Exosome-mediated horizontal gene transfer occurs in double-strand break repair during genome editing

Ryuichi Ono1, Yukuto Yasuhiko1, Ken-ichi Aisaki1, Satoshi Kitajima1, Jun Kanno1,2 & Yoko Hirabayashi1

The CRISPR-Cas9 system has been successfully applied in many organisms as a powerful genome-editing tool. Undoubtedly, it will soon be applied to human genome editing, including gene therapy. We have previously reported that unintentional DNA sequences derived from retrotransposons, genomic DNA, mRNA and vectors are captured at double-strand breaks (DSBs) sites when DSBs are introduced by the CRISPR-Cas9 system. Therefore, it is possible that unintentional insertions associated with DSB repair represent a potential risk for human genome editing gene therapies. To address this possibility, comprehensive sequencing of DSB sites was performed. Here, we report that exosome-mediated horizontal gene transfer occurs in DSB repair during genome editing. Exosomes are present in all fluids from living animals, including seawater and breathing mammals, suggesting that exosome-mediated horizontal gene transfer is the driving force behind mammalian genome evolution. The findings of this study highlight an emerging new risk for this leading-edge technology.
Since 2000, three types of genome editing technologies have been developed: zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and CRISPR-Cas9. Of these, CRISPR-Cas9 features not only the easiest construct design but also high double-strand break (DSB) efficiency; however, CRISPR-Cas9 can cause DSBs at unintended sites. In mouse zygotes, most DSBs introduced by CRISPR-Cas9 are repaired by nonhomologous end joining (NHEJ) without homologous DNA oligos for homologous recombination (HR). NHEJ-mediated repair of DSBs is prone to error, causing small indels. In 2015, we reported that DSBs introduced by CRISPR-Cas9 can be repaired by the capture of retrotransposon sequences, reverse-transcribed spliced mRNA sequences (RMDR: RT-product-mediated DSB repair) and CRISPR-Cas9 vector sequences (non-RMDR: non-RT-product-mediated DSB repair) in mouse zygotes. Most captured DNA sequences are truncated at their 5′ and 3′ ends. Short microhomologies (1–4 bp) between the captured DNA sequence and the DSB-introduced site were observed in only half of the cases, suggesting that both RMDR and non-RMDR proceed via NHEJ. RMDR and non-RMDR have also been observed in DSBs induced by CRISPR-Cas9 in NIH-3T3 cells.

The capture of DNA sequences was also observed at the DSB site introduced by the I-SceI restriction enzyme in Saccharomyces cerevisiae, a human hepatoma cell line and human monocytic leukemia cells and at naturally occurring DSB sites in Daphnia, Drosophila, and Aspergillus. Ty1 retrotransposon insertions into DSB sites were induced by I-SceI in HR-deficient S. cerevisiae. In the case of the hepatoma cell line LMH, I-SceI induced the insertion of truncated infected hepatitis B virus into DSB sites. Endogenous nucleotide sequence insertions were also induced by I-SceI in the human monocytic leukemia cell line U937. In Daphnia, Drosophila, and Aspergillus, greater than half of recently naturally gained introns originated from the repair of staggered DSBs.

These capture of unintentional DNA sequences at DSB sites might be an evolutionary driving force of mammalian genomes, including horizontal gene transfer. In this report, comprehensive analyses of DSB sites introduced by CRISPR-Cas9 in vivo and in vitro were performed to identify the relationships between DSB repairs and genome evolution and verify the risk for the leading-edge technology. Our results highlight exosome-mediated horizontal gene transfer, which occurs in DSB repair, during genome editing and represents a potential new risk for genome editing.

Results

Determination of indels by deep sequencing. First, we accurately determined the lengths of the indels introduced by the CRISPR-Cas9 system in vivo and in vitro by deep sequencing of PCR products amplified with two primers across the target DSB site (Fig. 1a).

