



Original Article

Diverse unintended on-target mutations induced by zygote genome-editing using CRISPR/Cas9 system

Toshime Igarashi, Yukuto Yasuhiko, Ryuichi Ono, Erika Tachihara, Miki Uchiyama, Atsuya Takagi, Yu Takahashi, Makiko Kuwagata and Satoshi Kitajima

Division of Cellular & Molecular Toxicology, Center for Biological Safety & Research, National Institute of Health Sciences, 3-25-26 Tono-machi, Kawasaki-ku, Kawasaki, Kanagawa 210-9501, Japan

(Received September 27, 2021; Accepted October 1, 2021)

ABSTRACT — With the advent of the CRISPR/Cas9 system, genome editing in various fields is advancing. Unintended mutation in off-target regions is a major problem of genome editing using the CRISPR/Cas9 system, and it is being reviewed. However, we found a high frequency and various unintended mutations in the “on-target” region when we generated a “knock-in” mouse with point mutation using this technique to develop a supernumerary rib model. Additionally, an inserted sequence of unknown origin was observed. Furthermore, these mutations were transferred to the next generation, even if tandem knock-in or large deletions occurred. These strongly suggest that a proper selection that meets the purpose is essential when considering the safety of foods and medicines using the genome-editing technology.

Key words: CRISPR/Cas9 system, Cas9 protein, Genome editing, Zygote electroporation, Knock-in, Mosaic

INTRODUCTION

Genome editing using the CRISPR/Cas9 system is simple and highly efficient and rapidly becoming widespread in various fields, including biology, medicine, and food development. In toxicology, this technology should greatly promote the development of genetically modified animal models. Developing gene knock-out or knock-in animals helps elucidate the molecular basis of the mechanism of toxicity expression because it enables us to approach causal relationships at the molecular level rather than correlations in biological phenomena. In rodent developmental toxicity studies, “supernumerary rib” (SNR) has been discussed for its toxicological significance due to its high spontaneous frequency. Recently, altered expression of the Hox genes has been indicated in 5-fluorocytocine-induced thoracolumbar SNR (Kumamoto *et al.*, 2020). The involvement of mecha-

nisms related to abnormal region determination along the anterior-posterior body axis has been suggested; however, many remain unclear. To clarify its molecular basis, we applied this genome-editing technique to develop SNR model mice by modifying the H1 enhancer Hox binding site (HBS) of *Myf5* locus, a transcription factor controlling rib formation, into a snake-type sequence. Briefly, Hox10 impedes the Pax3-Hoxb6 complex to sit on the H1 enhancer in the murine lumbar region, impairing rib formation by *Myf5* suppression. Snake-type H1 enhancer harbors a T > C substitution in HBS, allowing persistent expression of *Myf5* under the existence of Hox10 (Vinagre *et al.*, 2010; Guerreiro *et al.*, 2013; Mansfield *et al.*, 2013). Therefore, we generated knock-in mice bearing snake-type HBS using the CRISPR/Cas9 system.

Additionally, the safety of CRISPR/Cas9 genome-editing technology is a crucial issue in toxicology. Although the off-target effects are the most focused in genome-

