CRISPR off-target analysis in genetically engineered rats and mice

Keith R. Anderson[1], Maximilian Haeussler[2], Colin Watanabe[3], Vasantharajan Janakiraman[1], Jessica Lund[1], Zora Modrusan[1], Jeremy Stinson[1], Qixin Bei[1], Andrew Buechler[1], Charles Yu[1], Sobha R. Thamminana[1], Lucinda Tam[1], Michael-Anne Sowick[4], Tuija Alcantar[4], Natasha O'Neil[4], Jinjie Li[4], Linda Ta[4], Lisa Lima[4], Merone Roose-Girma[1], Xin Rairdan[4], Steffen Durinck[1] and Søren Warming ●[1]*

Despite widespread use of CRISPR, comprehensive data on the frequency and impact of Cas9-mediated off-targets in modified rodents are limited. Here we present deep-sequencing data from 81 genome-editing projects on mouse and rat genomes at 1,423 predicted off-target sites, 32 of which were confirmed, and show that high-fidelity Cas9 versions reduced off-target mutation rates in vivo. Using whole-genome sequencing data from ten mouse embryos, treated with a single guide RNA (sgRNA), and from their genetic parents, we found 43 off-targets, 30 of which were predicted by an adapted version of GUIDE-seq.

Several studies have analyzed CRISPR off-targets in cell lines, and a number of existing methods can be used to generate lists of off-targets in an unbiased manner: high-throughput genome-wide translocation sequencing[1], GUIDE-seq (genome-wide, unbiased identification of double-strand breaks enabled by sequencing)[2], Digenome-seq (in vitro Cas9-digested whole-genome sequencing)[3], BLESS (labeling of double-strand breaks followed by enrichment and sequencing)[4] and, most recently, SITE-seq (a biochemical method that identifies DNA cut sites)[5] and CIRCLE-seq (an in vitro method for identifying off-target mutations)[6]. To our knowledge, these methods are not currently used to test sgRNAs prior to the generation of rodent models. Instead, analysis for off-target risk in rodent genome editing relies on CRISPR design algorithms[7–9].

Most facilities that generate genome-edited mice and rats adhere to a minimal set of guidelines to limit off-target effects. Although such guidelines are useful, even the most optimal sgRNAs have a predicted list of potential off-targets, which is valuable only if the resulting animals are subjected to experimental testing. Because of the limited understanding of long-range enhancers and other regulatory elements, potential off-targets outside annotated genes cannot be disregarded[10]. Key advantages of CRISPR for genome editing include the speed with which animal models with new mutations can be generated, and it is desirable to avoid the need for back-crossing to the original strain for more than one generation. Furthermore, the success of knock-in projects depends on the proximity of the initial double-strand break to the target site[11], and in some cases it is therefore necessary to compromise on sgRNA specificity.

Our CRISPR workflow for rodent genome editing projects is outlined in Supplementary Fig. 1a. A similar sequencing approach was described previously[12], but in that report data from only five sgRNAs

and 56 potential off-targets are presented, and no off-target mutations were identified. In addition to the identification of lower-frequency off-targets that might otherwise be missed (Supplementary Note 1), deep-sequencing analysis of the mosaic G0 founders rather than their G1 progeny also allows for the selection of founders with high contributions of the desired allele for more efficient breeding.

Twenty-three percent (19/81) of our animal model projects had off-targets (Fig. 1a), defined by at least one animal with at least one allele with Cas9-induced mutations in >3% of the sequence reads in at least one of the analyzed off-target loci. Some animal model projects require the use of two sgRNAs, for example, to generate a large knockout deletion. Eighteen percent of the sgRNAs investigated (21/119; Supplementary Fig. 1b) had off-target activity, and 8 of those 21 sgRNAs had activity at more than one predicted off-target. CRISPR–Cas9 activity was detected at 2% (32/1,423) of predicted off-target loci with informative sequence reads (Supplementary Fig. 1c). Fifty-six percent (18/32) of the off-target loci were in introns or within a conservatively defined promoter/untranslated region (10 kb upstream or downstream) of known genes (Fig. 1b). Summaries for the 21 off-target-positive sgRNAs and a complete list of all sgRNAs analyzed in this study with predicted off-targets are provided in Supplementary Figs. 2 and 3 and Supplementary Table 1, respectively.

An overview of the 32 identified off-targets is provided in Supplementary Fig. 4, and the distribution of on- and off-target mutation frequencies is presented in Fig. 1c–f and Supplementary Fig. 5. A breakdown of allele frequencies for all G0 founders from off-target-positive sgRNAs is provided in Supplementary Figs. 6 and 7 and Supplementary Table 2. For five projects, G0 founders with off-target alleles were bred, and data for transmission of off-target alleles from these founders to their G1 progeny are provided in Supplementary Fig. 6. Our data show that even off-target alleles representing about 10% of reads can be transmitted to the next generation. Examples of genomic alignments are provided in Supplementary Fig. 8. The majority of Cas9-induced insertion/deletion mutations would be small enough to be identified by our analysis[13]. Therefore, apart from the 113 predicted off-target loci that lacked sequence data, it is unlikely that predicted and top-ranked off-target-positive loci were missed in this study.

