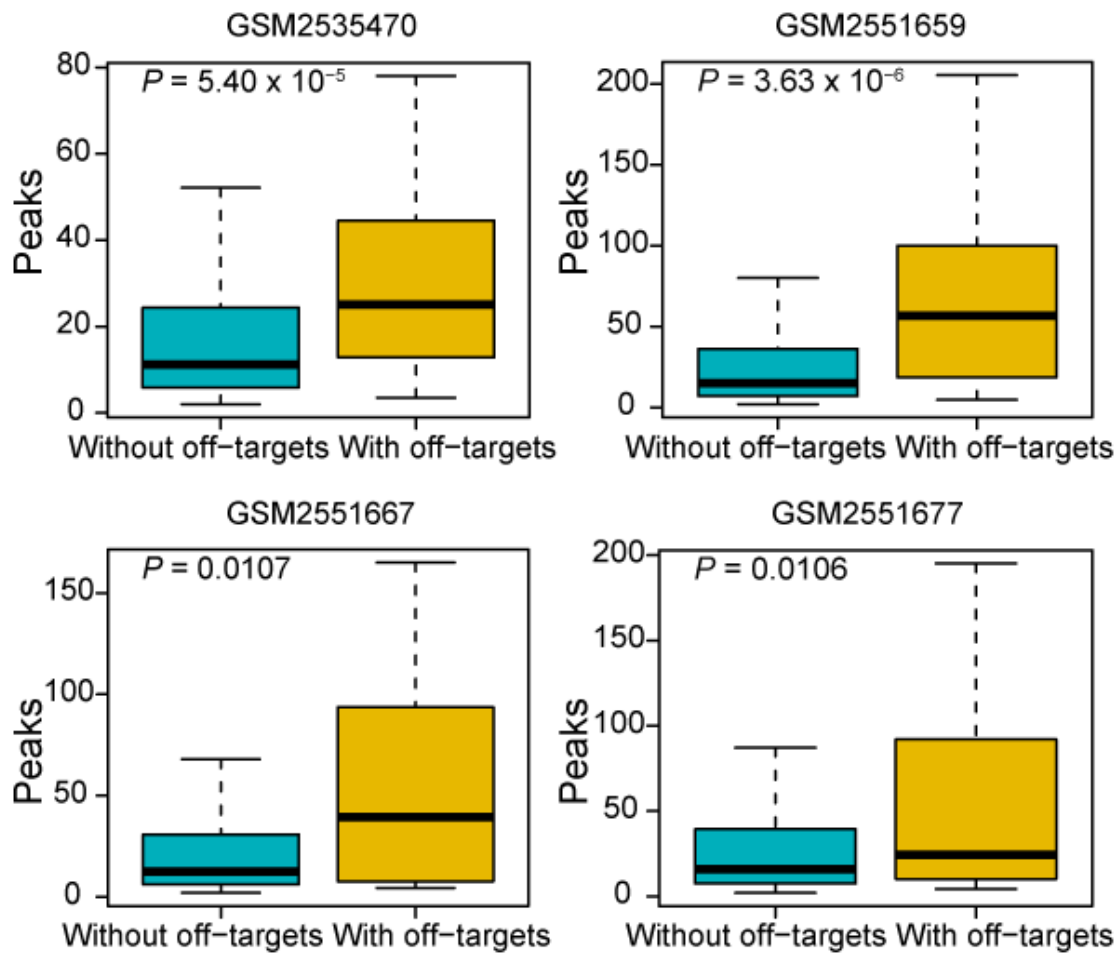


Figure S11. The comparison of peak regions with or without the identified off-targets from four datasets of Cistrome. The numbers on top of each boxplot indicated the GEO accession of applied datasets. *P*-values were calculated with Wilcoxon rank sum test.



Figure S12. Sanger sequencing chromatograms of SNVs located on exons detected by WGS in BE3-treated embryos. SNVs from the indicated samples were validated by Sanger sequencing. Green arrows, wild-type nucleotide; red arrows, mutated nucleotide. Primers were listed in table S16.

