Deeper sequencing at unexpected CRISPR/Cas9 off-target sites *in vivo:* A response to Editas, Intellia, Beacon, ToolGen and others

Kellie A. Schaefer[1,2] *, Wen-Hsuan Wu[3,4]*, Benjamin W Darbro, M.D., PhD.,[5], 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

The authors have no direct or indirect financial connections to any CRISPR company or to any related companies.

A CRISPR/Cas9 pattern at mutation sites in the WGS data. There is broad consensus that CRISPR/Cas9 has off-target effects, but the best methods to address this *in vivo* are uncertain. We appreciate the interest of Editas Inc., Intellia Inc., ToolGen Inc., Beacon Genomics and others in our observation and the questions they raise about the potential for off-target mutations by CRISPR.[1] Editas, Intellia, ToolGen, and Beacon have suggested that the variation found in the CRISPR treated animals, compared to the non-treated animal, is simply the result of parental inheritance and not off-target CRISPR effects. The parents of F03 and F05 and the FVB control mouse were purchased through the JAX genetic stability program, all belonged to generation F95pF100-F95pF101, and they were never bred in our facility. To further address this question, we inspected the variants originally reported in our Correspondence, as well as newly generated variant calls from comparison of each animal's WGS to the mouse reference genome (mm10). In addition, we performed TOPO cloning followed by Sanger sequencing to detect multiple alleles. We detected sites where numerous alleles were present, more so than would be expected by simple inheritance and with a pattern consistent with DNA breakage followed by repair (examples in Figures 1-9). In response to ToolGen's criticism that these regions did not contain a PAM sequence, these heterozygous mutations were mostly within 7–10 bp adjacent to NGG or NGA nucleotide sequences, the preferred Protospacer Adjacent Motif (PAM), or recognition site, for the SpCas9.

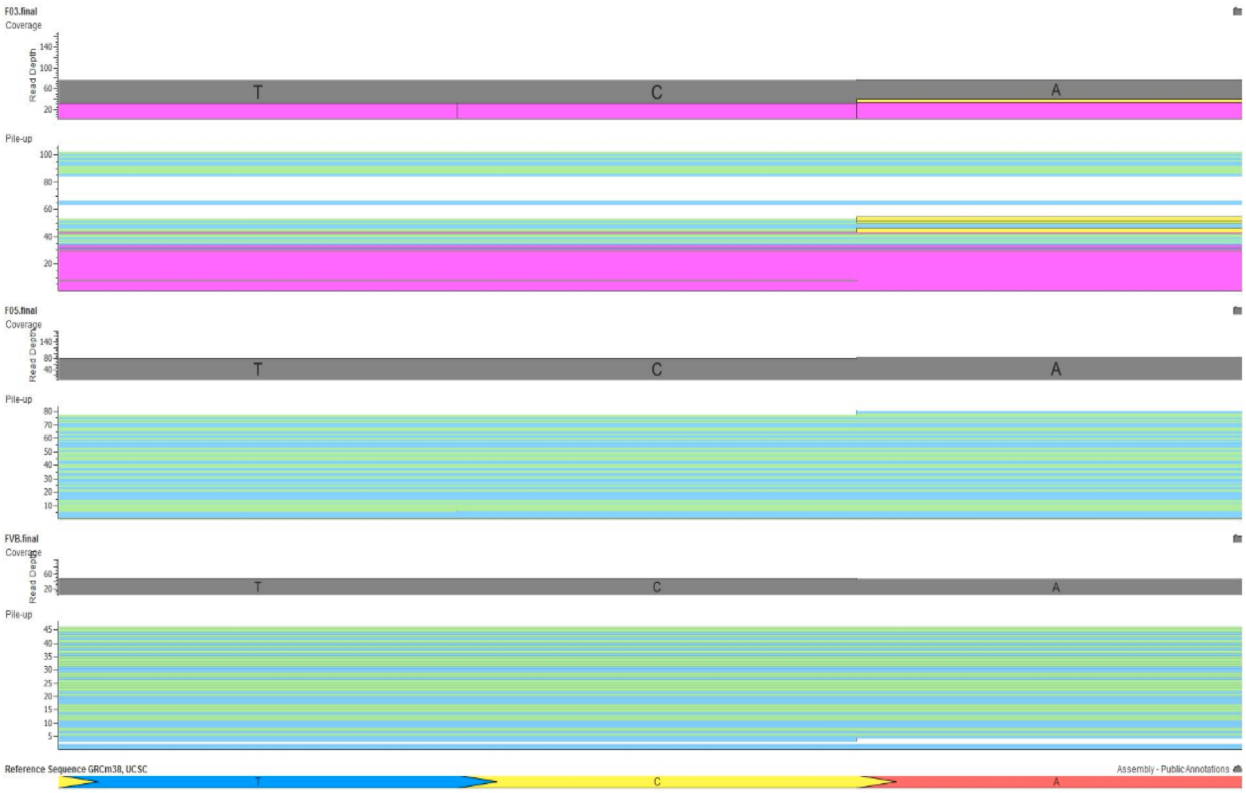


Figure 1. Whole genome sequence reveals more than 2 alleles at the chr1:86,357,000 – 86,357,003 deletion suggesting a CRISPR/Cas9-induced pattern of mutations. The $F_0$3 mouse has numerous indel calls (pink) as well as an A>C

SNV (yellow), while the $F_0$5 and FVB control mouse do not have multiple variants at this site (only wild-type). Images created using mouse BAM files and VarSeq (Golden Helix, Inc., Bozeman, Montana).