We introduced DSBs at the Peg10 gene locus by transfecting NIH-3T3 cells with a CRISPR plasmid encoding both Cas9 and gRNA targeting the Peg10 gene and a PGK-Puro plasmid. After transient selection with puromycin, DNA was extracted from the cells, and PCRs were performed to amplify the region containing the DSB site introduced into Peg10 (Fig. 1a). Then, the PCR products were subjected to high-throughput next-generation sequencing analyses. Greater than half of the sequence reads contained ±1–2 bp indels as previously described. These populations may have been repaired by error-prone NHEJ as previously reported. Greater than 90% of the deletions (3–64 bp) exhibited microhomologies (1–4 bp) at the junction, suggesting that these deletions were also mediated by NHEJ. (Fig. 1b, Supplementary Fig. 1, Table 1, Supplementary Data 1).

Long insertions (>33 bp) were observed in 4% of sequence reads from DSB-induced NIH-3T3 cells (Fig. 1b, c). Greater than half of the long insertion sequences were derived from plasmid DNA (Fig. 1c, d). In total, 16% and 2% of long insertions were identical to mouse genomic DNA and mRNAs (Fig. 1c, d, Supplementary Data 2). These results are comparable to previous results obtained by gel extraction and subcloning/Sanger sequencing.

Capture of bovine and E. coli genomic DNA by horizontal gene transfer. One of the two novel findings of our high-throughput sequencing analyses is that 21% of the long insertions were derived from Escherichia coli genomic DNA. These sequences are identical to the E. coli K12 strain, suggesting that they are derived from contamination by the host E. coli genomic fragments used to amplify the CRISPR-Cas9 vectors (Fig. 1d). DNA sequences from E. coli with or without microhomologies were captured (Supplementary Fig. 2). Finally, 2% of the long insertions mostly (88%) consists of mouse repeats, including mouse short interspersed nuclear elements (SINEs), mouse long interspersed nuclear element-1s (L1s), mouse endogenous retroviruses, and mouse satellite repeats and simple repeats, whereas the remaining 12% of the insertions were derived from bovine genomic DNA, including bovine SINE and bovine satellite repeats (Fig. 1d–g, Table 2a, Supplementary Data 2).

Exosome-mediated horizontal gene transfer. Most of the inserted bovine DNA was derived from bovine satellite DNA sequences, such as BTSAT2, BTSAT3, and BTSAT4, and bovine SINE sequences, such as Bovc-tA2 (Table 2a). Dulbecco’s Modified Eagle Medium (DMEM) containing 10% fetal bovine serum (FBS) was used to culture NIH-3T3 cells. Thus, DNA or RNA from FBS in the form of cell-free DNA/RNA, including exosomal DNA/RNA might be the source of the bovine DNA sequences captured by the DSB sites in the cultured mouse cells.