Materials and Methods

Mice

Female C57BL/6 mice (4 weeks old) and heterozygous Ai9 (full name B6.Cg-Gt(ROSA)26Sortm9(CAG-td-Tomato)Hze/J; JAX strain 007909) male mice were used for embryo collection. ICR females were used for recipients. 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.

Generation of Cas9 mRNA, BE3 mRNA, Cre mRNA and sgRNA

T7 promoter was added to Cas9 coding region by PCR amplification of px260, using primer Cas9 F and R. T7-Cas9 PCR product was purified and used as the template for *in vitro* transcription (IVT) using mMESSAGING mMACHINE T7 ULTRA kit (Life Technologies). T7 promoter was added to sgRNA template by PCR amplification of px330. The T7-sgRNA PCR product was purified and used as the template for IVT using MEGA shortscript T7 kit (Life Technologies). T7 promoter was added to Cre template by PCR amplification. T7-Cre PCR product was purified and used as the template for *in vitro* transcription (IVT) using mMESSAGING mMACHINE T7 ULTRA kit (Life Technologies). Cas9 mRNA, Cre mRNA and sgRNAs were purified using MEGA clear kit (Life Technologies) and eluted in RNase-free water.

sgRNA sequences

Locus	Sequence (5'-3')
Tyr-A	GCGAAGGCACCGCCCTCTTTTGG
Tyr-B	CCAGAAGCCAATGCACCTATCGG
LacZ	TGCGAATACGCCACGCGATGGG
Peb6b	CCAACCTAAGTAGCAGAAAGTGG
Tyr-C	GACCTCAGTTCCCCTTCAAAGGG
Tyr-D	CTGTGCCAAGGCAGAAACCCTGG

Primers

Name	Sequence (5'-3')
Cas9 IVT F	TAATACGACTCACTATAGGGAGATTTTCAGGTTGGACCGGTG
Cas9 IVT R	GACGTCAGCGTTCGAATTGC
Tyr-A IVT F	TAATACGACTCACTATAGGGGCCGAAGGCACCGCCCTCTTTGTTTTA GAGCTAGAAATAG
Tyr-B IVT F	TAATACGACTCACTATAGGGCCAGAAGCCAATGCACCTATGTTTTA GAGCTAGAAATAG
LacZ-IVT F	TAATACGACTCACTATAGGGTGCGAATACGCCACGCGATGTTTTA GAGCTAGAAATAG
Tyr IVT R	AAAAGCACCGACTCGGTGCC
Cre IVT F	TAATACGACTCACTATAGGGAGACAGATCACCTTTCCTATCAACC

Cre IVT R	TCGGTATTTCCAGCACACTGGA
BE3 IVT F	TCCGCGGCCGCTAATACGACT
BE3 IVT R	TGGTTCTTTCCGCCTCAGAAGCC
Tyr-C IVT F	TAATACGACTCACTATAGGGGACCTCAGTTCCCCTTCAAAGTTTTA GAGCTAGAAATAG
Tyr-D IVT F	TAATACGACTCACTATAGGGCTGTGCCAAGGCAGAAACCCGTTTTA GAGCTAGAAATAG
Tyr IVT R	AAAAGCACCGACTCGGTGCC

2-cell Embryo Injection, Embryo Culturing, and Embryo Transplantation

Super ovulated C57BL/6 females (4 weeks old) were mated to heterozygous Ai9 (full name B6.Cg-Gt(ROSA)26Sortm9(CAG-td-Tomato)Hze/J; JAX strain 007909) males, and fertilized embryos were collected from oviducts 23 h post hCG injection. For 2-cell editing, the mixture of Cas9 mRNA (50 ng/μl) or BE3 mRNA (50 ng/μl), sgRNA (50 ng/μl) and Cre mRNA (2 ng/μl) was injected into the cytoplasm of one blastomere of 2-cell embryo 48 h post hCG injection in a droplet of HEPES-CZB medium containing 5 μg/ml cytochalasin B (CB) using a FemtoJet microinjector (Eppendorf) with constant flow settings. The injected embryos cultured in KSOM medium with amino acids at 37 °C under 5% CO₂ in air for 2 hours and then transferred into oviducts of pseudopregnant ICR females at 0.5 dpc.