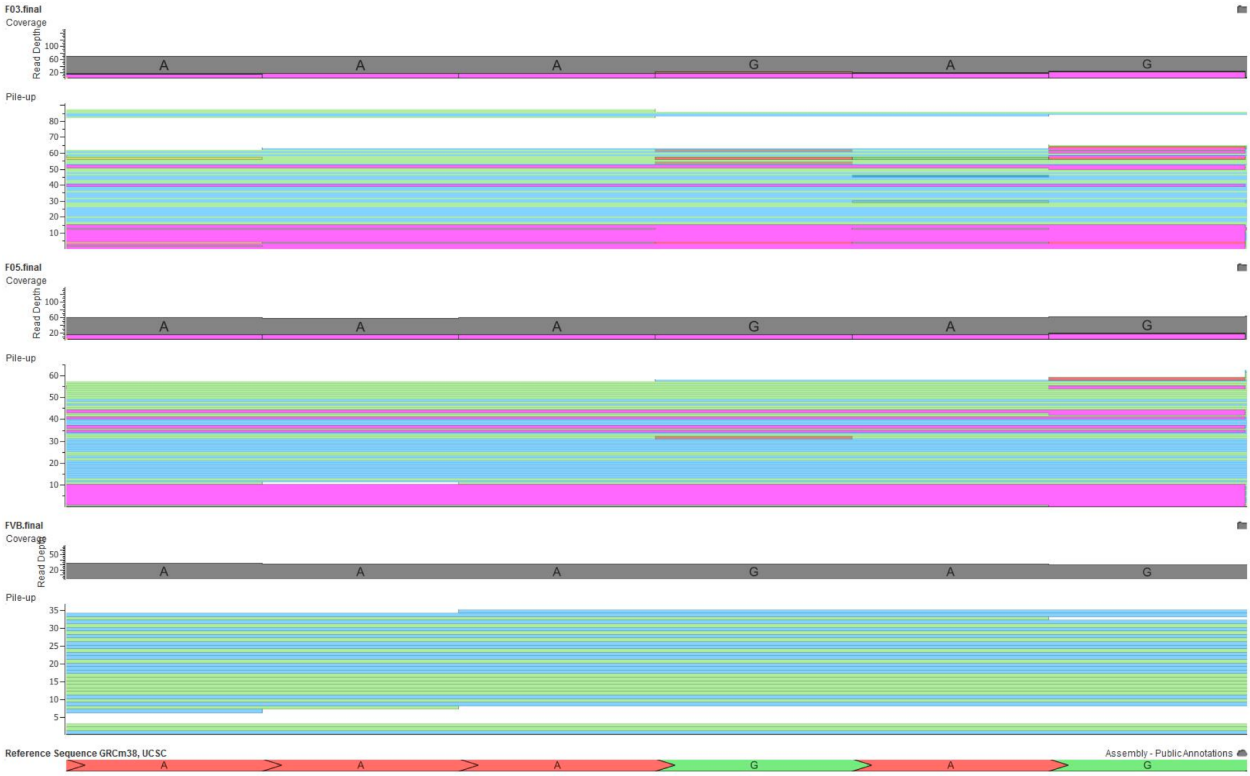
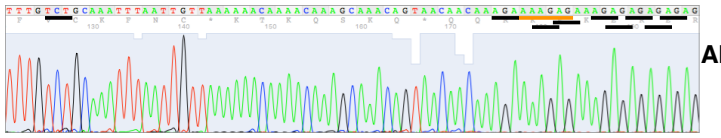
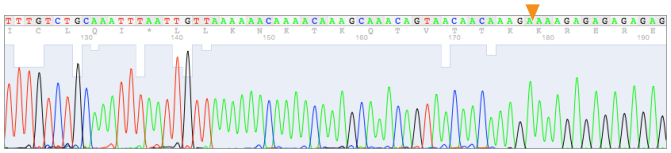


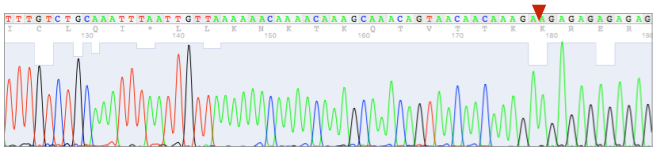
Figure 2. Whole genome sequencing reveals more than 2 alleles at the chr3:88,647,832 – 88,647,838 deletion suggesting a CRISPR/Cas9-induced pattern of mutations. $F_0$3 and $F_0$5 both have more than 2 alleles at the deletion site while FVB has only the wild-type allele at this site. Images created using mouse BAM files and VarSeq (Golden Helix, Inc., Bozeman, Montana).



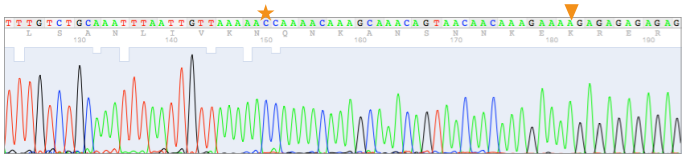
Allele 1: WT



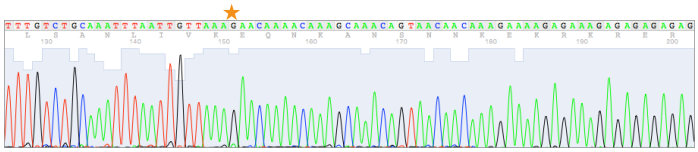
Allele 2: Deletion



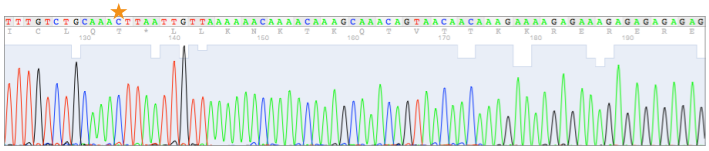
Allele 3: Alternative Deletion



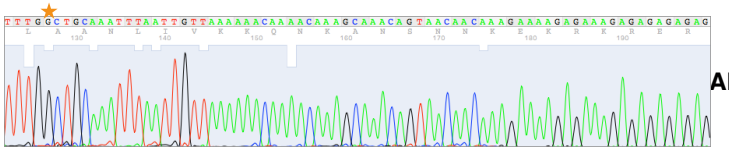
Allele 4: Deletion + SNP A>C



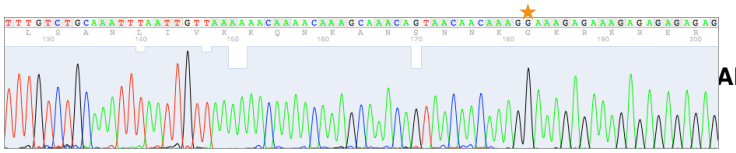
Allele 5: SNP A>G



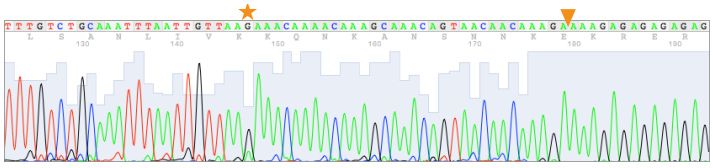
Allele 6: SNP T>C



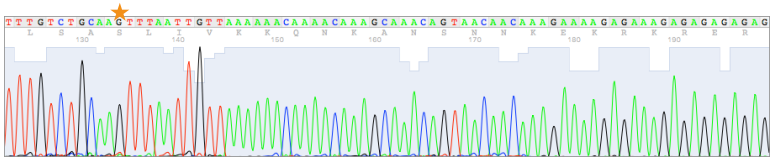
Allele 7: SNP T>G



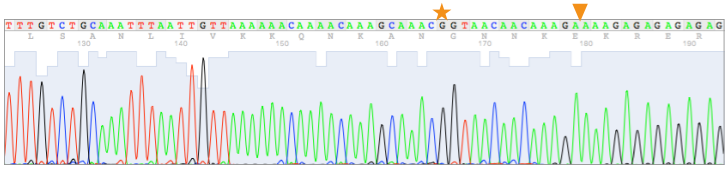
Allele 8: SNP A>G



Allele 9: Deletion + SNP A>G



Allele 10: SNP A>G



Allele 11: Deletion + SNP A>G

Figure 3: Sanger Sequencing reveals more than 2 alleles at the chr3:88,647,832 deletion suggesting a CRISPR/Cas9-induced pattern of mutations adjacent to NGG and NGA which are underlined in black on the wild-type allele. Deletion site is underlined in orange on the wild-type allele and is noted by the orange arrow in alternative alleles. Alternative deletion is noted by the red arrow. Different SNVs are denoted by stars.