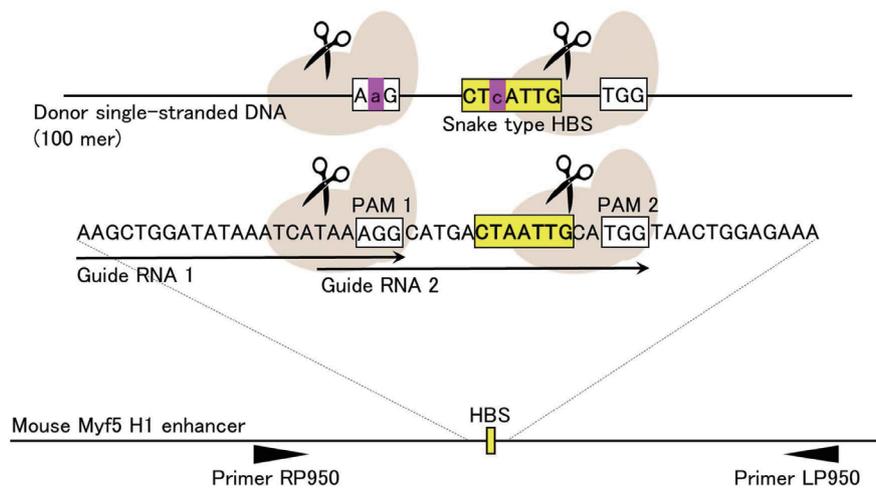


Fig. 1. Schematic of the murine Myf5 H1 enhancer and nucleotides designed for knock-in using the CRISPR/Cas9 system. The donor single-stranded DNA (100 mer) was designed with two nucleotide substitutions, same as the snake-type Hox binding site (HBS) sequence between the homology arm of murine Myf5 H1 enhancer. The guide RNAs are indicated with arrows. Arrowheads show the positions of the forward and reverse primers for PCR. CRISPR/Cas9 introduces double-strand breaks (DSB) at several bases upstream of the PAM site specified by the guide RNA. DSBs are repaired by nonhomologous end joining or homologous recombination (HR) with homologous DNA oligos using existing host cell systems. We aimed to knock-in donor DNA containing the snake-type HBS (T > C) sequence using the HR repair mechanism.

editing safety issues, concerns have been raised about so-called unintended “on-target” insertions (Ono *et al.*, 2015, 2019). We also found more on-target mutations than expected when generating knock-in mice. Thus, it was considered a great significant phenomenon in understanding the essence of this technology. Here we detail the diversity of the unintended on-target mutations and the transition to the next generation.

MATERIALS AND METHODS

C57BL/6NCrSlc female mice (3–4 weeks old) were purchased from SLC Japan (Shizuoka, Japan) to obtain zygotes. Recipient MCH(ICR)/Jcl female mice (10–12 weeks old) were purchased from CREA Japan (Tokyo, Japan). All mice were bred in a well-controlled, specific pathogen-free barrier facility. All experimental protocols involving the laboratory mice used in this study were reviewed and approved by the Committee for Proper Experimental Animal Use and Welfare, a peer-review panel established at the National Institute of Health Sciences (NIHS) with the experimental code No. 723. Additionally, the NIHS Safety Committee for Recombinant DNA Experiments approved the protocol used for genome-editing mice generation in this study (experimental code No. 1045).

To generate knock-in mice bearing T > C conversion in HBS, zygote genome editing was performed using AltR CRISPR/Cas9 system, which contained a donor single-stranded DNA (ssODN) (Integrated DNA Technologies, Inc., Coralville, IA, USA), and Genome Editor electroporator (BEX Co. Ltd., Tokyo, Japan). Ribonucleoprotein complexes of Cas9 nuclease and two guide RNAs (gRNAs), flanking HBS, consisted of crRNA and tracrRNA, alongside the donor ssODN, were introduced into one-cell fertilized eggs by electroporation (Fig. 1 and Fig. 2) (Hashimoto *et al.*, 2016). The CARD series medium and buffers (Kyudo Co. Ltd., Saga, Japan) were used for *in vitro* fertilization and culture. Each final concentration of Cas9 protein, crRNA, tracrRNA, ssODN was 100, 70, 133, 300 ng/μL, respectively (Nepa Gene Co., Ltd., Chiba, Japan).

The electroporated two-cell eggs were transplanted into the oviduct in pseudo-pregnant foster mothers and N0 mice were delivered. Furthermore, N1 mice were obtained by mating N0 with the wild type (WT), and then germline transmission was investigated in seven of 15 N0 mice. The target region was amplified by polymerase chain reaction (PCR) using tail DNA as a template, and TA cloning was performed to analyze the sequence. Amplified PCR products (expected 837 bp) were subcloned into Invitrogen pCRII vector (Thermo Fisher