sgRNAs that result in off-target activity are not more active than sgRNAs with no such activity (Supplementary Fig. 9a). We did not observe a strong correlation between MIT specificity score
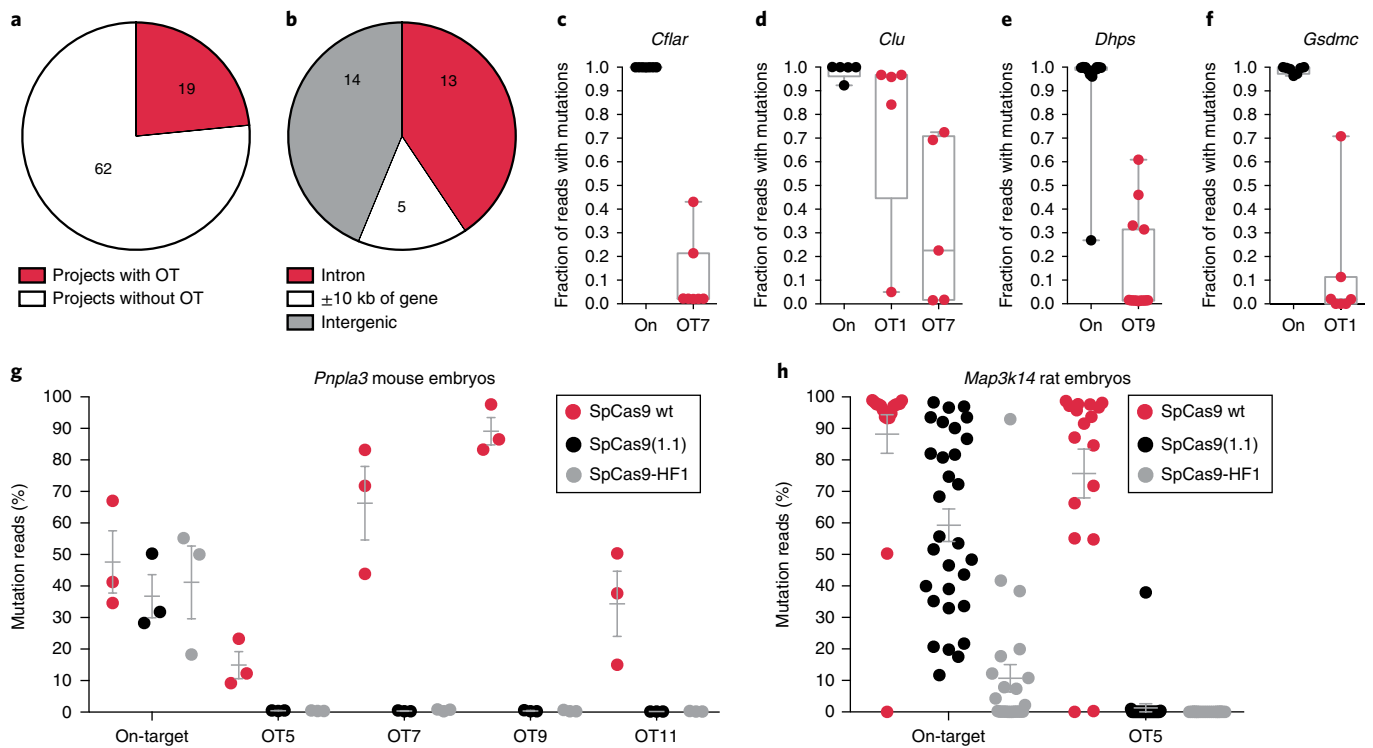
Fig. 1 | Off-targets identified from CRISPR projects and in vivo reduction of off-target frequency with re-engineered Cas9. a, The fraction of projects with and without off-targets (OT). **b**, The distribution of identified off-targets with respect to known coding genes. In **a** and **b**, numbers within the graphs represent the total for each condition. **c–f**, Box-and-whisker plots showing the distribution of 4/21 targets (On; black dots) and 5/32 off-targets (red dots). Y-axes indicate the fraction of next-generation sequencing reads with evidence of Cas9 activity at each locus (1 – fraction of wild-type reads). **c**, *Cflar*, $n = 7$ mice. **d**, *Clu*, $n = 5$ mice. **e**, *Dhps*, $n = 15$ mice. **f**, *Gsdmc*, $n = 7$ mice. Plots represent the median (center line), minimum–maximum range (whiskers), and lower and upper quartiles (edges). Data points represent individual animals. **g**, Percentage of mutation reads at mouse *Pnpla3* on- and off-targets. Each data point represents pooled blastocysts from one microinjection of SpCas9 wild type (wt; $n = 45, 32, 56$ blastocysts), SpCas9(1.1) ($n = 50, 51, 65$) or SpCas9-HF1 ($n = 49, 45, 73$). Unpaired two-tailed t-tests were carried out for on-target means being identical (df = 4): wild type versus 1.1, $t = 0.9025$, $P = 0.42$; wild type versus HF1, $t = 0.4260$, $P = 0.69$; 1.1 versus HF1, $t = 0.3259$, $P = 0.76$. **h**, Percentage of mutation reads at the rat *Map3k14* on- and off-targets. Each data point represents results of next-generation sequencing analysis of one rat embryo: SpCas9 wild type, $n = 17$; SpCas9(1.1), $n = 30$; SpCas9-HF1, $n = 24$. Unpaired two-tailed t-tests were carried out for on-target means being identical: wild type versus 1.1, df = 45, $t = 3.488$, $P = 0.0011$; wild type versus HF1, df = 39, $t = 10.66$, $P < 0.0001$; 1.1 versus HF1, df = 52, $t = 6.994$, $P < 0.0001$. **g,h**, Data are shown as mean ± s.e.m. **c–h**, Off-targets are numbered according to rank in the top-15 list of predicted off-target locations.