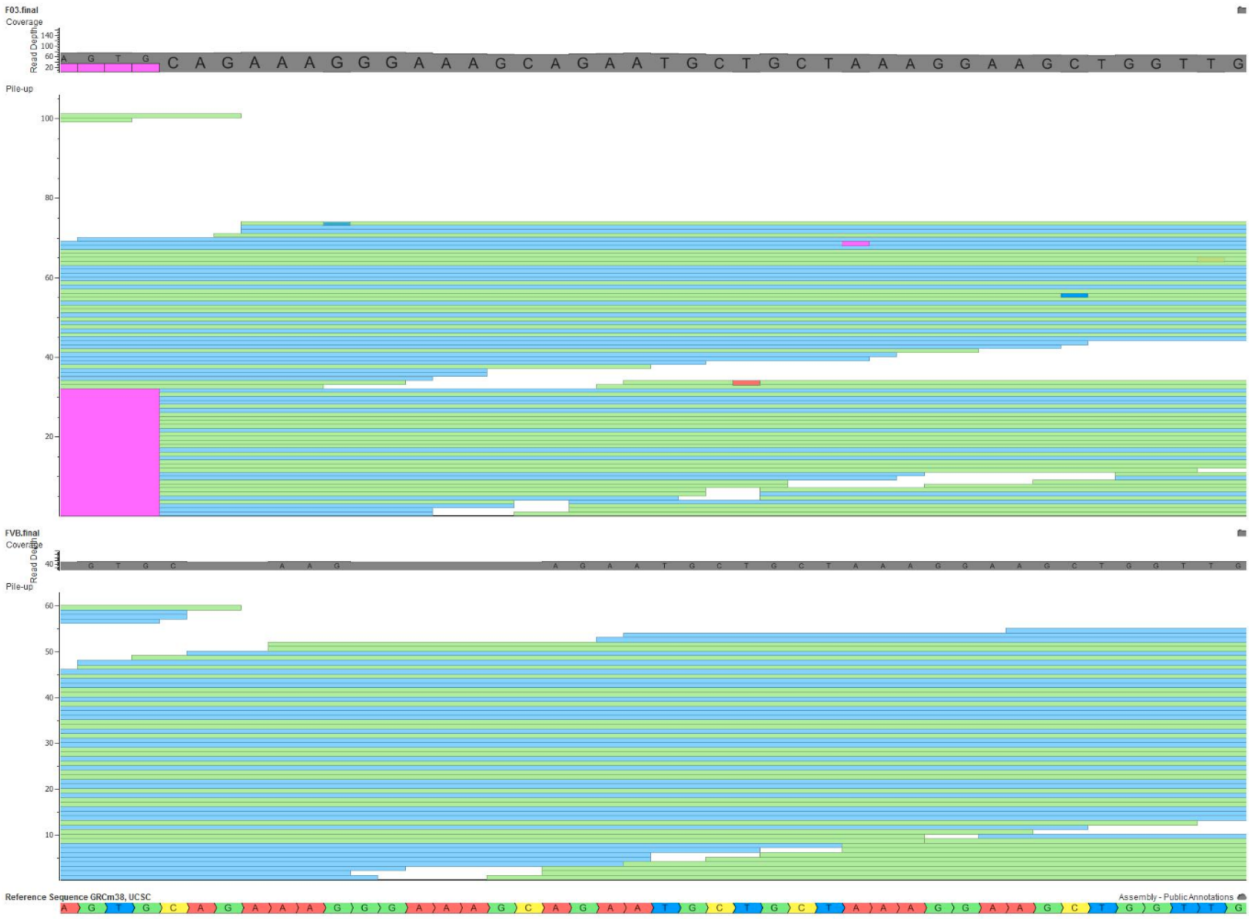


Figure 4. Whole genome sequencing reveals more than 2 alleles at the chr4:66,453,495 – 66,453,539 deletion suggesting a CRISPR/Cas9-induced pattern of mutations. $F_0$3 has more than 2 alleles at the deletion site while FVB only has the wild-type allele at this site. Note that the right-most T>C SNV in $F_0$3 was subsequently validated by Sanger sequencing (see Figure 5). Images created using mouse BAM files and VarSeq (Golden Helix, Inc).

Sanger Sequencing Validation of a CRISPR/Cas9 pattern at mutation sites in the WGS data. Since rare WGS variants might be attributed to WGS artifact, we next employed an alternate method (Sanger sequencing) to explore the possibility that CRISPR/Cas9 induced the pattern of mutations at variation sites we described in our original Correspondence. Primer pairs were designed for each site, and PCR followed

by TOPO cloning and sequencing was performed (see Methods). We found greater than 2 alleles at multiple sites. One site included the chr3:88,647,832 site for which we observed multiple alleles including an alternate indel at the same region as our previously reported indel, a signature of CRISPR (Figure 2). This indel was also detected by Sanger sequencing (Figure 3). Another site included the chr4:66,453,495 site, for which we had observed multiple alleles (Figure 4). Sanger sequencing at chr4:66,453,495 (Figure 5) demonstrated at total of 8 different alleles, including both the anticipated wild-type and deletion alleles as well as 6 other alleles, including novel SNVs. Several of these novel alleles were not visualized in the WGS reads (Figure 4), but of note, allele 8 was present in the WGS read. Similarly, we observed multiple alleles by WGS around the deletion at chrX:123,734,765 (Figure 6). Sanger sequencing at chrX:123,734,765 demonstrated 7 alleles at this site (Figure 7). Alleles 2 and 7 found by Sanger sequencing (Figure 7) can be seen in WGS reads (Figure 6). Another mutation site, chr14:94,681,371, showed a total of 13 different alleles detected by Sanger sequencing, including two additional deletions (Figure 9). At least one of these alleles was present in the WGS read (Figure 8). Thus our combined deep WGS analysis and Sanger sequencing experiments at our originally described CIRSPR/Cas9-induced mutation sites demonstrated multiple alleles that are inconsistent with Mendelian inheritance.