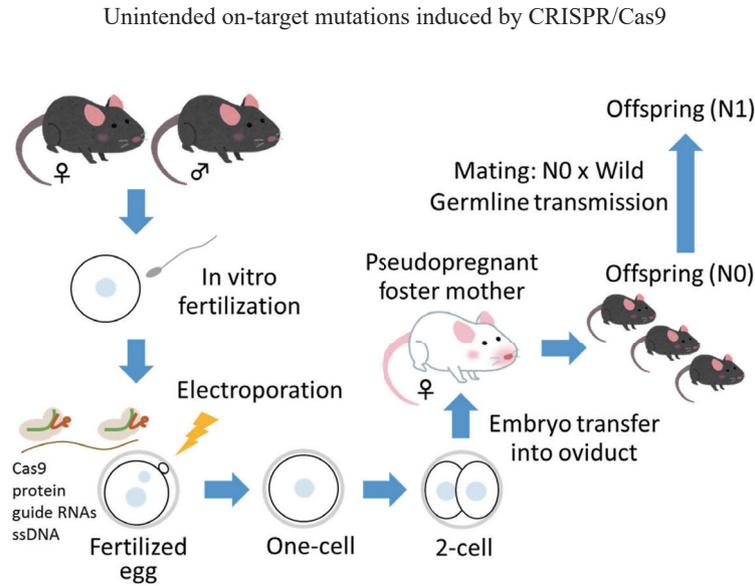


Fig. 2. Generation of genome-edited mice by zygote electroporation using CRISPR/Cas9 system.

Scientific K.K., Tokyo, Japan). Purified plasmid DNA was sequenced using M13 reverse primer, and more than ten clones were sequenced for each N0 and N1 mouse.

The sequences of each gRNA, ssODN, and PCR primer used in this study (Fig. 1) are as follows; two gRNAs, including specific 20-mer and PAM site (NGG) with underline, gRNA1, AAGCTGGATATAAAATCATAAAGG (20 + 3 mer); gRNA2, TAAAGGCATGACTAATTGCATGG (20 + 3 mer); and ssODN, including two nucleotide substitutions indicated by lower case letters, GTATGTTT-GTTGGAAAGGCCTTTAAGCTGGATATAAAATCAT-AAAaGCATGACcAATTGCATGGTAACTGGAGAAAT-GCTTTCTCTCTCTCTGGGGTGAAG (100 mer); primer LP950, CTGCTGGCAACATACCGAAA, primer RP950, CACAAGGAGCCATCTACTCT.

RESULTS AND DISCUSSION

Transplanting the two-cell eggs ($n = 117$) resulted in 15 offspring (N0 mice), which seems comparable to what was reported previously (Hashimoto *et al.*, 2016). Results of PCR genotyping and on-target sequences are shown in Fig. 3 and Table 1, respectively. As a result of detailed sequence analysis using N0 genomic DNA, our experiment found that on-target unintended mutations were significantly diverse and frequent in genome editing. The intended mutation was detected in six of 15 N0 mice and successfully knocked in the HBS with T > C conversion (Fig. 4 and Fig. 5). However, there were 35 unintended mutations across N0 mice. Table 2 shows the incidence of each mutation in N0 mice. There were point mutations

at double-strand break sites and more severe mutations. Large deletions of the on-target region (50–400 bases shorter than the WT) and tandem knock-in, i.e., multi-copy insertion of donor DNA HBS (up to 271 bases longer than the WT), were induced. Insertion of an irrelevant sequence was also observed in one case, N0 ID No. F6. Surprisingly, the inserted sequence, CAGAGAGTTTCAGAGAGTTTCAGAGAGTTTC, including three-times repeated motif, was 100% consistent with the *Salmo trutta* genome (see Table 1 footnote). However, the ten-base motif, CAGAGAGTTT, matched the sequences of many mouse genomes. Although this study did not reveal the origin of this sequence, the sequence derived from another species may be inserted since it has been reported that horizontal gene transfer occurred during double-strand break repair during genome editing (Ono *et al.*, 2019).

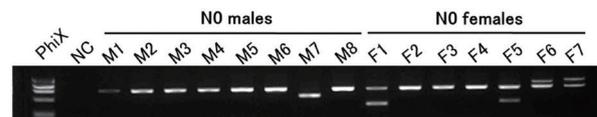


Fig. 3. Results of PCR genotyping using N0 tail DNA and LP950/RP950 primers. Compared with the expected size, 837b, large deletions in No. M7, F1, and F5, insertion in No. F6 and F7 were detected. They were confirmed by sequencing. The PCR products were subcloned into pCRII vector by TA cloning, and purified plasmids were sequenced using M13 reverse primer. PhiX: DNA size marker PhiX176/HaeIII, NC: negative control (DW).

Table 1. Sequence analysis of on-target region in N0 and N1 mice.