Single-cell PCR Analysis

To pick up and transfer single cells, we used a glass capillary under a dissection microscope. Eight-cell mouse embryos were digested with acid Tyrode solution to remove the zona pellucida, and then the embryos were transferred into 0.25% trypsin and gently pipetted to separate individual blastomere. Finally, the blastomeres were washed into 0.25% trypsin for 7 to 10 times and transferred into a PCR tube. 1.5 μl lysis buffer containing 0.1% tween 20, 0.1% Triton X-100 and 4 μg/ml proteinase K was then pipetted into the tube. Each tube was centrifuged to facilitate the mix. The lysis was incubated at a temperature of 56°C for 30 min, followed by 95°C for 5 min. The products of the lysis program were used as templates in a nest PCR analysis. All the operations were operated carefully in order not to pollute the samples.

Nest PCR primer

	Outer F: GTTATCCTCACACTACTTCTG
	Outer R: GTAATCCTACCAAGAGTCTCA
Tyr	Inner F: TCCTCACACTACTTCTGATG
	Inner R: GTCTCAAGATGGAAGATCAC

TA Cloning and Genotyping

The PCR product was purified and ligated to pMD18-T vector and transformed to competent

E. coli strain DH5 α . After overnight culture at 37 °C, randomly selected clones were sequenced by the Sanger method. The genotypes of mutant E14.5 embryos were determined by PCR of genomic DNA extracted from cells. ExTaq was activated at 95°C for 3 min, and PCR was performed for 34 cycles at 95°C for 30 sec, 62°C for 30 sec, and 72°C for 1min, with a final extension at 72°C for 5 min. For blastocysts, after washing 6 times with KSOM, the single blastocyst was transferred directly into PCR tubes containing 1.5 μ l embryo lysis buffer (0.1% tween 20, 0.1% Triton X-100 and 4 μ g/ml proteinase K) and incubate for 30 min at 56°C, heat inactivate proteinase K at 95°C for 10 min. PCR amplification was performed using nest primer sets. ExTaq was activated at 95°C for 3 min, and PCR was performed for 34 cycles at 95°C for 30 sec, 62°C for 30 sec, and 72°C for 1min, with a final extension at 72°C for 5 min. Secondary PCR was performed using 0.5 μ g primary PCR product and nested inner primer. PCR was carried out in the same reaction mixture. PCR product was gel purified and cloned using pMD-19t cloning kit (Takara) following the manufacturer's instructions. Colonies were picked from each transformation and then Sanger sequencing to detect mutations. The primers used were:

Name	Sequence (5'-3')
Tyr F	GTTATCCTCACACTACTTCTG
Tyr R	GTAATCCTACCAAGAGTCTCA
Tyr-OF	GTCTGTGACACTCATTAACC
Tyr-OR	CATAGGAGGTGCTAACAATAC
Tyr-IF	GTATTGCCTTCTGTGGAGTT
Tyr-IR	TGAACCAATCAGTCCTTGTT

FACS

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- μ m 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.