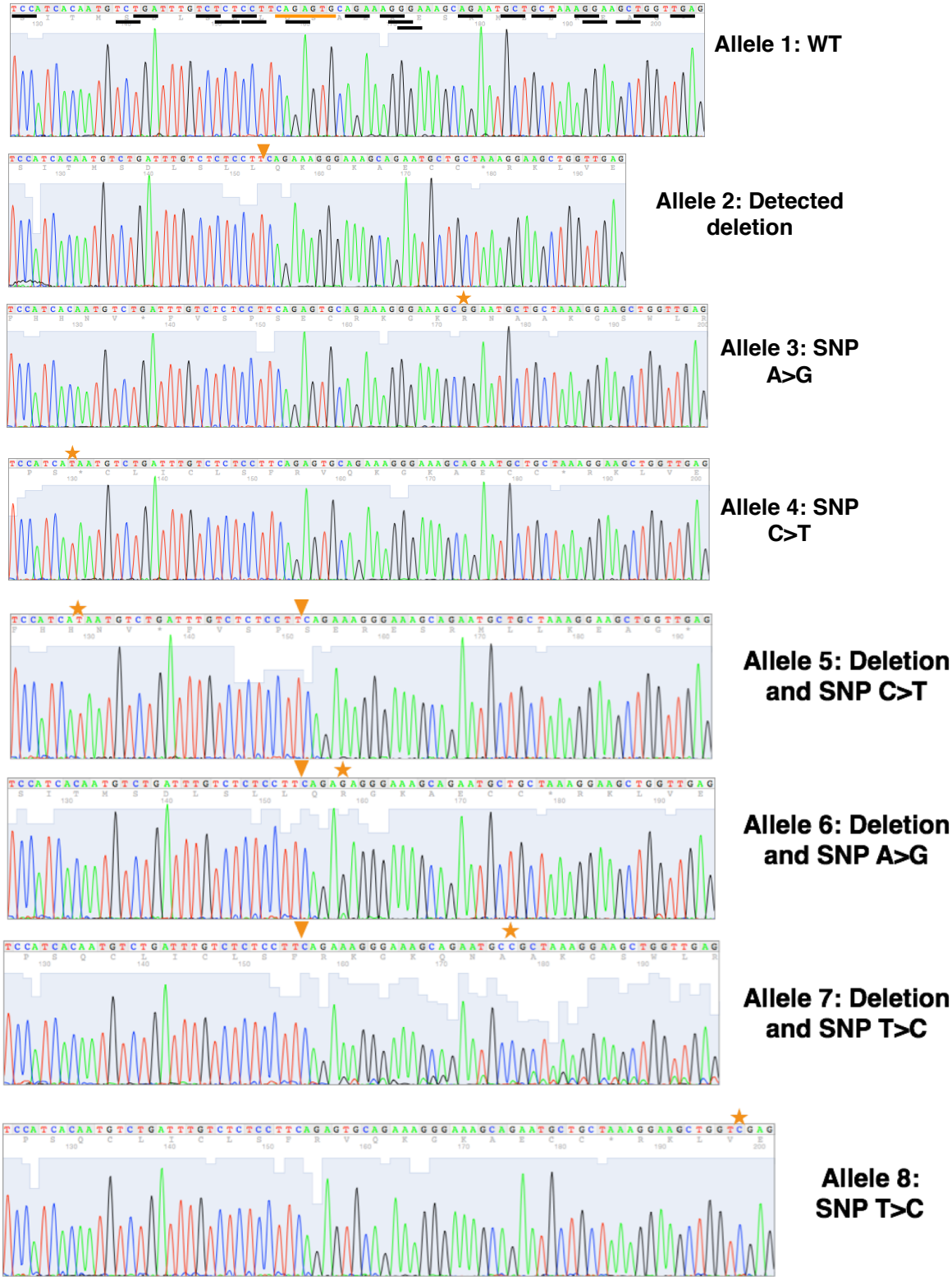


Figure 5. Sanger Sequencing reveals more than 2 alleles at the chr4:66,453,495 deletion suggesting a CRISPR/Cas9-induced pattern of mutations adjacent to NGG and NGA which are underlined in black on the wild-type allele. Deletion site is underlined in orange on the wild-type allele and is noted by the orange arrow in alternative alleles. Different SNVs are denoted by stars. Note that the SNV in allele 8 was also present in the WGS data (see Figure 4).

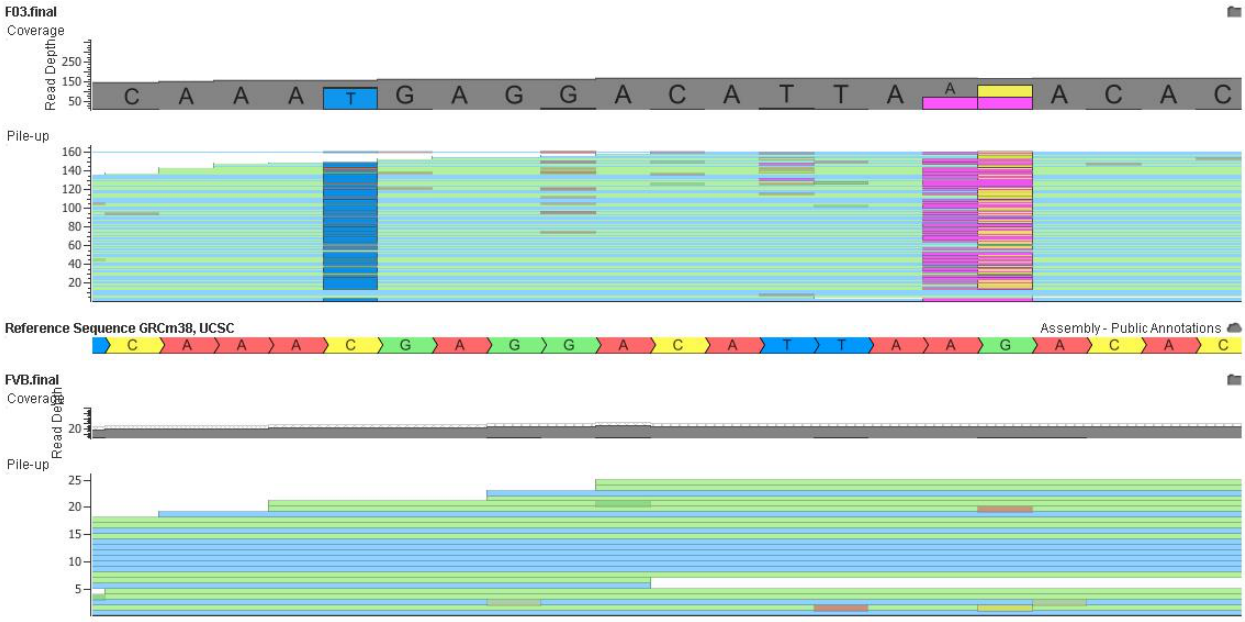


Figure 6. Whole genome sequencing reveals more than 2 alleles at the chrX: 123,734,750- 123,734,770 deletion suggesting a CRISPR/Cas9-induced pattern of mutations. $F_0$3 has more than 2 alleles at the deletion site while FVB only has the wild-type allele at this site. Images created using mouse BAM files and VarSeq (Golden Helix, Inc). Note that one C>T and A>G SNV in $F_0$3 was subsequently validated by Sanger sequencing (see Figure 7).