N0 ID	Alleles ⁽¹⁾	Types ⁽²⁾	Sequences ⁽³⁾	N0 ⁽⁴⁾	N1 ⁽⁵⁾
M1	3	UP	TATAAATCA TAAAGGCATGAC TAATxGCATGGTAACTGGAGAAA	8/10	NE
		UP	TATAAAT TAAAGGCATGAC TAATxGCATGGTAACTGGAGAAA	1/10	NE
		UP	TATAAATCA TAAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	1/10	NE
M2	4	WT	TATAAATCATAAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	0/10	2/24
		IM	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	6/10	11/24
		UP	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	1/10	0/24
M3	2	UD, -12 bp	TATAAATCA TAAAGGxxxxxxxxxxxxGCATGGTAACTGGAGAAA	3/10	11/24
		WT	TATAAATCATAAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	9/10	NE
		UP	TATAAATCATAAAGGCATGAC T TATT TGCATGGTAACTGGAGAAA	1/10	NE
M4	3	IM	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	8/10	NE
		UP	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	1/10	NE
		UP	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	1/10	NE
M5	4	UP	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	5/10	NE
		UP	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	1/10	NE
		UD, -9 bp; UP	xxxxxxxxTAAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	3/10	NE
M6	4	UD, -9 bp; UP	xxxxxxxxTAAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	1/10	NE
		IM	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	8/10	5/6
		UP (a)	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	1/10	0/6
M7	1	UP (b)	TATAAATCATAAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	0/10	1/6
		ULD, -86 bp	TATAxx	1/10	0/6
		ULD, -208 bp	TATAAATCAxx	10/10	8/22*
M8	2	UD, -8 bp	TxxxxxxxxTAAAGGCATGAC TAAxGCATGGTAACTGGAGAAA	5/10	NE
		UD, -17 bp	TATAAATCAxx	5/10	NE
F1	2	ULD, -420 bp	Txx	9/10	7/16*
		UD, -7 bp	TATxxxxxxxxAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	1/10	NE
F2	3	UD, -6 bp; UP	TATAAAT TGTAA TGGTAAAGxxxxxxxxGCATGGTAACTGGAGAAA	5/10	NE
		UD, -16 bp	TATAAATCATxx	4/10	NE
		UD, -16 bp; UP	TATAAATCATxx	1/10	NE
F3	3	UP	TATAAATCATAAAGGCATGAC C TATT TGCATGGTAACTGGAGAAA	4/10	NE
		UP	TATAAATCATAAAGGCATGAC C TATT TGCATGGTAACTGGAGAAA	1/10	NE
		UP	TATAAATCA TAAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	5/10	NE
F4	5	UD, -3 bp; UP	TATAAxxxxTAAAGGCATGAC TAATxGCATGGTAACTGGAGAAA	4/10	NE
		UD, -3 bp; UP	TATAAxxxxTAAAGGCATGAC TAATxGCATGGTAACTGGAGAAA	1/10	NE
		UD, -2 bp; UP	TATAAATxxTAAAGGCATGAC TAxTxGCATGGTAACTGGAGAAA	3/10	NE
F5	5	UD, -2 bp; UP	TATAA TxxTAAAGGCATGAC TAxTxGCATGGTAACTGGAGAAA	1/10	NE
		UP	TATAAATCA TAAAGGCATGAC TAxTxGCATGGTAACTGGAGAAA	1/10	NE
		IM	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	3/10	0/6
F6	4	UD, -400 bp (outside)	TATAAATCATAAAGGCATGAC CA... (48 bp)...#	3/10	4/6
		UP	TATAAATCATAAAGGCATGAC CA... (48 bp)...#	1/10	1/6
		UP	TATAAATCATAAAGGCATGAC CA... (48 bp)...#	1/10	0/6
F7	5	WT	TATAAATCATAAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	2/10	1/6
		IM	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	0/10	7/20
		IM	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	1/10	0/20
F6	4	UMI, +208 bp, 4-copy	TATAAATCATAAAGGCATGAC CA... (48 bp)...# TCATAAAGGCATGAC CAGAA TCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA... (79 bp)...	2/10	1/20
		UP	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	7/9	12/20
		WT	TATAAATCATAAAGGCATGAC TAATT TGCATGGTAACTGGAGAAA	0/10	4/18
F7	5	IM	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	9/10	3/18
		UP (a)	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	1/10	9/18
		UP (b)	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	1/10	9/18
F7	5	UMI, +271 bp, 5-copy	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA... (28 bp)... TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA... (17 bp)... TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA... (19 bp)... TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA... (44 bp)...	1/10	9/18
		UP (a)	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	0/10	1/18
		UP (b)	TATAAATCATAAAGGCATGAC CAATT TGCATGGTAACTGGAGAAA	0/10	1/18

(1) Total No. of detected alleles, including both WT and mutants.

