Response to Editas: Unexpected mutations after CRISPR-Cas9 editing *in vivo*

Kellie A. Schaefer[1,2] *, Wen-Hsuan Wu[3,4]*, Diana F. Colgan[1], Stephen H. Tsang M.D., Ph.D.[3,4], Alexander G. Bassuk M.D., Ph.D.[5,6] §, Vinit B. Mahajan M.D. Ph.D.[1,7] §

[1]Omics Laboratory, Stanford University, Palo Alto, CA.
[2]Interdisciplinary Graduate Program in Genetics, University of Iowa, Iowa City, IA.
[3]Jonas Children's Vision Care, and Bernard & Shirlee Brown Glaucoma Laboratory, Departments of Ophthalmology, Pathology & Cell Biology, Institute of Human Nutrition, College of Physicians and Surgeons, Columbia University, New York, NY 10032.
[4]Edward S. Harkness Eye Institute, New York-Presbyterian Hospital, New York, NY 10032
[5]Department of Pediatrics, University of Iowa, Iowa City, IA.
[6]Neurology, University of Iowa, Iowa City, IA.
[7]Byers Eye Institute, Stanford University, Palo Alto, CA.

*Contributed equally
[§]Co-corresponding Authors:
Vinit B. Mahajan, M.D., Ph.D., Byers Eye Institute, Stanford University, Palo Alto, CA, 94303. Phone: 650-721-6888. E-mail: mahajanlab@gmail.com

Alexander G. Bassuk, M.D., Ph.D., Department of Pediatrics, The University of Iowa, Iowa City, IA, 52242. Phone: 319-356-7726. E-mail: alexander-bassuk@uiowa.edu

We appreciate the interest of Editas Inc. in our observation and the questions they raise about the potential for off-target mutations by CRISPR. Restoration of sight in *Pde6b*^{rd1} mice was the primary outcome of our original study that began in 2015 to test CRISPR homology directed repair (HDR) of a single point mutation.[1] Off-target analysis was a secondary outcome reported in our Correspondence.[2]

In our study, only two of the eleven founders showed successful HDR.[2] Tissue from these two and a colony control underwent WGS. Thus, as Editas points out, the sample size in our report was small—one control and two cases. This number is nearly identical to that of Iyer et al.'s *Nature Methods* correspondence,[3] which is commonly cited to indicate Cas9 has limited off-target effects *in vivo*. Neither our study nor the Iyer et al. study used parental controls. Indeed, Iyer et al. (2015) states that, "To control for strain-specific variants, we also sequenced a C57BL/6J and a CBA animal from our breeding colonies." The reason for this approach is practical: injecting CRISPR-Cas9 requires multiple zygotes, typically gathered from many females mated with many males. In our case, 56 zygotes were harvested from six pregnant females bred to six stud males and injected with CRISPR-Cas9. Exact parentage is difficult to assess, due to this technical aspect as well as the highly inbred nature of this strain. We agree that future studies where off-targeting is determined by WGS should be designed with parental controls.

The Iyer et al. paper, which surveyed CRISPR-treated F1 of the hybrid mice for off-target mutations, focused on indels. In contrast, we looked for indels and SNVs. Additionally, Iyer et al. studied the off-target effects after non-homologous end-joining (NHEJ), while we studied off-targeting after homology directed repair (HDR). HDR requires a donor template, in our case a single-stranded oligonucleotide DNA molecule (ssODNA), which itself might be mutagenic, and even more so in combination with CRISPR-Cas9.[4,5] The difference in our results may reflect differences in gRNAs, technique etc., (as we addressed) but cannot be attributed solely to our use of colony controls.

There is significant heterozygosity observed between $F_0$3 and $F_0$5. Genetic drift is not something that could plausibly account for the observed heterozygosity, due to the experimental design: Based on our standard practice for murine transgenesis, a standard procedure was followed by ordering 3-week old oocyte donors and 8-week old stud males from Jackson Labs. We did not breed these mice in-house. All the stud males and oocyte donors were ordered within a few weeks of one another. In fact, this is what JAX recommends to avoid genetic drift issues as part of their Genetic Stability Program. These freshly ordered mice were used exclusively for the purpose of *rd1* repair and were not kept past 3-4 months of age. Based on the JAX order, it's likely that parents that produced both the stud and oocyte donor were probably siblings, as it is common practice to use sibling matings to generate a colony of inbred mice. Thus, $F_0$3 and $F_0$5 could essentially be considered clones of one another and would be expected to have a high degree of homozygosity. Instead, we observed extensive heterozygosity (Figure, Table). The heterozygosity in $F_0$3 and $F_0$5 cannot be parentally inherited.