To confirm the possibility of such horizontal gene transfer from the cell culture medium, we repeated these experiments using goat serum instead of FBS (Fig. 2). As noted in the experiments with FBS, mouse genomic DNA, mRNA, or plasmid DNA and E. coli genomic DNA were detected as expected (Fig. 2a, b). As expected, goat DNA sequences were captured in the DSB sites of NIH-3T3 cells (Fig. 2c–e, Table 2b, Supplementary Data 3). These data demonstrate that horizontal gene transfer can occur from the serum used in the culture medium. To clarify the origin of the captured bovine DNA sequences, i.e., whether these sequences arose from cell-free nucleic acids or nucleic acids in exosomes, we introduced DSBs by CRISPR-Cas9 in NIH-3T3 cell lines cultured with exosome-free 10% FBS. Bovine DNA sequences originating from cell-free nucleic acids should still be introduced at the DSB sites if horizontal gene transfer was mediated by cell-free nucleic acids. In contrast, a reduction in the insertion of bovine DNA sequences in the presence of exosome-free serum would indicate that trans-species gene transfer is mediated by exosomes. The insertion rates of endogenous mouse DNA sequences, vector sequences, and E. coli sequences were comparable in cells cultured with exosome-free FBS or normal FBS; however, most of the bovine DNA insertions were abolished by culture with exosome-free 10% FBS/DMEM (Figs. 3a–c and 4, Table 2c, Supplementary Data 4). Furthermore, exosomal RNA collected from FBS by ultracentrifugation and treated with or without RNase and DNase/RNase were comprehensively sequenced. Bovine satellite sequence RNAs and bovine...
Fig. 1 High-throughput indel identification to detect CRISPR-Cas9-induced mutations at the Peg10 gene locus in NIH-3T3 cells cultured in 10% FBS/DMEM. a Schematic representation of the sgRNA, Cas9, and primers. DSBs were repaired with deletions, mutations (small indels), and large insertions. The PCR products amplified using the primers were subjected to high-throughput sequencing. White box: UTR (untranslated region), yellow box: ORF1; blue box: ORF2. b The size of the original WT PCR product is presented as 0 bp. The lengths of the insertions are presented as the Plus number, and the lengths of the deletions are presented as the Minus number. Two independent high-throughput sequencing experiments were performed: FBS-V1 and FBS-V2. The total sequence reads of FBS-V2 were normalized to those of FBS-V1. c Distribution of indels at CRISPR-Cas9-induced DSB sites in NIH-3T3 cells (FBS-V1). Of the sequence reads, 35% were deletions, and 4% were large insertions (more than 33 bp; red region). d Of the large insertions (red region in c), 59% corresponded to partial sequences of the transfected plasmid DNA. An additional 16% and 2% of the reads were identical to mouse genomic DNA and mRNA sequences, respectively, and 21% of the large insertions corresponded to E. coli genomic DNA. The remaining 2% of the total reads are described in e (blue region). e 12% of the reads classified as others (blue region in d) were from Bos taurus (bovine), including genome, SINEs, and satellite DNA sequences. Structures of de novo inserted bovine sequences at the Peg10 loci (f, g). Both the post- and preintegration sequences are presented. The sgRNA sequence and the PAM sequences are presented in red and bold red characters, respectively. The black lines indicate the junction sites between pre- and postintegration sequences. The sequences in the blue boxes are overlapping microhomologies and are marked with black dotted lines. Each insertion was truncated at both the 5' and 3' ends. f Truncated Bov-tA1, BCS, and bovine SINEs were inserted with 6 and 1-bp microhomologies. g A truncated BTSAT3b, a bovine satellite, and a partial BERV2, bovine endogenous retrovirus, were inserted with a 1-bp overlapping microhomology.
retrotransposon RNAs were highly expressed in FBS under all the conditions, suggesting that bovine satellite sequence RNAs and bovine retrotransposon RNAs were within the exosomes (Supplementary Fig. 4).

These data support exosome-mediated trans-species gene transfer; however, it is possible that these horizontal gene transfer events were mediated by cell-free nucleic acids. Because exosomes and cell-free nucleic acids are reportedly present in all fluids from living animals, trans-species gene transfer events may also occur in mouse embryos in which DSBs are introduced by injection of CRISPR-Cas9 mRNA into zygotes. Thus, DNA was extracted from day 10 embryos in which CRISPR-Cas9 mRNA and Peg10 sgRNA were injected at the zygote stage and analyzed by high-throughput sequencing. One of 12 embryos (#20) captured BTAUL1, a bovine SINE (Fig. 5a, b, Supplementary Fig. 5, Supplementary Data 5). The KSOM medium used to culture the mouse zygotes contains bovine serum albumin (BSA) fraction V, which may contain exosomes or cell-free nucleic acids.

**Table 1** Most frequent sequences after CRISPR-Cas9 treatment

| Frequent sequences after CRISPR/Cas9 treatment (>2-bp deletions) | Number of sequence reads | Rate of sequences (%) |
|---------------------------------------------------------------|--------------------------|-----------------------|
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 15,263 | 2.07 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCACCATTTCCGCCGACACGTGTCCAGCGAAAGCCTC | 11,539 | 1.80 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCACCATTTCCGCCGACACGTGTCCAGCGAAAGCCTC | 11,517 | 1.76 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 7264 | 1.13 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 6252 | 0.98 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 5405 | 0.84 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 4395 | 0.77 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 3279 | 0.51 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 3116 | 0.49 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 2251 | 0.35 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 2153 | 0.34 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 1884 | 0.29 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 1602 | 0.25 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 1407 | 0.22 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 1301 | 0.20 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 1271 | 0.20 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 1180 | 0.18 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 1155 | 0.18 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 1143 | 0.18 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 1072 | 0.17 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 956 | 0.15 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 940 | 0.15 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 819 | 0.13 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 799 | 0.12 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 777 | 0.12 |
| AGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCGACACGTGTCCAGCGAAAGCCTC | 725 | 0.11 |

