Response to "Unexpected mutations after CRISPR–Cas9 editing *in vivo*"

To the Editor: Recently, Schaefer *et al.*[1] reported that whole-genome sequencing (WGS) of two Cas9-treated, gene-corrected mice and a wild-type control mouse unveiled 1,397 single-nucleotide variations (SNVs) and 117 small insertions and deletions (indels) present commonly in the two Cas9-treated mice "but absent in the uncorrected control" and from a database of mouse SNVs and indels. There was essentially no sequence homology between the on-target site and these SNVs and indel sites, most of which lacked a protospacer-adjacent motif (PAM) sequence. Nevertheless, the authors concluded that these variations were caused by CRISPR–Cas9 without validating these unexpected off-target effects in an independent experiment *in vitro* or *in vivo*.

A major concern in Schaefer *et al.*[1] is the experimental design to identify *de novo* variants. Schaefer *et al.*[1] indicated that all three mice belong to the same inbred FVB/NJ strain and, possibly assuming no genomic variation in inbred strains, postulated *de novo* mutations in the Cas9-treated mice without sequencing their parents. Furthermore, an analysis of variants that are present in the control 'cohoused' mouse but absent in the two gene-edited mice, to determine initial genomic variation between the investigated mice, is missing.

Our reanalysis of their data (**Supplementary Notes 1** and **2**), however, shows that the two gene-corrected mice (F03, SRR5450997 and F05, SRR5450996) were genetically distant from the control mouse (WT, SRR5450998) and suggests that the SNVs and indels were not caused by CRISPR–Cas9 but were inherited from their unidentified parents. We used Strelka[2] and Mutect[3] employed in Schaefer *et al.*[1] to find SNVs and indels between two samples in any pairwise combination of F03, F05, and WT. We found the lowest number of sample-specific variants when comparing the two samples F03 and F05 (745 SNVs); while in comparisons including WT, the number of sample-specific variants was considerably higher (3,251 SNVs; **Supplementary Fig. 1a**). Strikingly, this was not only the case for variants present in F03 or F05 (and not in WT), but also for variants present in WT (and not in F03 or F05). To preclude a calling bias caused by the different sequencing depth of samples, we used the multisample variant caller Platypus[4] and found a similar number of variants as above (**Supplementary Fig. 1b**). Furthermore, most variants not present in WT were shared between F03 and F05 and were homozygous in both of the samples. It is virtually impossible that any mutagenic process causes such a pattern, since it would require the independent introduction of exactly the same variants into four copies of the genome. Finally, we focused on heterozygous variants in the three mice. Mice from inbred strains are generally not expected to contain substantial numbers of heterozygous variants, whereas mutagenic processes will in the vast majority cause heterozygous mutations. While there were indeed heterozygous variants in F03 and F05, the number of heterozygous variants was higher in WT. The largest fraction of heterozygous variants was shared between all three mice, which strongly suggests that these variants were inherited and not individually acquired in the investigated animals.

Target specificities of CRISPR systems have been extensively studied *in vitro*[5] and in animals and cell lines[6–13] (**Supplementary Note 1**). Schaefer *et al.*[1] did not find any off-target mutations in the two Cas9-treated mice at the top 50 most likely off-target sites with 3- to 4-nucleotide mismatches, in line with these previous reports. Given the specificity of CRISPR–Cas9, it is unlikely that Cas9 cleaves sites that differ by over 10 nucleotides from the on-target sequence. Altogether, these results demonstrate that there is considerable genetic heterogeneity between the three mice, with F03 and F05 being genetically closer to each other than to WT. We conclude that the data presented by Schaefer *et al.*[1] do not provide any evidence for CRISPR off-target effects *in vivo*.

Data availability. All reanalyzed data will be available upon request from the first author.

ACKNOWLEDGMENTS
J.-S.K. is supported by a grant from the Institute for Basic Science (IBS-R021-D1).

COMPETING INTERESTS
J.-S.K. is a cofounder of and holds stocks in ToolGen, Inc., a company focused on

therapeutic genome editing. All the other authors declare no competing financial interests.

Sang-Tae Kim[1,7], Jeongbin Park[2,3,7], Daesik Kim[4], Kyoungmi Kim[1], Sangsu Bae[5], Matthias Schlesner[2,6] & Jin-Soo Kim[1,4]

[1]Center for Genome Engineering, Institute for Basic Science, Seoul, South Korea. [2]Division of Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg, Germany. [3]Faculty of Biosciences, Heidelberg University, Heidelberg, Germany. [4]Department of Chemistry, Seoul National University, Seoul, South Korea. [5]Department of Chemistry, Hanyang University, Seoul, South Korea. [6]Bioinformatics and Omics Data Analytics, German Cancer Research Center (DKFZ), Heidelberg, Germany. [7]These authors contributed equally to this work.
e-mail: m.schlesner@Dkfz-Heidelberg.de or jskim01@snu.ac.kr

1. Schaefer, K.A. *et al. Nat. Methods* **14**, 547–548 (2017).
2. Saunders, C.T. *et al. Bioinformatics* **28**, 1811–1817 (2012).
3. Cibulskis, K. *et al. Nat. Biotechnol.* **31**, 213–219 (2013).
4. Rimmer, A. *et al. Nat. Genet.* **46**, 912–918 (2014).
5. Kim, D. *et al. Nat. Methods* **12**, 237–243 (2015).
6. Cho, S.W., Kim, S., Kim, J.M. & Kim, J.-S. *Nat. Biotechnol.* **31**, 230–232 (2013).
7. Cho, S.W. *et al. Genome Res.* **24**, 132–141 (2014).
8. Smith, C. *et al. Cell Stem Cell* **15**, 12–13 (2014).
9. Veres, A. *et al. Cell Stem Cell* **15**, 27–30 (2014).
10. Suzuki, K. *et al. Cell Stem Cell* **15**, 31–36 (2014).
11. Iyer, V. *et al. Nat. Methods* **12**, 479 (2015).
12. Tsai, S.Q. *et al. Nat. Biotechnol.* **33**, 187–197 (2015).
13. Frock, R.L. *et al. Nat. Biotechnol.* **33**, 179–186 (2015).