Further support is provided by the highly inbred line used in our study.[5-7] Inbreeding leads to a reduction in heterozygosity within the population. In 1988, FVB/N mice (which are blind because of the *Pde6b*^{rd1}mutation) were imported from NIH to Dr. Taketo at The Jackson Laboratory. In 1991, these were re-derived at F50 into the foundation stocks facility at The Jackson Laboratory (FVB/NJ). There is no evidence for widespread SNVs between mice in this line. No heterozygosity has been described. In contrast, the Oey et al. paper cited in the Editas Inc. letter, which reported variation between littermates, is based on a line that is a C57BL/6J x C3H/HeJ cross. These mice carry the *agouti viable yellow* (A^{vy}) allele (this is why their mice show *agouti* coat colors and not black like the C57BL/6J strain). The number of backcrosses done in their colony is never reported.[8] Moreover, the

A^{vy} line is known to have a poor DNA-repair mechanism, and a high spontaneous cancer rate.[9,10] Hence, the colony used in Oey et al. is predisposed to SNVs and mutations. Table 2 of Oey et al. notes at least 1130 heterozygous variants shared by their two littermates, suggesting theirs is not a typical inbred line. An inbred, essentially clonal strain is not the same as a strain that was insufficiently backcrossed and crossed to a line predisposed to mutation. Moreover, in our observation, over 50% of the nearly 2035 total SNVs (339 unique to F03, 299 unique to F05, and 1,397 shared between the two) and over 30% of the over 160 total indels (47 unique to F03, 11 unique to F05 and 117 shared between the two) were reported at unexpected off-target sites, were read as heterozygous, and were absent in the control (see Figure 1 and Table 1). Again, heterozygous SNVs and indels should be an exceedingly rare event in this inbred line. Furthermore, the number of observed SNVs, if due to genetic drift, is estimated to take over 3.5 years (without any backcrossing) and would still be expected to be homozygous.

With regard to the relatedness between $F_0$3 and $F_0$5, the clonality between $F_0$3 and $F_0$5 can be discerned in our posted WGS data by the identity at all non-mutant call alleles. The WGS filtering pipeline in our Correspondence was not designed to determine all of the sequencing differences between the cases and controls. Many of the differences that would have reflected such genetic drift between cases and controls were rejected in the way our pipeline was designed. For example, nucleotides known to be commonly mutated in the germline were all rejected and did not appear in the final list of mutant genes (see Methods from the original Correspondence). If we were to assume long-standing genetic drift between the cases and the control, which are both from the original inbred line, we would expect these changes to be homozygous, and the most expedient way to eliminate variant calls that were due to this drift would be to add a filtering step that removes all homozygous calls. While this extra filtering step might lead to some false negative calls of true homozygous mutations, it would still leave over 1000 heterozygous mutations (which is more than 50% of the total mutations reported, Figure, Table). These heterozygous mutations cannot be explained by long standing differences between inbred cases and control, as such differences would be homozygous. Therefore, genetic drift does not account for the number of mutations, most specifically the level of heterozygosity observed, leading one to consider the source as CRISPR therapy intervention.

With regard to the differential sequence read depth between the cases and the colony control, when we originally designed the HDR study, we fully expected to observe little to no off-targeting in the CRISPR/Cas9 treated mice. The FVB/NJ control inbred line genome was already publically available at 50x coverage in the mouse genome project (http://www.sanger.ac.uk/science/data/mouse-genomes-project, ftp://ftp-mouse.sanger.ac.uk/REL-1303- SNPs_Indels-GRCm38/) based on a published WGS study.[5] However, we chose to sequence an available colony control to rule out any mutations that might be introduced because of differences in our local sequencing protocols and apparatus. Therefore, to save resources, we sequenced the control mouse at 30x coverage and the cases at 50x. We noted in the original correspondence that all mutation calls in the 50x sequenced cases had a read depth of at least 23x. For the 30x sequenced control, approximately 97% (2145/2210) of the wild-type reads were greater or equal to 20x covered. Of the remaining sites, 53/65 of wild-type reads were sequenced at greater than 15x. The remaining 12 mutation loci (7 SNV and 5 indels) reads were greater than 10x. It is possible that these few lower read loci are false positives. It is also possible that many of the reads in our cases that fell slightly below the 23x cutoff and were not called are actually false negatives, and that the true mutation rate is even higher than we reported. To secondarily test some of these loci, we performed Sanger sequencing for some of the mutations in the original Correspondence and have included more in the present Correspondence (Figure).

Concern was expressed that despite the fact our CRISPR-Cas9-treated mice were mosaics, there was high similarity between WGS read depths in the SNVs. While this could be explained by parental inheritance, this could alternatively be explained by Cas9/ssODNA introducing mutations during early embryonic development, specifically at the 1-, 2-, or 4-cell stages when levels of Cas9 are high. HDR may have occurred at a later stage in development resulting in a different degree of mosaicism. This could also account for the novel indels between the two animals (at regions not predicted by current algorithms), many of which are read as heterozygous (Table).