(a composite 0–100 quality score for sgRNAs)[7] and the fraction of animals with off-target hits ($r^2 = 0.034$; Supplementary Fig. 9b). Nevertheless, our data suggest, in agreement with previous observations[9], a specificity score[7] cutoff of 66. The odds ratio for identifying an sgRNA without off-target mutations was 18 when the score was ≥66 ($P = 0.0001$) (Supplementary Note 2). Although the identified off-targets tended to rank higher among the top 15 predicted off-targets (Supplementary Fig. 9c), several ranked lower, which suggests that true off-targets could be missed if analysis were restricted to the top 15 predicted genomic sites.

Using two of the sgRNAs found to have off-targets (mouse *Pnpla3* and rat *Map3k14*), we compared engineered Cas9 variants with improved specificity[14,15] to wild-type Cas9 in mouse and rat embryos (Fig. 1g,h) and cell lines (Supplementary Fig. 10). Both Cas9 variants, eSpCas9(1.1) and SpCas9-HF1, reduced the off-target mutation frequency compared with that observed with wild-type Cas9. For *Pnpla3* embryos, we did not observe a reduction in on-target efficiency for the variants (Fig. 1g). In cells, SpCas9-HF1 activity was slightly higher than that of wild-type Cas9 (Supplementary Fig. 10a). For *Map3k14*, engineered Cas9 on-target efficiency was more variable and lower than wild-type Cas9 activity (Fig. 1h, Supplementary Fig. 10b), in agreement with other findings (Supplementary Note 3). We recommend that eSpCas9(1.1)[14] and SpCas9-HF1[15] or other engineered Cas9 variants (e.g., HypaCas9[16])

be considered for the routine generation of animal models, especially for projects where lower on-target efficiency is acceptable.

Because we observed 4 off-targets among the predicted top 15 for the mouse *Pnpla3* sgRNA, and because the specificity score was very low (26.4), we next used an unbiased approach to identify all true off-targets for this sgRNA, which allowed us to more precisely test the predictive value of existing algorithms. Using a mouse cell line, we first carried out target-enriched GUIDE-seq (TEG-seq), a variant of GUIDE-seq[2] adapted for the Ion Torrent sequencing platform (Methods), and identified 170 DNA-tag insertion sites. Deep-sequencing analysis of amplicons (AmpliSeq) confirmed 105 of the sites as off-targets (Supplementary Table 3). As shown in Supplementary Fig. 11, 63 of the 105 AmpliSeq-validated off-targets overlap with our list of 10,360 CRISPOR-predicted[9] off-targets having up to five mismatches. Forty-two off-targets had between six and eight mismatches, and for some, bulging[17] of base pairs was required for proper alignment (Supplementary Table 3).

It is unknown how well GUIDE-seq analysis predicts off-targets generated in vivo. We therefore carried out in vitro fertilization and microinjection of *Pnpla3* sgRNA and Cas9 mRNA into zygotes from C57BL/6J mice, followed by whole-genome sequencing of ten embryos and their genetic parents, at an average of 80× coverage (Methods). Filtering (Supplementary Notes) included subtraction of all variants found in the genetic parents, and we identified
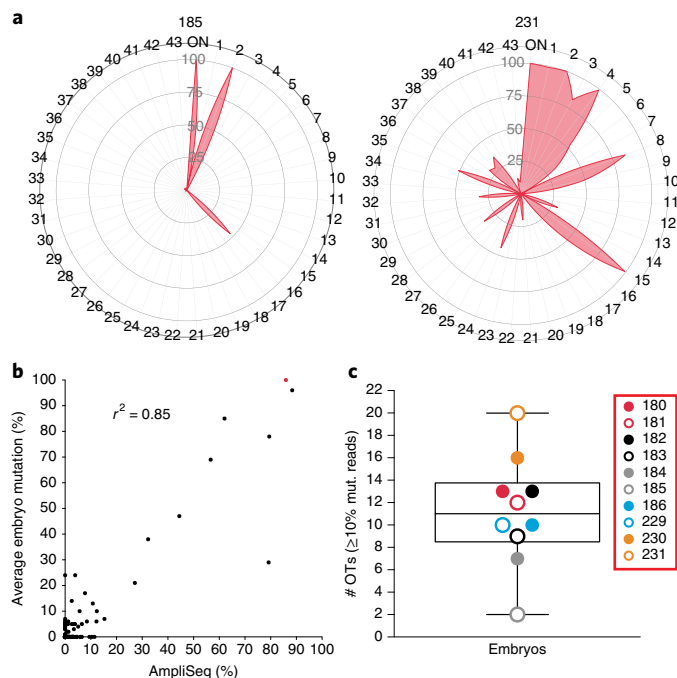
Fig. 2 | TEG-seq is a good predictor of in vivo activity. a, Radar plots showing the distribution of on- and off-target activity across the target and 43 off-target loci identified by whole-genome sequencing. Off-target numbers are indicated on the periphery. Data from two embryos are shown (185, smallest number of off-targets; 231, greatest number of off-targets). Percent mutation read quartiles are indicated by gray circles. **b**, Correlation between AmpliSeq mutation frequency in Neuro-2a cells and average embryo mutation frequency ($n=119$ loci; 118 off-targets (black dots) and 1 *Pnpla3* target (red dot)). The r^2 value (coefficient of determination) is shown (correlation coefficient $r=0.9225$; two-tailed t-test, r value not significantly different from zero, $P<0.0001$). **c**, Box-and-whisker plot depicting the number of off-targets (OTs) with at least 10% mutation (mut.) reads per embryo; $n=10$ mouse embryos. The plot shows the median (center line), minimum–maximum range (whiskers), and lower and upper quartiles (edges).