**Table 2** Number of sequence reads inserted into the Peg10 DSB locus

| Name of inserted DNA | Number of sequence reads | Percentage of total reads with long insertion |
|----------------------|--------------------------|-------------------------------------------|
| (a) Number of sequence reads, including murine repetitive elements and bovine DNA sequences, inserted into the Peg10 DSB locus (FBS VI)  |
| Mouse SINE | 13 | 0.59% |
| Mouse L1 | 83 | 0.39% |
| Mouse endogenous retroviruses | 84 | 0.56% |
| Mouse satellite repeats | 137 | 0.36% |
| Simple repeats | 90 | 0.36% |
| Bovine genome | 5 | 0.02% |
| Bovine SINE | 34 | 0.36% |
| Bovine satellite repeats | 15 | 0.06% |
| (b) Number of sequence reads, including murine repetitive elements and goat DNA sequences, inserted into the Peg10 DSB locus |
| Mouse SINE | 98 | 0.30% |
| Mouse L1 | 83 | 0.25% |
| Mouse endogenous retroviruses | 103 | 0.32% |
| Mouse satellite repeats | 319 | 0.98% |
| Simple repeats | 145 | 0.44% |
| Bovine genome | 10 | 0.03% |
| Bovine SINE | 52 | 0.16% |
| Bovine satellite repeats | 14 | 0.04% |
| (c) Number of sequence reads, including murine repetitive elements and bovine DNA sequences, inserted into the Peg10 DSB locus |
| Mouse SINE | 21 | 0.19% |
| Mouse L1 | 53 | 0.30% |
| Mouse endogenous retroviruses | 40 | 0.36% |
| Mouse satellite repeats | 88 | 0.80% |
| Simple repeats | 47 | 0.43% |
| Bovine genome | 0 | 0.00% |
| Bovine SINE | 0 | 0.00% |
| Bovine satellite repeats | 0 | 0.00% |

Discussion

In this report, we demonstrated that horizontal gene transfer assisted by CRISPR-Cas9 occurs in NIH-3T3 cells and mouse embryos. This phenomenon might be the driving force behind mammalian genome evolution. In fact, mice with fusions between the murine Peg10 gene and a bovine SINE were obtained (Supplementary Fig. 5). A number of possible trans-species horizontal gene transfer events have been reported in mammals. Chromosomal domains (chromatin organization modifier), a protein structural domain, are highly conserved in chromoviruses, and SCAN domains might originate from GYPSYDR-1 retrotransposons. Sirh-family genes, which are conserved in mammals, contain a gag-like domain from the Ty3/Gypsy-type retrotransposon of fugu fish22-26. Recently, in silico analyses demonstrated horizontal transfer of Bov (non-LTR retrotransposon from Bos taurus) and L1 retrotransposons (B. taurus) in eukaryotes27. In this study, we revealed that Bov and L1 were abundant in exosomes and that goat Bov was horizontally transferred into mouse NIH-3T3 cells. These data support that horizontal gene transfer events are mediated by exosomes.

CRISPR-Cas9 itself exhibits some propensity for inducing off-target mutations2. The DSBs produced by CRISPR-Cas9, whether on target or off target, were repaired by the capture of unintentional DNA sequences2. Although the risk of unintentional
**Fig. 2** Trans-species horizontal gene transfer at the Peg10 gene locus from the serum included in the culture medium. (a) Distribution of indels at CRISPR-Cas9-induced DSB sites in NIH-3T3 cells cultured using DMEM containing 10% goat serum instead of FBS. In addition, 38% of the sequence reads were deletions. Large insertions (greater than 33 bp) represented 4% of the total sequence reads (red region). (b) Here, 51% of the large insertions corresponded to partial sequences of the plasmid DNA that was transfected into the NIH-3T3 cells. In addition, 16% and 1% of the reads were identical to mouse genomic DNA and mRNA sequences (MM10), respectively. Moreover, 29% of the large insertions corresponded to *E. coli* genomic DNA. The remaining 3% of the total reads are described in (c) (blue region). (d) Partial goat DNA sequences from chromosome 28 were inserted with a 1-bp microhomology. (e) A truncated goat satellite DNA sequence was inserted with a 2-bp overlapping microhomology.