(2) The types of detected sequences were described using the following abbreviation. WT: wild type, IM: intentional mutation, UD: unintended deletion, ULD: unintended large deletion (> 50bp), UMI, unintended multicopy insertion, UP: unintended point mutation.

(3) Diverse on-target mutations and motifs were highlighted with the following colors. Yellow: Hox binding site, pink: conversion, green: insertion, gray: deletion, black frame: PAM site. #: In N0 mice ID No. F6, an irrelevant sequence was inserted on the 3' side of the first HBS. The sequence of 31-base, including three-time repeats, was a 100% match to *Salmo trutta* genome assembly, chromosome 27 (GenBank Sequence ID: LR584424.1, range: 1748337-1748367, CAGAGAGTTTCAGAGAGTTTCAGAGAGTTTC).

(4) No. of detected clones per examined clones in TA cloning using N0 PCR products. One hundred and fifty clones and more were sequenced. Red: intended mutation, Blue: WT.

(5) No. of detected N1 mice per examined N1 mice, which are offspring of N0 × Wild C57BL/6N. Ninety-four N1 mice were examined by sequencing or PCR. *: Detected by PCR analysis. NE: not examined.

Unintended on-target mutations induced by CRISPR/Cas9

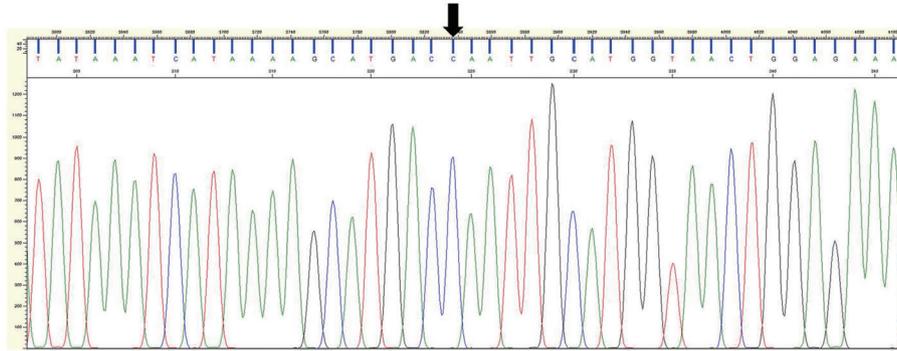


Fig. 4. Intended mutation in N0 mouse, ID No. M2. The arrow indicates the HBS T > C conversion. The knocked-in snake HBS was confirmed by sequencing.

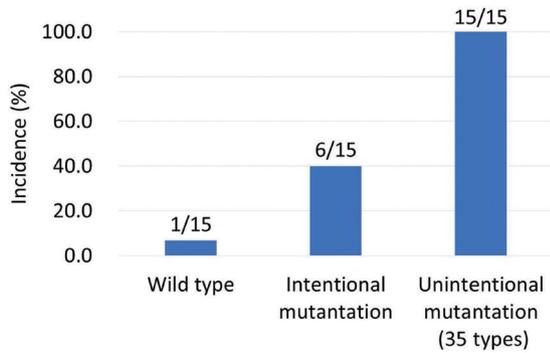


Fig. 5. Incidence of genotypes of alleles detected in N0 mice.