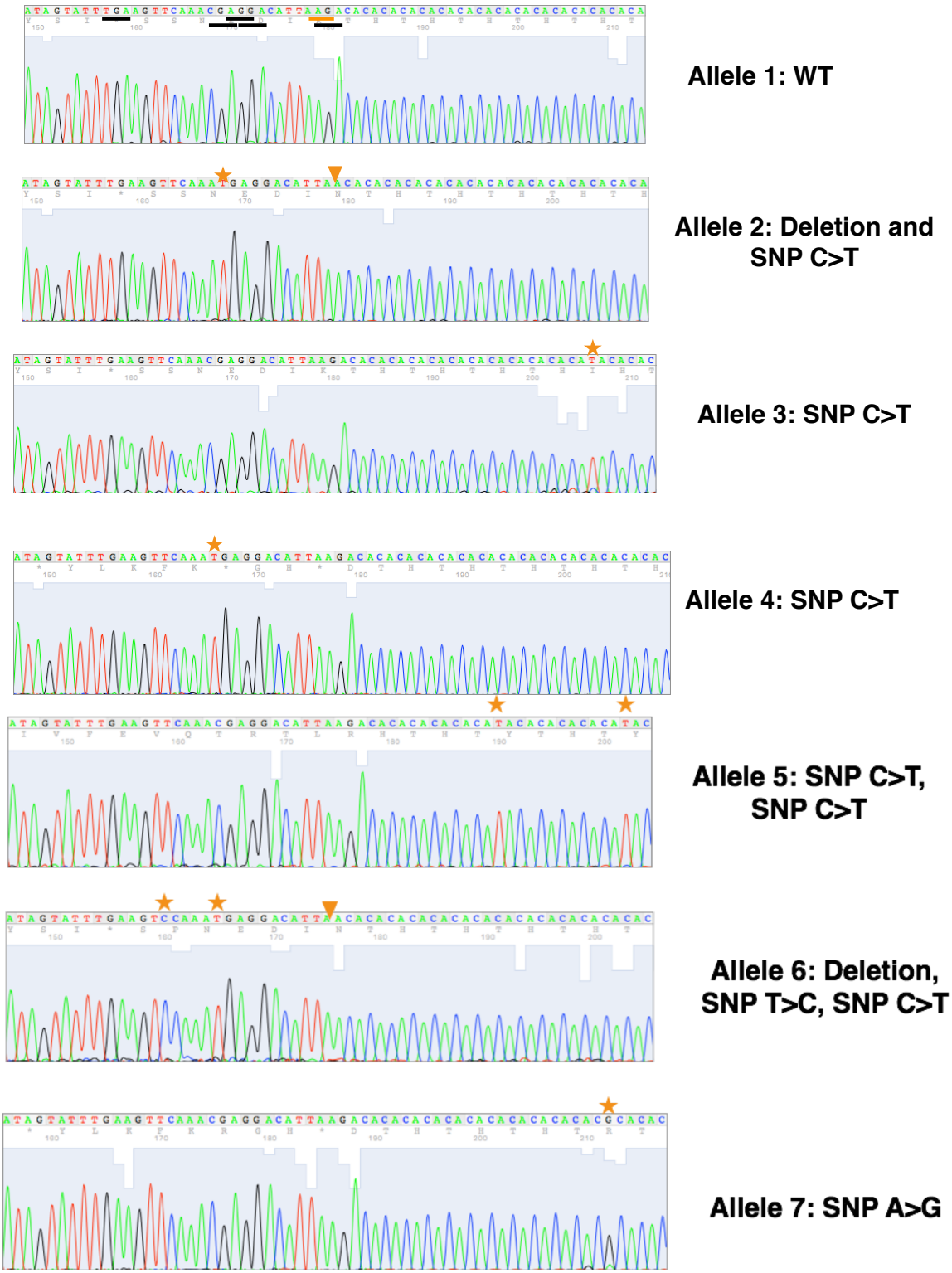


Figure 7. Sanger sequencing reveals more than 2 alleles at the chrX:123,734,765 deletion suggesting a CRISPR/Cas9-induced pattern of mutations adjacent to NGG and NGA which are underlined in black on the wild-type allele. Deletion site is underlined in orange on the wild-type allele and is noted by the orange arrow in alternative alleles. Different SNVs are denoted by stars. Note that the SNVs in allele 2 and 7 were also present in the WGS data (see Figure 6).

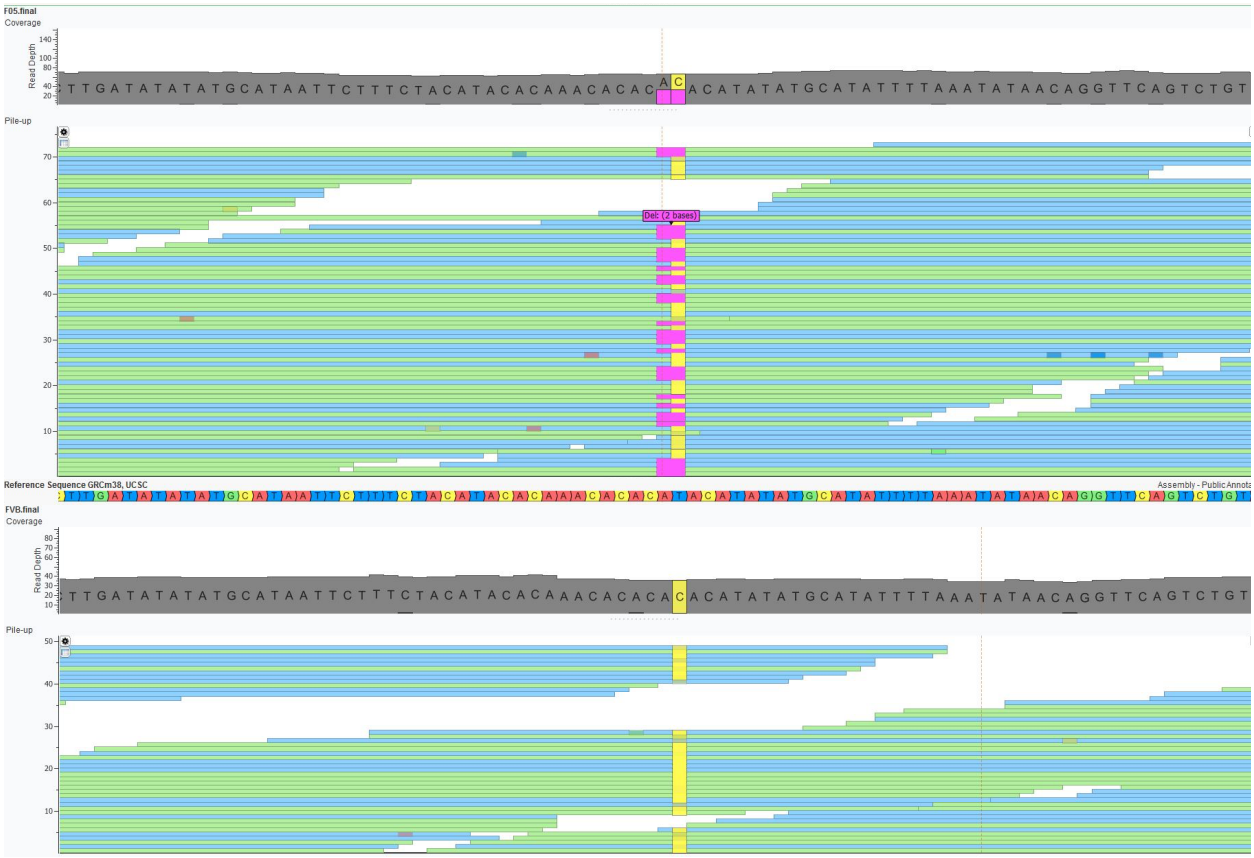
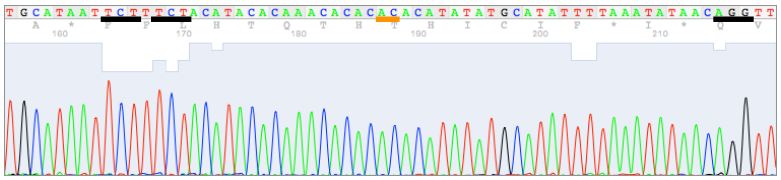
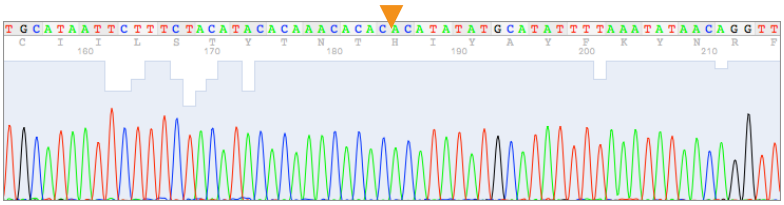