**Fig. 3** Repression of horizontal gene transfer at the Peg10 gene locus by using exosome-free FBS. (a) Distribution of indels at CRISPR-Cas9-induced DSB sites in NIH-3T3 cells cultured using DMEM containing 10% exosome-free FBS. Here, 37% of the sequence reads were deletions. Large insertions (greater than 33 bp) represented 4% of the total sequence reads (red region). (b) In addition, 55% of the large insertions corresponded to partial sequences of the plasmid DNA that were transfected into the NIH-3T3 cells. In addition, 17% and 2% of the reads were identical to mouse genomic DNA and mRNA sequences (MM10), respectively. Approximately 24% of the large insertions corresponded to *E. coli* genomic DNA. The remaining 2% of the total reads are described in (c) (blue region). (d) No bovine sequence reads were detected at induced-DSB loci.
insertions is greater than 4%, considerable efforts have focused on reducing off-target effects. The pair of CRISPR-Cas9 D10A (nickase) and a high-fidelity CRISPR-Cas9 nuclease reduce genome-wide off-target effects.28,29. These efforts hold promise because DSBs at off-target sites could be neglected. However, unintentional insertions at on-target DSB site cannot be suppressed by these off-target-reducing methods. Therefore, gene therapy using these genome-editing technologies may capture unintentional insertions. DSBs are typically repaired by NHEJ or HR. NHEJ is the predominant pathway in mammals,30,31 and Drosophila,32,33 whereas HR is the major pathway in S. cerevisiae.34. Another DSB repair mechanism, microhomology-mediated end-joining (MMEJ), repairs DSBs via the use of substantial microhomology. MMEJ uses microhomologies of 5–25 bp during the alignment of two broken ends, whereas NHEJ frequently proceeds through the annealing of short (1–4 bp) microhomologies.13. Most of the insertion sequences identified in the present study displayed short microhomologies (1–4 bp) or no microhomology with the introduced DSB site, suggesting that they were captured by NHEJ rather than MMEJ.

SCR7, an inhibitor of NHEJ, improves the efficiency of HR in genome editing.35,36 Increasing the efficiency of HR may be a key strategy to reduce the risk of unintended insertions.

**Methods**

**Animals.** All animal studies were conducted in accordance with the guidelines approved by the animal care committee of the National Institute of Health Sciences (No. 1004). The animal welfare committee of National Institute of Health Sciences (No. 539) approved the protocol. Animals had access to a standard chow diet and water ad libitum and were housed in a pathogen-free barrier facility with a 12L:12D cycle, as previously described.6

**Plasmid preparation.** The plasmids expressing both hCas9 and Peg10-ORF1-sgRNA were prepared as previously described. Briefly, Peg10-ORF1-sgRNA (5'-TGTCTCTACGTGCCGCAAATGGG-3') oligo DNA was ligated into the Bsal site of pX330 (http://www.addgene.org/42230/). Plasmid preparations were performed using the QiAGEN Plasmid Maxi kit (Qiagen, Hilden, Germany). All oligo sequences in this study are shown in Table 3.

**Production of hCas9 mRNA and Peg10-ORF1-sgRNA.** To produce the Cas9 mRNA, the T7 promoter was added to the Cas9 coding region of the pX330 plasmid (http://www.addgene.org/42230/). Plasmid preparations were performed using the QiAGEN Plasmid Maxi kit (Qiagen, Hilden, Germany). All oligo sequences in this study are shown in Table 3.

**Cell culture.** NIH-3T3 cells (RIKEN BRC-Cell Bank: RRBC-RCB2767) were cultured with 10% FBS (Invitrogen)/DMEM, 10% goat serum (Cosmo Bio, Tokyo, Japan)/DMEM or 10% exosome-free FBS (System Biosciences, Palo Alto, CA)/DMEM. pX330-Peg10-ORF1 plasmid and pGK-puro plasmid (500 ng each) were introduced into 2 x 10^5 NIH-3T3 cells/well in a six-well plate using Lipofectamine LTX reagent (Thermo Fisher Scientific, Waltham, MA). At 24 h after transfection, 10 μg/ml puromycin (Thermo Fisher Scientific, Waltham, MA) was added to the wells. Two days after transfection, the cells were collected, and genomic DNA was extracted as previously described.