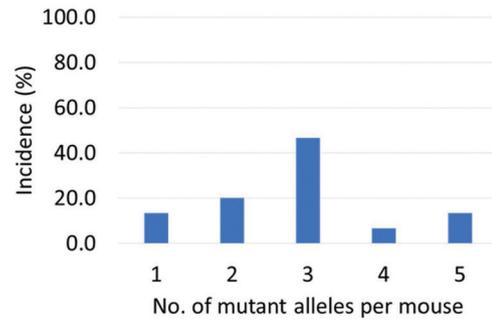


Fig. 6. No. of mutant alleles per N0 mouse, indicating mosaic genome. Mean ± SD: 2.87 ± 1.19

Most N0 mice did not have a WT allele and showed a mosaic genome, including 2–5 mutated sequences (Fig. 6). Up to five mutant alleles were generated, suggesting that the Cas9 activity remained until the eight-cell stage (Fig. 7). In contrast, nonmosaic knock-in mice have been generated using a similar method: genome editing by early pronuclear zygote electroporation (Hashimoto *et al.*, 2016). The high mosaic rate in our study may be related to the fact that the amount of Cas9 protein was twice that of the previous study.

Moreover, we conducted sequence analysis of N1 mice to investigate the germline transmission of these mutations detected in N0. The number of mutant allele types detected in most N1 mice was lower than that in N0 mice corresponding to each N1 mouse (Table 3). However, the knock-in type and unintended mutations, such as large deletion and tandem knock-in types, and insertion of fish genome-like sequence, were detected in N1 (Table 4). Therefore, under the conditions of this study, most on-target mutations were transmitted to the next

Table 2. Incidence of each unintended mutation in N0 mice.

Types	Incidence	N0 mice ID
Point mutation	10/15 (66.7%)	M1, M2, M3, M4, M5, M6, F3, F4, F5, F6
Small deletion (5–50 bases)	6/15 (40.0%)	M2, M5, M8, F1, F2, F5
Large deletion (> 50 bases)	4/15 (26.7%)	M6, M7, F1, F5
Tandem knock-in (multicopy insertion)	2/15 (13.3%)	F6*, F7

Incidence = No. of mice detected/total No. of mice examined.

All N0 mice (n = 15) were analyzed. *Insertion of the fish genome-like sequence was observed, suggesting horizontal gene transfer.

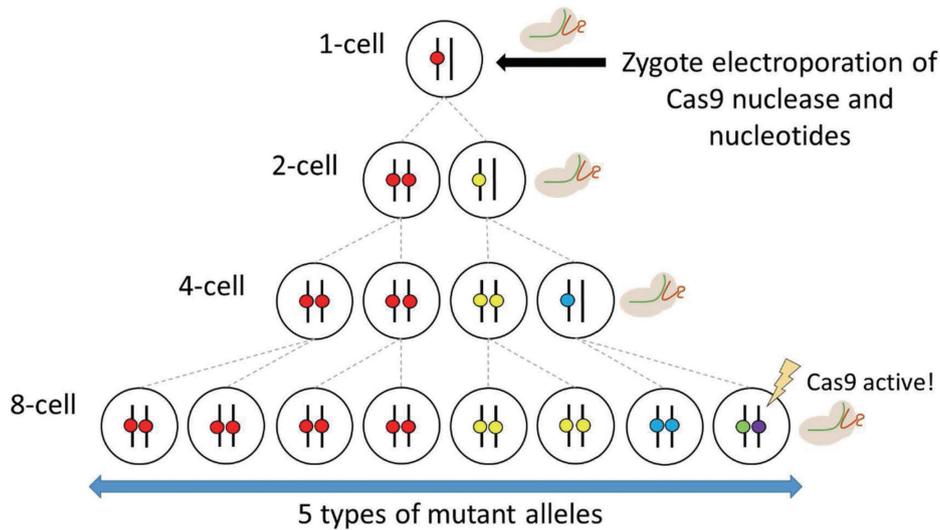


Fig. 7. Schema of the hypothetical model of mosaic genome induction. The activity of Cas9 nuclease until the eight-cell stage can induce five mutant alleles.

Table 3. Comparison of the number of mutant alleles detected in N0 and N1 mice.

N0 mice (n = 7) ID No.	Mutant alleles		Wild allele	
	N0	N1	N0	N1
M2	3	2	–	+
M6	3	2*	–	–
M7	1	1**	–	NE
F1	1	1**	–	NE
F5	5	3	–	–
F6	3	2	–	+
F7	2	4***	–	+

On the target region of N1, genomic DNA was sequenced after TA cloning. +: detected, –: not detected, NE: not examined. *Contains one allele with undetected point mutations in N0. **PCR analysis was only performed using flanking primers (LP950 and RP950). Large deletions were detected in N1 mice derived from N0 mice with ID No. M7 and No. F1. ***It contains two alleles with undetected point mutations in N0. N1 had three undetected mutant alleles in N0; however, each allele was only one clone.