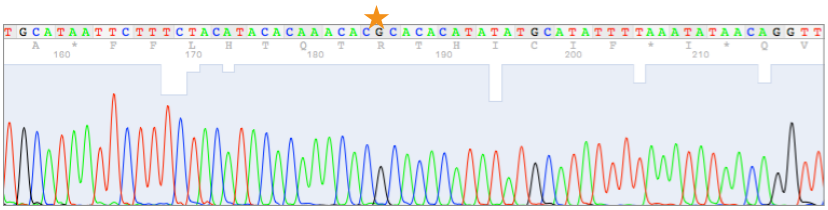
Figure 8: Whole genome sequencing reveals more than 2 alleles at the chr14:94,681,330-94,681,411 deletion suggesting a CRISPR/Cas9-induced pattern of mutations. $F_0$5 has more than 2 alleles at the deletion site while FVB only has the wild-type allele at this site. Images created using mouse BAM files and VarSeq (Golden Helix, Inc). Note that a C>T SNV in $F_0$5 was subsequently validated by Sanger sequencing (see Figure 9).



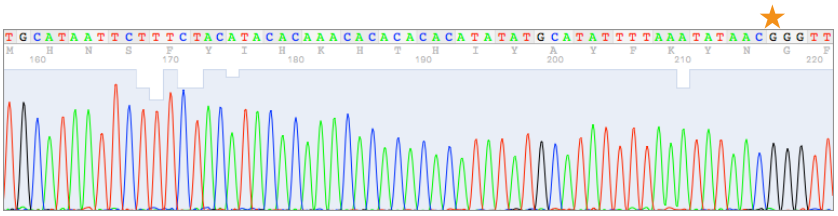
Allele 1: WT



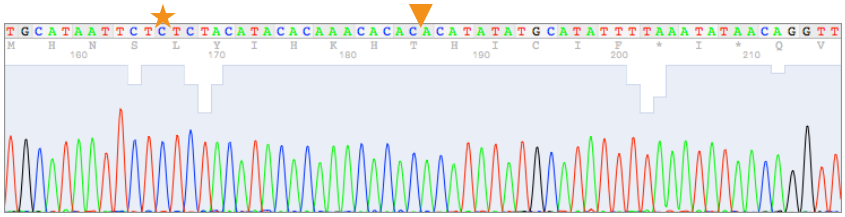
Allele 2: Expected Deletion



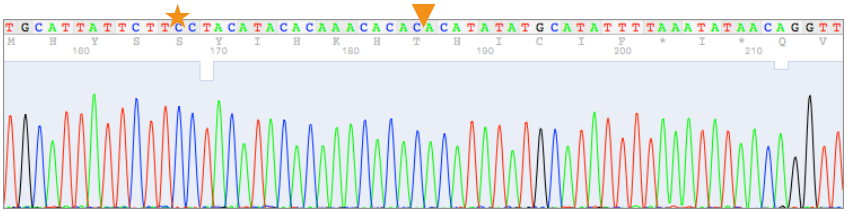
Allele 3: SNP A>G



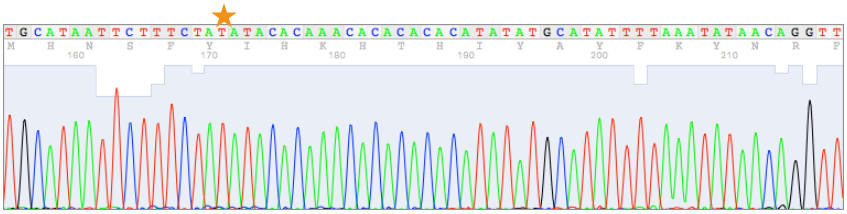
Allele 4: SNP A>G



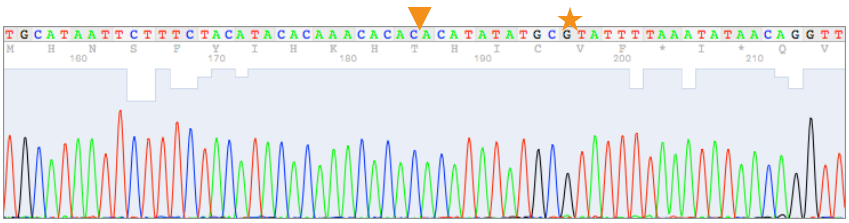
Allele 5: Expected Deletion + SNP T>C



Allele 6: Expected Deletion + SNP T>C



Allele 7: SNP C>T



Allele 8: Expected Deletion + SNP A>G

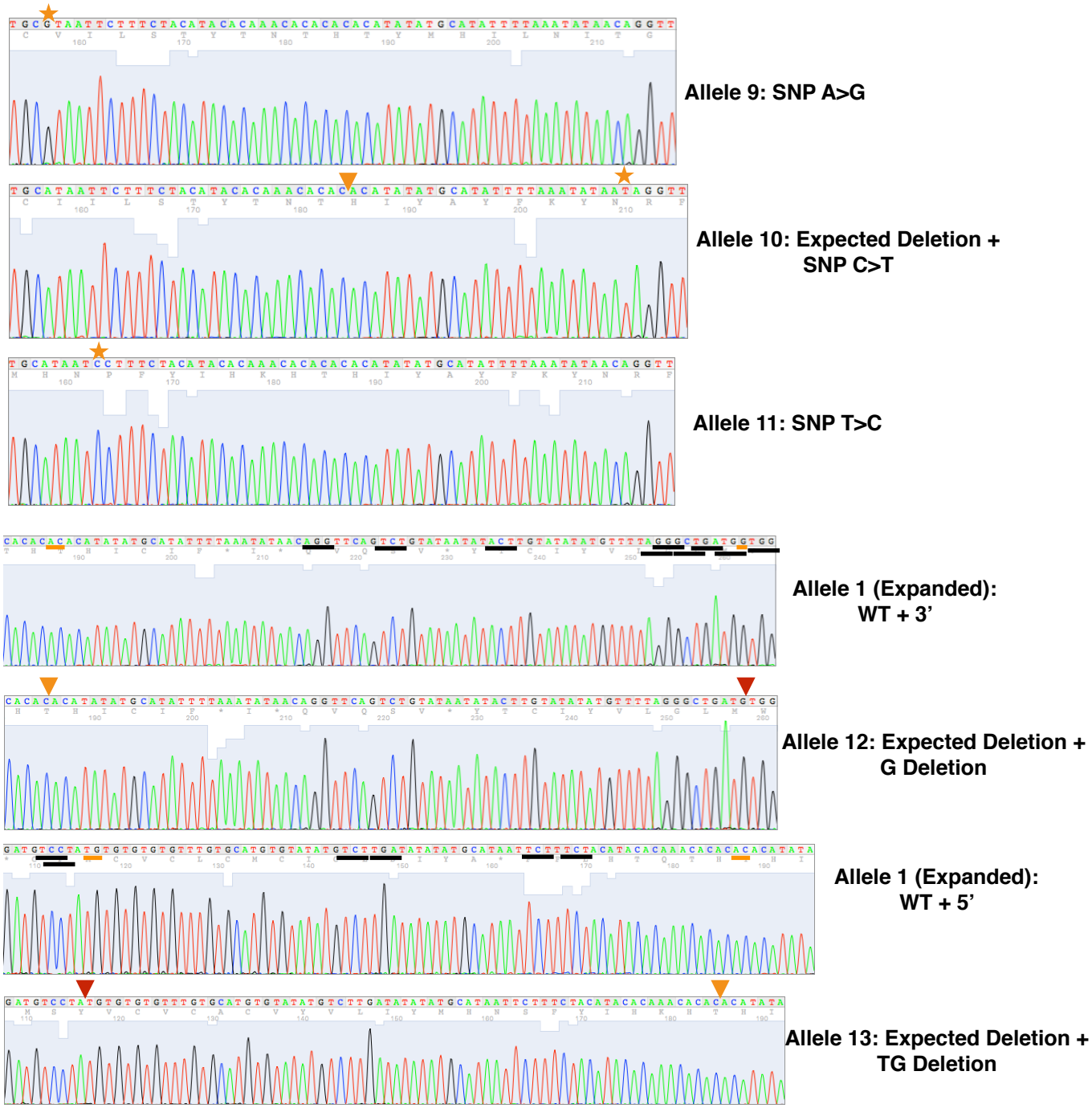


Figure 9. Sanger sequencing reveals more than 2 alleles, including SNVs and deletions, at the chr14:94681371 deletion suggesting a CRISPR/Cas9-induced pattern of mutations adjacent to NGG and NGA which are underlined in black on the wild-type allele. Deletion site is underlined in orange on the wild-type allele and is noted by the orange arrow in alternative alleles. Additional deletions are noted by red arrows. Different SNVs are denoted by stars. Note that the C>T SNVs in allele 10 was also present in the WGS data (see Figure 8).

Possible interpretations and mechanisms of the combined WGS and Sanger sequencing data. One possible alternative interpretation of these findings is that both WGS and Sanger findings are due to artifact. However, given that we have found these multiple alleles using more than one method, here we posit the interpretation that the finding of multiple, novel alleles at our originally described CRISPR/Cas9 mutation sites by two different methods suggests that these are in fact CRISPR/Cas9-induced variants. One possible mechanism underlying our findings is that the displaced non-target strand is prone to hydrolysis at the 5' end of the PAM sequence.[2] Insufficient repair of this hydrolysis could explain why SNVs were detected. Such mechanisms would be much more difficult to detect, if at all, *in vivo* and without WGS. An alternate but related mechanism would be that the observed CRISPR-induced SNVs might be due to translesion synthesis after double strand breaks, which has been described in E. Coli.[3]

A third explanation for the number of mutations detected, specifically SNVs, could be due to break-induced repair (BIR); one way the cell can repair breaks in DNA. Cas9 D10A-nickase can lead to BIR, which is prone to high levels of genetic instability including 1,000-fold increase in mutations when compared to normal DNA replication.[4] Based on the mechanism of BIR, any mutations, specifically SNVs, made in the leading-strand can become fixed during synthesis of the lagging strand as there is no opportunity for mismatch correction.[5] Furthermore, one study analyzing CNVs discovered an unexpectedly high mutations spectrum represented by both SNVs and template switches that could be detected in up to 52% of the CNVs at the locus studied.[6] This included long-distance template-switching events that resulted in mutations far away from the breakpoint junctions.[6]