43 true Cas9-generated off-targets (Supplementary Fig. 12 and Supplementary Table 3). Examples of *Pnpla3* off-target genomic alignments are provided in Supplementary Fig. 13, and all 43 can be viewed from a data track hub on the UCSC Genome Browser (Methods). Thirty of the 43 off-targets overlap with the validated TEG-seq hits, and 25 of those 30 also overlap with the predicted off-targets (Supplementary Fig. 11 and Supplementary Note 4).

The distribution of the 43 off-targets among all ten embryos is shown in Fig. 2a and Supplementary Figs. 14 and 15a. We observed strong correlation ($r^2=0.85$, $P<0.0001$) between the AmpliSeq mutation frequencies and the average embryo mutation frequency (Fig. 2b). The correlation between number of TEG-seq reads and average embryo mutation frequency was less strong ($r^2=0.58$, $P<0.0001$; Supplementary Fig. 15b), and we did not observe a strong correlation between average embryo mutation frequency and MIT[7] ($r^2=0.30$, $P=0.0001$) or cutting frequency determination (CFD)[8] ($r^2=0.38$, $P<0.0001$) off-target score (Supplementary Fig. 15c,d).

To calculate AmpliSeq's ability to predict true off-targets in vivo, we chose off-targets with at least 10% average mutation frequency in mosaic mouse embryos, as lower-frequency alleles are less likely to be transmitted to the next generation (Supplementary Fig. 6). Figure 2c shows the number of off-targets with ≥10% frequency per embryo; we note that only 15/43 off-targets had an average frequency greater than 10%. All 15 of these off-targets were observed in at least two embryos. Of the 15 off-targets, 1 had six mismatches, 2 had five mismatches, and the remaining 12 had one to four mismatches. Supplementary

Figure 15e shows the fraction of the 15 off-targets captured at a given AmpliSeq mutation frequency cutoff. With a 2.5% AmpliSeq cutoff, 14 out of 15 off-targets would be identified (93%). Off-target 21 (Supplementary Table 3) was identified by TEG-seq but was not subsequently confirmed by AmpliSeq. By comparison, algorithm off-target score cutoffs of 0.1 (1,423 loci; MIT) or 0.2 (250 loci; CFD) were needed to capture the majority (73% and 93%, respectively) of these 15 hits (Supplementary Fig. 15f,g and Supplementary Note 5).

Although the *Pnpla3* sgRNA produced a considerable number of off-target mutations, it probably represents a worst-case scenario (Supplementary Note 6). The use of sgRNAs with higher specificity scores and/or Cas9 with increased fidelity should make it possible to reduce the risk and effects of off-targets. However, to avoid unexpected phenotypes altogether, we also recommend prediction of off-targets by unbiased methods followed by AmpliSeq screening of G0 founder animals.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41592-018-0011-5.

References

1. Frock, R. L. et al. *Nat. Biotechnol.* **33**, 179–186 (2015).
2. Tsai, S. Q. et al. *Nat. Biotechnol.* **33**, 187–197 (2015).
3. Kim, D. et al. *Nat. Methods* **12**, 237–243 (2015).
4. Crosetto, N. et al. *Nat. Methods* **10**, 361–365 (2013).
5. Cameron, P. et al. *Nat. Methods* **14**, 600–606 (2017).
6. Tsai, S. Q. et al. *Nat. Methods* **14**, 607–614 (2017).
7. Hsu, P. D. et al. *Nat. Biotechnol.* **31**, 827–832 (2013).
8. Doench, J. G. et al. *Nat. Biotechnol.* **34**, 184–191 (2016).
9. Haeussler, M. et al. *Genome Biol.* **17**, 148 (2016).
10. Hnisz, D. et al. *Science* **351**, 1454–1458 (2016).
11. Paquet, D. et al. *Nature* **533**, 125–129 (2016).
12. Singh, P., Schimenti, J. C. & Bolcun-Filas, E. *Genetics* **199**, 1–15 (2015).
13. Dow, L. E. et al. *Nat. Biotechnol.* **33**, 390–394 (2015).
14. Slaymaker, I. M. et al. *Science* **351**, 84–88 (2016).
15. Kleinstiver, B. P. et al. *Nature* **529**, 490–495 (2016).
16. Chen, J. S. et al. *Nature* **550**, 407–410 (2017).
17. Lin, Y. et al. *Nucleic Acids Res.* **42**, 7473–7485 (2014).

Author contributions

K.R.A. and S.W. designed the study and wrote the manuscript, with input from all other coauthors. K.R.A., M.H., C.W. and S.W. analyzed the data. S.D., C.W. and Q.B. generated the primer design and deep-sequencing analysis pipeline. S.D. analyzed the whole-genome sequencing data. K.R.A., A.B., M.R.-G., L. Tam, C.Y. and S.R.T. designed the CRISPR strategies and analyzed the founders and G1 progeny. X.R., T.A., N.O., J. Li, L. Ta and L.L. performed in vitro fertilization, microinjections and tissue/embryo collection. M.-A.S. coordinated mosaic founder breeding and colony management. V.J., J. Lund, Z.M. and J.S. generated sequencing libraries and performed deep sequencing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41592-018-0011-5.

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

Correspondence and requests for materials should be addressed to S.W.