**Exosome collection and exosome RNA isolation.** Exosomes were prepared by a stepwise centrifugation-ultracentrifugation method as described previously with minor modifications. Briefly, 1.4 ml FBS was concentrated at 10,000xg for 30 min to remove the cell debris and then centrifuged at 100,000xg for 70 min using a...
Table 3 List of oligo DNA sequences for sgRNA, IVT, and amplicon sequence

| sgRNA sequence | PCR primers for IVT | PCR primers for amplicon sequence |
|----------------|--------------------|----------------------------------|
| Peg10 ORF-sgRNA3 | TGTCTCTACGTGGCAAGTG | AATGATACGGCTCACTCAGTTGAGGTCTCTCTCACTTGGGAATGG |
| Peg10 ORF-IVT-F | AATGATACGGCTCACTCAGTTGAGGTCTCTCTCACTTGGGAATGG | AATGATACGGCTCACTCAGTTGAGGTCTCTCTCACTTGGGAATGG |
| Peg10 ORF-IVT-R | AATGATACGGCTCACTCAGTTGAGGTCTCTCTCACTTGGGAATGG | AATGATACGGCTCACTCAGTTGAGGTCTCTCTCACTTGGGAATGG |

Barcode Splitter program. The lengths of the filtered sequences were counted by the Compute sequence length program. Frequent reads were identified by FastQC.

**Sequence analyses.** PCR products longer (>32 bp) than WT were analyzed by the BLASTN program from the NCBI server (http://www.ncbi.nlm.nih.gov/BLAST/) and the CENSOR program from the GENETIC INFORMATION RESEARCH INSTITUTE (http://www.girinst.org/censor/index.php). The reference sequences are *Mus musculus* (mouse); MM10 (NCBI) for mouse genome, GRCm38 (NCBI) for mouse cDNA and ENSMUST (Ensembl) for mouse mRNA; *B. taurus* (bovine): bovine cDNA (NCBI) for bovine genome; *C. porcellus* (pigs): SDRat1.0.0 (NCBI) for goat genome; and *E. coli* str.K-12 (NCBI) for *E. coli* genome. These data were obtained from the NCBI (https://www.ncbi.nlm.nih.gov/genome/) and Ensembl (https://asia.ensembl.org/info/data/fpl/index.html) databases.

**Data availability**

The authors declare that the data supporting the findings of this study are available within the article and its Supplementary Information files.

Received: 1 November 2018 Accepted: 11 January 2019
Published online: 08 February 2019

**References**

1. Gaj, T., Gersbach, C. A. & Barbas, C. F. 3rd ZFN, TALEN, and CRISPR-Cas-based methods for genome engineering. *Trends Biotechnol.* 31, 397–403 (2013).
2. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826 (2013).
3. Wang, H. et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910–918 (2013).
4. Ono, R. et al. Double strand break repair by capture of retrotransposon sequences and reverse-transcribed spliced mRNA sequences in mouse zygotes. *Sci. Rep.* 5, 12281 (2015).
5. Li, W., Tucker, A. E., Sung, W., Thomas, W. K. & Lynch, M. Extensive, recent nucleotide sequence insertions derived from distant regions of the genome. *Proc. Natl. Acad. Sci. USA* 112, 7729–7734 (2015).
6. Teng, S. C., Kim, B. & Gabriel, A. Retrotransposon reverse-transcriptase-mediated decay enables intron gain in Drosophila. *Science* 326, 1260–1262 (2009).
7. Farlow, A., Meduri, E., Dolezal, M., Hua, L. & Schlotterer, C. Non-sense-mediated decay enables intron gain in Drosophila. *PLoS Genet.* 6, e1000819 (2010).
8. Zhang, L. Y., Yang, Y. F. & Niu, D. K. Evaluation of models of the mechanisms underlying intron loss and gain in *Aspergillus fumigatus*. *J. Mol. Evol.* 71, 364–373 (2010).
9. Moore, J. K. & Haber, J. E. Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. *Nature* 383, 644–646 (1996).
10. Wang, H. et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910–918 (2013).
11. Ono, M. et al. Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. *Proc. Natl. Acad. Sci. USA* 111, 7729–7734 (2014).
12. Zhang, L. Y., Yang, Y. F. & Niu, D. K. Evaluation of models of the mechanisms underlying intron loss and gain in *Aspergillus fumigatus*. *J. Mol. Evol.* 71, 364–373 (2010).
13. Moore, J. K. & Haber, J. E. Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. *Nature* 383, 644–646 (1996).
14. Bill, C. A. & Summer, J. Genetic DNA double-strand breaks are targets for headpavadin DNA integration. *Proc. Natl. Acad. Sci. USA* 101, 11135–11140 (2004).
15. Onozawa, M. et al. Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. *Proc. Natl. Acad. Sci. USA* 111, 7729–7734 (2014).
16. Yang, Z. et al. Fast and sensitive detection of indels induced by precise gene targeting. *Nucleic Acids Res.* 43, e59 (2015).
17. McVey, M. & Lee, S. E. MMEJ repair of double-strand breaks (director’s cut): deleted sequences and alternative endings. *Trends Genet.* 24, 529–538 (2008).
22. Ono, R. et al. A retrotransposon-derived gene, PEG10, is a novel imprinted gene. *RMC Bioinformatics* 7, 474 (2006).