Are there other reasons we may have detected off-target mutations? Editas suggests the guide RNA was suboptimal; and this may be correct. We used the online software from Benchling (San Francisco, CA) to design several gRNAs, and achieved high on-target cleavage rate with only one *in vitro*. This one gRNA was used *in vivo*. Since we aimed to rescue sight by repair of a specific *rd1* sequence by HDR, our *rd1* specific gRNA had to target a relatively short sequence, and our sequence optimization options were limited. In contrast, for a gene-disruption strategy, use of non-homology end joining (NHEJ; which can target many regions across a gene) typically gives the flexibility to choose from far more gRNAs. Although, a less perfect gRNA might be expected to hit more off-target sites, it would still be predicted to be restrained to homologous sites. Instead, we observed mutations to sites that showed little homology to the gRNA.

The summary statements in our Correspondence reflect observations of a secondary outcome following successful achievement of the primary outcome using CRISPR to treat blindness in *Pde6b*^{rd1} mice. As the scientific community considers the role of WGS in off-target analysis, future *in vivo* studies are needed where the design and primary outcome focuses on CRISPR off-targeting. We agree that a range of WGS controls are needed that include parents, different gRNAs, different versions of Cas9, and different *in vivo* protocols. We look forward to the publication of such studies. Combined, these results will be essential to fully understand off-targeting and can be used to create better algorithms for off-target prediction. Overall, we are optimistic that some form of CRISPR therapy will be successfully engineered to treat blindness.

1. Wu, W.H. *et al*. CRISPR repair reveals causative mutation in a preclinical model of retinitis pigmentosa. *Mol. Ther.* **24**, 1388–1394 (2016).

2. Schaefer, K.A. *et al.* Unexpected mutations after CRISPR–Cas9 editing *in vivo*. *Nat Meth.* **14**, 547–548 (2017).

3. Iyer, V. *et al*. Off-target mutations are rare in Cas9-modified mice. *Nat. Methods.* **12**, 479 (2015).

4. Woolf T.M. *et al.* The stability, toxicity and effectiveness of unmodified and phosphorothioate antisense oligodeoxynucleotides in Xenopus oocytes and embryos. *Nucleic Acids Research*. **18**, 1763-1769 (1990).

5. Wong, K. *et al.* Sequencing and characterization of the FVB/NJ mouse genome. *Genome Biol.* **13***,* R72 (2012).

6. Taketo, M. *et al.* FVB/N: An inbred mouse strain preferable for transgenic analyses. *Proc. Natl. Acad. Sci.* **88**, 2065-2069 (1991).

7. Beck, J.A. *et al.* Genealogies of mouse inbred strains. *Nat Genet.* **24**, 23-25 (2000).

8. Oey, H. *et al.* Genetic and epigenetic variation among inbred mouse littermates: identification of inter-individual differentially methylated regions. *Epigenetics Chromatin.* **8**, 1-12 (2015).

9. Schwartz, A.G. Inhibition of spontaneous breast cancer formation in female C3H(Avy/a) mice by long-term treatment with dehydroepiandrosterone. *Cancer Res.* **39**, 1129-1132 (1979).

10. Hansen, L.A. Effect of the viable-yellow (A(vy)) agouti allele on skin tumorigenesis and humoral hypercalcemia in v-Ha-ras transgenic TGxAC mice. *Carcinogenesis.* **19**, 1837-1845 (1998).

Table. Total heterozygous off-target mutations in CRISPR treated mice.

	F03	F05	Shared Mutations
SNVs	910 (52% of total)	954 (56% of total)	675
Indels	58 (35% of total)	46 (35% of total)	37

Off-target mutations that passed all 3 pipelines were called "heterozygous" if reads were equal between the mutant allele and reference (+/- 10%).
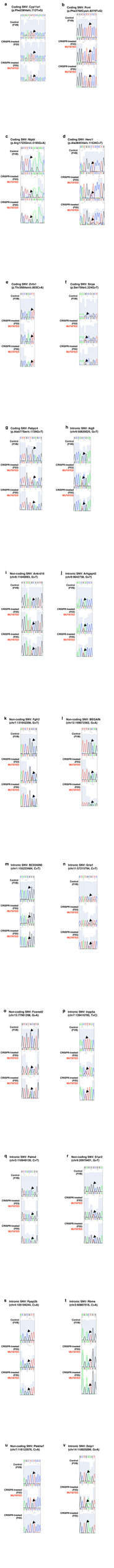
Figure. Sanger sequencing confirms heterozygous mutants detected by WGS in CRISPR-treated mice.