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

Methods

Animals. All mice and rats were generated at Genentech and maintained in accordance with American Association of Laboratory Animal Care (AALAC) guidelines. The experiments were conducted in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the Genentech Institutional Animal Care and Use Committee.

sgRNA design and off-target prediction. Depending on the type of animal model to be engineered, sgRNAs in the genomic region of interest were identified initially (first nine animal models) by a previously published scoring algorithm (http://crispr.mit.edu[7]) and subsequently by a Molecular Biology CRISPR design tool (Benchling) that uses the same algorithm to provide 'MIT' specificity scores for each sgRNA, as well as the top 50 predicted off-target loci and corresponding MIT off-target scores, and also highlights mismatches. The first edition of the Benchling tool masked repeat regions similarly to the published tool from Hsu et al.[7]. The current version of Benchling's CRISPR design tool allows unmasking of repeat regions and includes any potential off-target hits in the top-50 list based on scores. Final sgRNAs used for editing were chosen on the basis of a qualitative balance of specificity scores, distance to desired mutation/insertion and manual assessment of the off-target list. In the off-target list assessment, we considered the preference to avoid sgRNAs with potential hits in coding regions, sgRNAs with off-target hits on the same chromosome as the intended target, and, when possible, any sgRNAs that had many predicted off-targets lacking mismatches in the seed region (10–12 nt proximal to the protospacer-adjacent motif (PAM)), as well as whether off-targets had NGG or NAG PAMs. Once an sgRNA decision was finalized, the off-target list was used to identify the top 11, and later top 15, off-target loci per sgRNA, and next-generation sequencing (NGS) amplicon primers were designed for the on-target locus and each of the off-targets (more details below).

Microinjection of mouse and rat embryos. sgRNAs were generated by MEGA shortscript T7 in vitro transcription (Thermo Fisher; AM1354) according to standard methods (performed in-house or purchased directly from Thermo Fisher). For all sgRNAs, a fragment was cloned into a plasmid backbone, sequence-verified, linearized and used as template for in vitro transcription. The scaffold for all sgRNAs in this study was 5′-gtttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgc-3′.

In vitro–transcribed sgRNAs were purified (MEGAclear kit; Thermo Fisher; AM1908) and purity quality control was performed with the Agilent small RNA kit (5067-1548), Agilent RNA 6000 nano kit (5067-1511) or Advanced Analytics standard RNA analysis kit (DNF-471). Using Bioanalyzer (Agilent; G2939B), we calculated a more accurate concentration as the area under the expected peak curve, which, depending on the sequence and strength of the secondary structure, ranged from 75 to 105 nt. Reagent concentrations for microinjection were as follows: 25 ng/μl Cas9 mRNA (Thermo Fisher; A29378) + 13 ng/μl each sgRNA. For the whole-genome sequencing project, pronuclear-stage embryos from C57BL/6 J mice (The Jackson Laboratory) were generated by in vitro fertilization (IVF). Two 3-week-old females were injected with 0.1 ml of HyperOva (Cosmo Bio)[18] followed by 5 IU of human chorionic gonadotropin (hCG; NHPP) 47 h later. IVF was performed 14 h after the hCG injection in modified Tyrode's solution (in g/L: 7.31 NaCl, 0.20 KCl, 0.04 NaH$_2$PO$_4$, 2.10 NaHCO$_3$, 0.10 MgCl$_2$ 6H$_2$0, 0.26 CaCl$_2$ 2H$_2$0, 1.00 glucose, 0.5 penicillin, and 4 BSA) supplemented with 1.25 mM reduced glutathione (Sigma). Oocytes (126 and 103, respectively) collected from the two females underwent IVF in separate dishes with sperm from a single 11-week-old male mouse. Spleen samples were collected from all three donor mice after oocyte or sperm collection, and the samples were snap-frozen in liquid nitrogen and stored at −80 °C until genomic DNA preparation. Approximately 6 h after IVF, we used an inverted microscope (Nikon Eclipse, TS100) to sort zygotes with two visible pronuclei. Groups of 25–50 pronuclear-stage embryos were transferred into an injection slide/chamber containing M2 medium (Zenith Bio) supplemented with 5 μg/ml cytochalasin B (Sigma) for 10 min[19]. Subsequently, cytoplasmic injection was performed with a microinjection needle crafted with a P-97 micropipette puller (Sutter Instruments). Embryos that survived microinjection were incubated (37 °C, 6% CO$_2$) overnight in KSOM + AA (Zenith Bio) drops covered with mineral oil (Sigma). Each e0.5 pseudopregnant ICR female mouse received 22–30 two-cell-stage embryos by oviduct transfer surgery. The injection mixture was prepared the day of microinjection with RNase- and DNase-free reagents. Injection buffer: 10 mM Tris-HCl, 0.25 mM EDTA, pH 8.0 (Thermo Fisher Scientific). Cas9 mRNA and sgRNA concentrations used are listed above. Embryos from each donor female were kept separate during IVF, microinjection and embryo transfer. Similar conditions were used for the additional microinjection projects, except natural mating rather than IVF was used to obtain zygotes.

Multiplex PCR amplicon NGS sample generation. On-target and off-target primer pairs were designed using NCBI Primer-Blast[20] with the following modifications to the default settings: amplicon size, 200–300; T_m min, 59 °C; T_m max, 61 °C; pair specificity against appropriate genome (reference assembly). If primers could not be designed to generate a unique amplicon in the 200–300-bp range, the repeat filter was turned off, "avoiding low-complexity regions" was