23. Ono, R. et al. Identification of a large novel imprinted gene cluster on mouse proximal chromosome 6. *Genome Res.* 13, 1696–1705 (2003).

24. Ono, R. et al. Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nat. Genet.* 38, 101–106 (2006).

25. Sekita, Y. et al. Role of retrotransposon-derived implanted gene, Rtl1, in the fet-o-maternal interface of mouse placenta. *Nat. Genet.* 40, 243–248 (2008).

26. Naruse, M. et al. Biological properties of extracellular vesicles and their physiological functions. *J. Extracell. Vesicles* 4, 27066 (2015).

27. Ivancevic, A. M., Kortschak, R. D., Bertozzi, T. & Adelson, D. L. Horizontal transfer of BovB and L1 retrotransposons in eukaryotes. *Genome Biol.* 19, 85 (2018).

28. Ran, F. A. et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389 (2013).

29. Kleinsteiver, B. P. et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495 (2016).

30. Takata, M. et al. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* 17, 5497–5508 (1998).

31. Rebuzzini, P. et al. New mammalian cellular systems to study mutations introduced at the break site by non-homologous end-joining. *DNA Repair* 4, 546–555 (2005).

32. Preston, C. R., Flores, C. C. & Engels, W. R. Differential usage of alternative pathways of double-strand break repair in Drosophila. *Genetics* 172, 1055–1068 (2006).

33. Beumer, K. J. et al. Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. *Proc. Natl. Acad. Sci. USA* 105, 19821–19826 (2008).

34. Jeggo, P. A. DNA breakage and repair. *Adv. Genet.* 38, 185–218 (1998).

35. Chu, V. T. et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* 33, 543–548 (2015).

36. Maruyama, T. et al. Increasing the efficiency of precise genome editing with CRISPR/Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* 33, 538–542 (2015).

Acknowledgements

The authors thank Toshiko Momiyama, Erika Tachihara, Miki Uchiyama, and Yusuke Furukawa for excellent technical assistance. This work was supported in part by the Research on Regulatory Science of Pharmaceuticals and Medical Devices (18bk01040730103) and the Research on Development of New Drugs (18ak010109j001) from Japan Agency for Medical Research and Development (AMED), Health Sciences Research Grants from the Ministry of Health, Labor, and Welfare, Japan (H30-KAGAKU-IPAN-002, H30-KAGAKU-SHITEI-001), and JSPS KAKENHI (18K19183 and 18K19315) to R.O.

Author contributions

R.O. conceived of the study. R.O., K.A., S.K., J.K., and Y.H. participated in the experimental design. R.O. performed most analyses. Y.Y. produced mutant mice by CRISPR/Cas. R.O. wrote the manuscript. All authors read and approved the final manuscript.

Additional information

Supplementary information accompanies this paper at https://doi.org/10.1038/s42003-019-0300-2.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

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

© The Author(s) 2019