Most of the mammalian genome (some approximations nearing 75% of bases) is transcribed, even though only 1-2% encodes for proteins[7], whereas 80-95% of prokaryote genomes encode proteins.[8] The rest of the transcriptome is comprised of tens of thousands of non-coding RNAs (ncRNAs). ncRNAs play many regulatory roles in eukaryotic cells including controlling multiple levels of gene expression through transcription and translation, RNA splicing, chromatin architecture and epigenetics.[8] Although ncRNAs are also present in prokaryotes and simple eukaryotes, the role of ncRNAs has increased dramatically in the genomes of more complex organisms, including small ncRNAs, which may influence Cas9 function as described below.

Perhaps most relevant is the potential role of miRNAs which are approximately 22 nucleotides long, the same size as the seed sequence of a crRNA. Since 5'-NGG-'3 for SpCas9 can on average be found every 8-12bp in the genome, both lncRNA and miRNA could act as crRNA. Additionally, there have been multiple reports of miRNAs in the nucleus, making it even more likely that Cas9 may be associating with crRNA-like endogenous sequences within the same subcellular compartment as DNA.[7] Furthermore, precursors of siRNAs, another type of small ncRNA, are transcripts with long stem-loop structures, a vital structure of the crRNA. All these ncRNA species are seen in substantially high numbers of embryogenesis.[9] This could implicate them as having a role in off-targeting of Cas9 injected embryos.

Proof of any of these particular mechanisms awaits further studies, but there is no obvious mechanism by which simple Mendelian inheritance can explain variants observed in the CRISPR-treated mice.

Furthermore, Intellia, ToolGen, Editas and Beacon all questioned the degree of relatedness of mice in our study. The FVB colony control and the parents (of founders) of the CRISPR-treated mice all belonged to generation F95pF100-F95pF101 (confirmed by JAX) obtained under the JAX genetic stability program. The FVB colony control and the FVB parents were never bred in-house, so genetic differences cannot be due to genetic drift.

Comparison to previous WGS CRISPR/Cas9 *in vivo* papers. There are surprisingly few studies using WGS after CRISPR, and casual comparisons miss important differences. The Iyer et al. paper surveyed CRISPR-treated F1 hybrid mice for off-target mutations and focused on indels.[10] In contrast, we surveyed CRISPR-treated F0 inbred mice 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.[11] The difference in results may reflect differences in the strains, filtering methods needed for hybrid mice with high levels of heterozygosity, gRNAs, technique (as we addressed) etc., but cannot be attributed solely to the use of colony controls. The two studies are similar to our study in terms of sample size and lack of parental controls.