unselected, and the amplicon size was broadened to 150–400 bp. Primers were not permitted to reside within a 50-nt region to either side of the expected sgRNA cut site to ensure coverage of varied deletion sizes. All on- and off-target primer pairs associated with each sgRNA were ordered as 50 μM RxnReady stocks (Integrated DNA Technologies). Multiplex PCR (one reaction/sgRNA per animal) was carried out with a polymerase kit (Qiagen; 206143) according to the manufacturer's protocol. A subset of early projects used primer sets that amplified each on- and off-target region individually with Advantage GC 2 polymerase mix (Clontech; 639119), and reactions were pooled after PCR. A single unrelated wild-type animal control (appropriate species and strain) was routinely included to account for locus-specific single-nucleotide polymorphisms (SNPs), insertions/deletions (indels), and problematic sequences not associated with CRISPR editing (data not included). When wild-type animals were not included, off-target-negative littermates served as negative controls and provided confirmation of reference sequence accuracy. Each multiplex PCR reaction (or reactions, if two sgRNAs were used) was purified with a DNA Clean & Concentrator-5 kit (Zymo Research; D4004) and eluted in 25 μl of H$_2$O. Amplicon DNA concentrations were obtained with a Qubit high-sensitivity kit (Invitrogen; Q32854). For single sgRNA knockout and some knock-in animal models, all G0 mosaic founders were screened by targeted amplicon deep sequencing. For dual sgRNA knockout and the remainder of knock-in animal models, G0 animals were first analyzed by PCR/gel electrophoresis or droplet digital PCR (Bio-Rad), and only G0 mosaic founders positive for the desired dropout deletion or intended mutation were screened by targeted amplicon NGS.

Preparation of libraries and deep sequencing. For library preparation of each sample, 66 ng of the PCR products were used with the Ovation Library System for Low Complexity Samples kit (NuGEN; 9092-256). Sequencing of the libraries was completed with an Illumina MiSeq or HiSeq 2500 in rapid mode and 200-cycle single-end runs with V2 chemistry reading dual barcodes.

Analysis of CRISPR on- and off-targets using NGS data. Sequencing reads were aligned to the genome (GRCm38/mm10 for mouse, RGSC 5.0/Rn5 for rat) using GSNAP[21] as packaged in gmap-2014-11-14 with the following options: -m 5 -i 1 -N 1 -B 5 --split-output=alignment/gsnap -E 4 -n 10 -w 200000 --quality-protocol=sanger --format=sam -t 18. We used only the uniquely mapped reads for downstream analysis. We computed indel allele frequencies by counting every type of indel with a specific start position and length. Only reads that fully crossed a 51-bp window around the predicted target site were counted. Reads containing only base mismatches that were probably introduced during PCR or sequencing were counted as wild-type reads. Unique mutant indel alleles with >3% of total reads were flagged as potential off-targets by the analysis pipeline. Potential off-targets were then analyzed by visual inspection in IGV and were considered true off-targets only if the indel occurred at the expected position upstream of the PAM site. Mutant reads with a frequency <3%, including reads with additional SNPs probably resulting from PCR or the sequencing reaction in addition to an indel, were pooled ("sum of alleles <3%" in Supplementary Figs. 6 and 7). To identify potential Cas9-induced SNPs, we visually inspected the wild-type reads in IGV, as SNPs would automatically be included in the wild-type bin.

Whole-genome sequencing. Spleens from the male (#188) and the two female parents (#187 and #232), as well as all viable e10.5 embryos (#180–186 and #229–231), were harvested and digested completely for genomic DNA extraction (DNeasy Blood & Tissue kit; Qiagen; 69506). Genomic DNA was sheared to lengths between 200 and 700 bp with Covaris (LE220), and fragments ranging between 400 and 600 bp were selected for the libraries, prepared from 100 ng of DNA using the Nano library kit (Truseq; FC-121-4001). Libraries were sequenced on Illumina HiSeq 4000 instruments with paired-end 1 × 150 bp read length at an average of 80× coverage. Sequencing reads were aligned to the mouse genome (GRCm38/mm10) with BWA 0.7.10 as follows: bwa mem -t 32 -M. The alignments were then processed through the GATK (Genome Analysis Toolkit) pipeline to produce a joint genotyped VCF file. All indel events with a GQ quality score of at least 80 in one of the offspring were then used as targets for our CRISPR analysis workflow described above, using a 21-bp window, and indel allele frequencies in each animal were computed. Indels present in at least one offspring and absent in parental mice were retained.

Cell culture. Hepa1-6 cells (ATCC; CRL-1830; derived from C57L/J mice[22]) were cultured in high-glucose DMEM with 10% FBS (Seradigm; 1500-100), 10 mM L-glutamine, penicillin (100 U/mL) and streptomycin (100 μg/mL) (Gibco; 15140-122) and used for mouse gRNA studies. Rat-2 cells (ATCC; CRL-1764; derived from Fisher rat embryo[23]) were cultured in RPMI media with 10% FBS (Seradigm; 1500-100), 10 mM L-glutamine, penicillin (100 U/mL) and streptomycin (100 μg/mL) (Gibco; 15140-122) and used for rat gRNA studies. For the Cas9 variant study transfections, 3 × 10^5 cells were nucleofected (Lonza Nucleofector 4D) with 1 μg of pRK-Cas9 (or variants eSpCas9(1.1) and/or SpCas9-HF1, cloned in the same configuration as wild-type Cas9 to ensure accurate comparison) and 0.5 μg of sgRNA plasmids using solution SF + program EN-138 or solution SG + DS-189 for Hepa1-6 or Rat-2 cells, respectively. Cells were recovered after nucleofection

in 100 μl of RPMI media for 30–60 min before being seeded in six-well plates. Genomic DNA from cells was harvested with a Quick-gDNA microprep kit (Zymo Research; D3021) 5 d after nucleofection.