Whole genome sequencing and data analysis

Genomic DNA was extracted from cells by using the DNeasy blood and tissue kit (catalog number 69504, Qiagen) according to the manufacturer's instructions. WGS was performed at mean coverages of 50x by Illumina HiSeq X Ten. BWA (v0.7.12) was used to map qualified sequencing reads to the reference genome (mm10). Picard tools (v2.3.0) was then applied to sort and mark duplicates of the mapped BAM files. To identify the genome wide *de novo* variants with high confidence, we conducted single nucleotide variation calling on three algorithms, Mutect2 (v3.5), Lofreq (v2.1.2) and Strelka (v2.7.1), separately(37-39). In parallel, Mutect2 (v3.5), Scalpel (v0.5.3) and Strelka (v2.7.1) were run individually for the detection of whole genome *de novo* indels(37, 39, 40). The overlap of three algorithms of SNVs or indels

were considered as the true variants. We also marked variants overlapped with the UCSC repeat regions and microsatellite sequences or existed in the dbSNP (v138) and MGP (v3) databases. All the sequencing data were deposited in NCBI Sequence Read Archive (SRA) under project accession SRP119022.

To validate the on-target efficiency, we blasted the aligned BAM files with the on-target sites with e-value set as 0.0001. Potential off-targets of targeted sites were predicted using two previous reported algorithms, Cas-OFFinder (<http://www.rgenome.net/cas-offinder/>) and CRISPOR (<http://crispor.tefor.net/>) with all possible mismatches(21, 22).

The SNVs and indels were annotated with annovar (version 2016-02-01) using RefSeq database(41). Proto-oncogenes and tumor-suppressor genes were retrieved from the UniprotKB/Swiss-Prot database (2018_09). We downloaded five ATAC-seq bed peak files from the CistromeDB database with biological sources as embryonic and passing all quality controls (28). The five datasets retrieved included CistromeDB IDs “79877” (GSM2551659), “79976” (GSM2551677), “80493” (GSM2535470), “81049” (GSM2551664) and “81052” (GSM2551667). The off-target sites were mapped to the peak regions in each bed file based on the chromosome positions, and then the peaks regions with or without the off-targets were compared with each other by two-sided Wilcoxon rank sum test.

Simulation of spontaneous mutations in the development of embryos

To estimate the number of spontaneous mutations from the 2-cell stage to E14.5 stage, the occurrence of single nucleotide mutations was simulated in silico considering the average sequencing coverage of 40 and allele frequency threshold of 10%. For each round of simulation, given the mutation rate of 1.8×10^{-10} and the size of mouse nuclear genome (2,785,490,220bp), we considered the replication process from the 2-cell stage to 16-cell stage, for mutations occurring after 16-cell stage will not be detected taken the allele frequency into consideration. Each cell may mutate or not during each replication, and once mutation occurs, the divided cells will inherit the mutation. Then the accumulated mutations and their wild-type alleles were randomly selected for sequencing with depth of 40, the selected mutations were summed as the number of spontaneous mutations for each round, and the same process was repeated for 10,000 times.

In vitro cleavage of DNA

The genomic region flanking the sgRNA target site for each gene was PCR amplified and PCR amplicons was purified with the Universal DNA Purification Kit (Tiangen) according to the manufacturer's instructions. Genomic DNA was purified from mice tail with the TIANamp Genomic DNA Kit (Tiangen) according to the manufacturer's instructions. The Cas9 ribonucleoproteins (1 μ g) and sgRNA (1 μ g) were pre-incubated at room temperature for 10 min to form RNP complexes. DNA (4 μ g) was incubated with RNP complexes in a reaction buffer for 3 h at 37°C. Digested DNA was purified again with Universal DNA Purification Kit (Tiangen) after RNase A (100 μ g/ml) was added to remove sgRNA.

Digenome sequencing analysis

Libraries were subjected to WGS using Illumina HiSeq X Ten sequencer and the WGS was performed at a sequencing depth of 30x to 40x. Qualified reads were aligned to the mouse

reference genome (mm10) by Isaac aligner with the following parameters: base quality cutoff, 15; keep duplicate reads, yes; variable read length support, yes; realign gaps, no. DNA cleavage sites were identified computationally using Digenome-seq2 (<https://github.com/chizksh/digenome-toolkit2>).

Statistical analysis

R version 3.5.1 (<http://www.R-project.org/>) was used to conduct all the statistical analyses in this work. All tests conducted were two-sided, and the significant difference was considered at $P < 0.05$.