Table 4. Incidence of each mutation transmitted to N1 mice.

Types of mutations	N0 ⁽¹⁾ Tail	N1 ⁽²⁾ transmission	N0 mice IDexamined ⁽³⁾
<i>Intentional mutation</i>			
Knock-in	6/15	4/5	<u>M2</u> , <u>M6</u> , <u>F5</u> , <u>F6</u> , <u>F7</u>
<i>Unintentional mutations</i>			
Small deletion (5–50 bases)	4/15	2/2	<u>M2</u> , <u>F5</u>
Large deletion (> 50 bases)	4/15	3/4	M6, <u>M7</u> , <u>F1</u> , <u>F5</u>
Tandem knock-in (multicopy insertion)	2/15	2/2	<u>F6*</u> , <u>F7</u>

(1) Incidence = (No. of N0 mice detected)/(No. of N0 mice [n = 15])

(2) Incidence = (No. of N1 mice detected)/(Total No. of N0 mice examined in each type)

(3) In underlined N0 mice, germline transmission of the corresponding mutation was observed as detection in N1 mice. *Insertion of the fish genome-like sequence was observed, suggesting horizontal gene transfer.

generation without exclusion. It was considered that N0 germline mosaicism contributed to the transmission of various mutations to the N1 generation. For the strains of each mutation in which germline transmission has been confirmed, future phenotypic analysis is expected to develop a useful SNR model.

Conclusively, the diversity of on-target unintentional mutations was revealed in generating knock-in mice using CRISPR/Cas9. Furthermore, such mutations were transmitted to offspring. These potential features of genome-editing technology strongly suggest that a proper selection is essential to ensure the safety of foods and medicines using the technology.

ACKNOWLEDGMENTS

The authors would like to thank Enago (www.enago.jp) for the English language review.

Conflict of interest---- The authors declare that there is no conflict of interest.

REFERENCES

- Guerreiro, I., Nunes, A., Woltering, J.M., Casaca, A., Nóvoa, A., Vinagre, T., Hunter, M.E., Duboule, D. and Mallo, M. (2013): Role of a polymorphism in a Hox/Pax-responsive enhancer in the evolution of the vertebrate spine. *Proc. Natl. Acad. Sci. USA*, **110**, 10682-10686.
- Hashimoto, M., Yamashita, Y. and Takemoto, T. (2016): Electroporation of Cas9 protein/sgRNA into early pronuclear zygotes generates non-mosaic mutants in the mouse. *Dev. Biol.*, **418**, 1-9.
- Kumamoto, T., Senuma, M., Todoroki, M., Kumagai, F., Imai, H., Suzuki, R., Ogawa, T. and Kuwagata, M. (2020): 5-Fluorocytosine induces fetal skeletal malformations in rats by altering expression of Homeobox genes. *Fundam. Toxicol. Sci.*, **7**, 97-103.
- Mansfield, J.H. (2013): Cis-regulatory change associated with snake body plan evolution. *Proc. Natl. Acad. Sci. USA*, **110**, 10473-10474.
- Ono, R., Ishii, M., Fujihara, Y., Kitazawa, M., Usami, T., Kaneko-Ishino, T., Kanno, J., Ikawa, M. and Ishino, F. (2015): Double strand break repair by capture of retrotransposon sequences and reverse-transcribed spliced mRNA sequences in mouse zygotes. *Sci. Rep.*, **5**, 12281.
- Ono, R., Yasuhiko, Y., Aisaki, K., Kitajima, S., Kanno, J. and Hirabayashi, Y. (2019): Exosome-mediated horizontal gene transfer occurs in double-strand break repair during genome editing. *Commun. Biol.*, **2**, 57.
- Vinagre, T., Moncaut, N., Carapuço, M., Nóvoa, A., Bom, J. and Mallo, M. (2010): Evidence for a myotomal Hox/Myf cascade governing nonautonomous control of rib specification within global vertebral domains. *Dev. Cell*, **18**, 655-661.