Target-enriched GUIDE-seq. TEG-seq is comparable to GUIDE-seq[2]. Both methods use protected short double-stranded DNAs to tag break sites in the genome. These tags are then used as universal primer sites for amplification and NGS mapping of the sequences flanking the break site. The TEG-seq protocol uses amplification and sequencing primer design that makes it specifically compatible with the Ion Torrent line of NGS platforms. Comparisons to data from Illumina-based GUIDE-seq experiments suggest that both methods yield similar results (P.-Z. Tang, B. Ding, L. Peng, V. Mozhayskiy, J. Potter and J.D. Chesnut, unpublished observations). TEG-seq was performed by Thermo Fisher. Briefly, 150,000 Neuro-2a mouse cells (ATCC; CCL131; derived from the A/J strain[24]) were transfected with Cas9 protein (1 μg) and a full-length synthetic *Pnpla3* sgRNA (10 pmol; Synthego) complexed as ribonucleoprotein along with a dsTag (1.25 pmol) using the Neon electroporation system and incubated for 3 d, after which genomic DNA was isolated (PureLink Genomic DNA; Thermo Fisher) and subjected to the TEG-seq/GUIDE-seq procedure. We used synthetic sgRNA for TEG-seq because we have found it to be more efficient than in vitro–transcribed sgRNA for ribonucleoprotein electroporation. Amplicon reads were aligned to the reference mouse genome (GRCm38/mm10). The mapped reads were further processed through Motif-Search, a plugin software for off-target search and read counts (Thermo Fisher). For easier comparison, we normalized total reads from the different samples by using reads per million (RPM): total reads from the sample multiplied by 1 million and divided by the total number of mapped reads of the NGS run. The off-target candidates with RPM > 1 were subjected to targeted amplicon resequencing (AmpliSeq). Primers flanking the cleavage sites and used for AmpliSeq are listed in Supplementary Table 3. Off-targets were confirmed by detection of either the dsTag or the presence of indels > 3 bp in the expected position upstream of the PAM site.

Statistical analysis. We used Prism 6 (GraphPad software) for all analysis. All box-and-whisker plots (Figs. 1c–f and 2c, Supplementary Figs. 5 and 15a) show median, minimum–maximum range, and first, second, third and fourth quartiles. Dot plots (Fig. 1g,h, Supplementary Figs. 9a and 10) show the mean value and s.e.m. All data points are shown. For Supplementary Fig. 9a, average on-target values were calculated in Microsoft Excel for each sgRNA from individual animal on-target efficiencies for that sgRNA. Each data point thus represents the average on-target efficiency for each sgRNA, and the means of the distributions represent the mean of average values. The distributions of on-target values for each individual sgRNA were not calculated. Two-tailed unpaired *t*-test was used to test for significance of difference between the means. For Fig. 1g,h and Supplementary Fig. 10, mean on-target values were compared in a pairwise manner using the two-tailed unpaired *t*-test for means being identical. For Fig. 2b and Supplementary Fig. 15b–d, the Pearson correlation coefficient (*r*) and r^2 value (coefficient of determination) were computed. Odds ratios (Supplementary Fig. 9b) were calculated by Fisher's exact test (two-tailed) for each specificity score, and the score with the highest odds ratio (66; OR = 18) was plotted in the figure.

Figures. Plots and graphs were generated in Prism 6 (GraphPad software). Tables included in Supplementary Figs. 2–4 and 12 were generated in Microsoft Word. IGV alignment data in supplementary figures were generated from snapshots of files loaded in IGV2.3.96. Radar plots were generated with ggplot2[25]. All figures were annotated and assembled in Adobe Illustrator CS6.

Plasmids. All plasmids used in this study are available from the corresponding author upon reasonable request.

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

Data availability. Whole-genome sequencing, TEG-seq and AmpliSeq data from this study are available through the NCBI Sequence Read Archive under accession number SRP124981. The BAM files for the 43 off-targets identified from whole-genome sequencing can be viewed as a track in the UCSC browser (http://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=max&hgS_otherUserSessionName=GenentechNoInsertion). All other data supporting the findings of this study are available within the paper and its Supplementary Information. Source data for Fig. 1g,h are available online.

References

18. Takeo, T. & Nakagata, N. *PLoS One* **10**, e0128330 (2015).
19. Hu, L.-L. et al. *Zygote* **20**, 361–369 (2012).
20. Ye, J. et al. *BMC Bioinformatics* **13**, 134 (2012).
21. Wu, T. D. & Nacu, S. *Bioinformatics* **26**, 873–881 (2010).
22. Darlington, G. J., Bernhard, H. P., Miller, R. A. & Ruddle, F. H. *J. Natl. Cancer Inst.* **64**, 809–819 (1980).
23. Topp, W. C. *Virology* **113**, 408–411 (1981).
24. Klebe, R. J. & Ruddle, F. H. *J. Cell Biol.* **43**, 69A (1969).
25. Wickham, H. *ggplot2* (Springer, New York, 2009).

nature research

Corresponding author(s):   Søren Warming (sorenw@gene.com)

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Amplicon NGS reads were aligned to the genome (GRCm38/mm10 for mouse, RGSC 5.0/Rn5 for rat) using GSNAP29 as packaged in gmap-2014-11-14 with the following options: -m 5 -i 1 -N 1 -B 5 --split-output=alignment/gsnap -E 4 -n 10 -w 200000 --quality-protocol=sanger --format=sam -t 18. WGS reads were aligned to the mouse genome (GRCm38/mm10) using BWA 0.7.10 as follows: bwa mem -t 32 -M . The alignments were then processed through the GATK pipeline to end up with a joint genotyped VCF file. All indel events with a GQ quality score of at least 80 in one of the offspring were then used as targets for our CRISPR analysis workflow described below with an exception that a 21bp window was used instead of a 51bp window.
Data analysis	InDel allele frequencies were computed by counting every type of indel with a specific start position and length. Only reads that fully cross a 51 bp window around the predicted target site are counted. Reads containing only base mismatches likely introduced during PCR or sequencing were counted as wildtype reads. Unique mutant InDel alleles with >3% of total reads were flagged as potential off-targets by the analysis pipeline. Potential off-targets were then analyzed by visual inspection in IGV and only considered true off-targets if the InDel occurred at the expected position upstream of the PAM site. Mutant reads with a frequency <3%, including reads with additional SNPs likely resulting from PCR or the sequencing reaction in addition to an InDel, were pooled (="sum of alleles <3%" in Supplementary Figs.5 and 6). To identify potential Cas9-induced SNPs, the wt reads were visually inspected in IGV2.3.96 as SNPs would automatically be included in the wildtype bin. Statistical analysis was done using the Prism (GraphPad) software. sgRNA design was done using Benchling.

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

TEG-seq, amplicon NGS validation, and WGS fastq files are available on NCBI Sequence Read Archive, SRA dataset SRP124981 (https://www.ncbi.nlm.nih.gov/sra/SRP124981). The BAM files for the 43 off-targets identified from WGS can be viewed as a track in the UCSC browser (http://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=max&hgS_otherUserSessionName=GenentechNoInsertion).
List of Figures with associated raw data:
Fig.1g,h (Excel file)
Supplementary Fig. 9a and 10 (Excel files)
Supplementary tables 1-3 provide raw data for Figs.1 and 2
No restrictions on data availability. Any additional data not included above that support the findings of this study are available from the corresponding author upon request

Field-specific reporting

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

☒ Life sciences   ☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design

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

Sample size	Sample size for G0 mosaic analysis was not pre-determined, we analyzed any rodents that were born and/or were positive for the intended on-target allele. Typically 150 zygotes were microinjected, however fluctuation in implantation and survival of embryos to term and beyond the post-natal period resulted in fluctuation of mice or rats analyzed across all projects. For G1 progeny, only animals from off-target+ G0 founders were analyzed and sample number depended on # of litters and litter size. For cell experiments, 3 individual transfections per condition were performed. Cas9 variants have been tested in cells previously and our work extends into animal embryos. To follow up the cell line work, we aimed to collect >20 mouse embryos to pool per replicate for 3 replicates and >15 individual rat embryos per condition (rat embryos cannot be cultured and therefore have to be implanted and collected later in development). We exceeded these sample sizes. For WGS, we aimed to analyze 10 individual embryos which required fully sequencing 2 litters, a common father and 2 individual mothers at 80X coverage which was an expensive and large effort. The 13 total samples with WGS data compares to other published data looking at 3 samples.
Data exclusions	Deep-sequencing data from un-related wildtype negative controls. When part of the analysis, this data was used to provide additional confidence in the identified off-targets. Data from off-target-negative litter-mates produced the same information. It is not possible to routinely include un-microinjected embryos as controls.
Replication	Cell line data was reliably reproduced. Most animal experiments did not require re-injection, but when same gRNA was injected on different days all replication of data were successful. For Pnpla3, multiple methods and experiments replicated the finding of a set of off-targets.
Randomization	Experiments were not randomized. Microinjections were performed by 1 of 3 injectionists over ~3 year period. For Cas9 variant comparison in mice, all 3 microinjectionists injected all 3 Cas9 types over 3 days. Embryos from each day that survived to blastocyst = 1 pooled sample/condition/day for a total of 3 samples. For Cas9 variant comparison in rats, microinjections occurred over multiple days until we had at least 15 embryos collected for each condition. For all routine animal model engineering projects randomization is irrelevant, embryos get injected with appropriate CRISPR reagents. Our data is a collection of our routine analysis prior to animal models reaching investigators for phenotypic study.
Blinding	Investigators were not blinded. However, all of our data was performed on NGS platforms requiring submission of samples with blinded ID numbers. NGS barcodes associate to those IDs and data is run through an unbiased analysis pipeline. We report on allele frequencies directly from our NGS software.

Materials & experimental systems

Policy information about availability of materials

n/a	Involved in the study
☒	☐ Unique materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☐	☒ Research animals
☒	☐ Human research participants

Antibodies

Antibodies used	HyperOva (cat#KYD-010-EX-X5) from CosmoBio
Validation	Used for superovulation of female mice, references: Takeo T, Nakagata N., Theriogenology. 2016 Sep 15;86(5):1341-6. PMID: 27242176 & Takeo T, Nakagata N., PLoS One. 2015 May 29;10(5):e0128330. PMID: 26024317

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Cell lines (Hepa1-6, CRL-1830; Rat-2 cells, CRL-1764; Neuro-2a, CCL-131) were obtained from ATCC.
Authentication	Cell lines not authenticated by short tandem repeats, but cell lines have been screened by other means. Our NGS data verified the species origin (mouse, rat) of the cell lines used.
Mycoplasma contamination	Cell lines were tested and found mycoplasma-negative prior to performing the experiments.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Research animals

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

Animals/animal-derived materials	Inbred mouse strains C57BL/6J, C57BL/6N, and BALB/cJ and Sprague Dawley rats were used. Genotyping of G0 founder or G1 progeny animals was performed on gDNA extracted from tail snip tissue samples from 1.5-3 week old animals sent from our animal care facility, therefore animal weights were not collected at any point and are irrelevant to our studies. Sex of the animals is also irrelevant to the genotyping outcome. For mouse Cas9 variant experiments, embryos were cultured until blastocyst stage. For rat Cas9 variant experiments, embryos were collected from pregnant females at e10.5. For the WGS experiment, C57BL/6J mice were used.

Method-specific reporting

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ Magnetic resonance imaging