Novel filtering methods for WGS explain differences in variant call rates. An advantage of the original study was the use of dedicated somatic variant callers in the WGS analysis pipeline. By design, any CRISPR induced mutations in the $F_0$3 and $F_0$5 mice would be expected to be mosaic in nature. This more closely resembles the mixture of alleles seen in tumor samples for which these somatic variant callers were originally developed. The use of such tools not only increases the sensitivity for detecting low allelic fraction variants but also allows for the accurate detection of multi-allelic variant sites that might otherwise be missed by conventional germline variant callers that assume a diploid state. The original analysis also enhanced specificity by using a consensus calling approach. Only variants that were called in all three of the somatic variant callers (muTect, LoFreq, and Strelka for SNVs and Strelka, Scalpel and Pindel for indels) were retained and reported. These combined methods differ from the more traditional WGS methods used by Editas and Intellia. The use of dedicated somatic variant callers combined with a consensus calling approach greatly increased the probability of detecting true positive off-target CRISPR mediated mutations.

Claims regarding widespread heterozygosity in the inbred FVB line. There is significant heterozygosity observed in $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 to 8-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 6 months of age. Based on the JAX order, the parents that generated both the stud and oocyte donor were likely to be siblings of the stud, 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 be homozygous. Instead, we observed extensive heterozygosity, which was validated by Sanger sequencing (Figures 3, 5, 7, 9 and 10, Table). The heterozygosity in $F_0$3 and $F_0$5 is unlikely to be parentally inherited. The colony control FVB/N was also purchased from JAX and not bred in-house.

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%).

The Editas and Intellia re-analysis of our sequencing data has limited relevance to our Correspondence. The chief issue is that Editas notes the many variances between cases or controls and the mm10 reference sequence from a C57BL/6 strain. The appropriate reference sequence is the FVB/N strain sequence. Nonetheless, because we originally wanted to rule out any common germline mutations, we specifically *excluded* all variants present in the mm10 sequences (C57BL/6) or any of the 35 other reference strains present in dbSNP.

Heterozygosity in mice varies by strain and breeding, and use of the highly inbred FVB/N mice from JAX without in-house breeding is an experimental advantage.[12-14] Inbreeding leads to a reduction in heterozygosity within the population.[15] 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 not reported.[16] Moreover, the A^{vy} line is known to have a poor DNA-repair mechanism, and a high spontaneous cancer rate.[17,18] 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 $F_0$3, 299 unique to $F_0$5, and 1,397 shared between the two) and over 30% of the over 160 total indels (47 unique to $F_0$3, 11 unique to $F_0$5 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 10 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.

Intellia and ToolGen claim that their re-analysis of the WGS data from our correspondence shows many heterozygous sites in each of the three FVB mice, with Intellia's analysis claiming hundred of thousands of sites. This number was not validated by any other method (e.g., Sanger sequencing of a commercially available FVB mouse at suspected heterozygous sites), and may represent a combination of insufficient filtering and false positive WGS reads. Evidence that Intellia and ToolGen's heterozygous sites are likely false positives comes from several sources. First, Wong et al, in the only paper to report WGS sequencing in the FVB mouse reported no heterozygous variants, and no other published report describes any widespread heterozygosity in FVB mice.[12] Second, Wong et al report only 115,228 total private SNPs in the FVB/NJ strain, so the total number of heterozygous sites reported by Intellia and ToolGen in each mouse is higher than the total number of private FVB SNPs previously described. Third, via Sanger sequencing confirmation, Wong et al describe that every 2/127 of their SNPs are false positives, but Intellia and ToolGen do not confirm any of their proposed variants by a secondary method. Fourth, to avoid false positives, our original Correspondence insisted on variants passing 3 different filtering methods, and then also not being known common germline variants, and not ever having been seen in dbSNP. Fifth, we Sanger sequence confirmed many of our variants (original Correspondence Supplementary Figure and Figures 3, 5, 7, 9 and 10 in the present Correspondence). In summary, the amount of heterozygosity in the inbred FVB line (the standard line for murine transgenesis) claimed by Intellia and ToolGen is not reflected in the literature and has not been validated by independent approaches.

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 simply determine all of the sequencing differences between the cases and controls. 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, Figures 3, 5, 7, 9 and 10, 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 at least consider the source as CRISPR therapy intervention.

Sequence read depth differences between cases and 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.[12] 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 (Figures 3, 5, 7, 9 and 10).

Identity of some mutations in cases. 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 very 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).

The finding of identical variants between the two CRISPR treated mice may also be explained by the filtering and/or the upper limits of the sensitivity of our study design. It is likely that CRISPR-Cas9 caused mutations at a particular off-target site at a high rate, and that many different alleles were created in this mutagenesis. However, since we sequenced at 50x and only accepted calls greater than 23x, we would not call many mutations that were lower frequency. The mutations called and reported in our original Correspondence may thus simply reflect the high frequency mutations and calls, but

there may be multiple other mutations at the same genomic loci. In the present Correspondence, we validate multiple alleles at several loci (Figures 1-9), confirming the utility of deep sequencing by multiple methods. These alleles contain additional indels, addressing Beacon's criticism of only identical indels. Future studies with alternate off-target calling methods (e.g., CIRCLE-Seq) or higher depth sequencing and different filtering protocols will directly answer this question.

Sample size (power) question. 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.[19] Off-target analysis was a secondary outcome reported in our Correspondence.[1]

In our study, only two of the eleven founders showed successful HDR.[19] 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,[10] 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 *in vivo* off-targeting is ascertained by WGS should be designed with additional controls, including parental controls.

Issues regarding gRNA guide design. 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 guide *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. This raises important questions. Are guide optimization studies performed by algorithms *in silico* or performed in immortalized cell lines predictive of guide function *in vivo*? Collection of more *in vivo* data using WGS will help address this question.

Consideration of other NGS studies. Our original correspondence was limited to five references; so two additional references were not included because of methodological differences.[20,21] For example, the Nakajima et al study used exome sequencing,

whereas our Correspondence used whole genome sequencing.[20] Mianne et al used 9x read depth and 1.5% assembly gaps for their WGS.[21] It is uncommon today for WGS to be performed at such low coverage, since filtering is likely to exclude so many regions due to poor quality. This may result in many false negative calls versus our 50x coverage. As in our study, in Mianne et al the sequenced control was a wild type mouse, not a parent. Mianne et al used a "standard mutation detection tool to search for potential sequence variations." The identity of these tools is not known to us. Mianne et al chiefly examined predicted off-target sites and sites surrounding the on-target site. Mianne et al then used another unidentified SNV detection tool to look only at coding sequences and found 42 SNVs. They went on to eliminate most of these because they had "low allele frequencies"- but again, this is with an average of 9x coverage and none were evaluated by Sanger sequencing before excluding them. Mianne et al Sanger sequenced 7 coding region SNVs- 6 of which were predicted to be false positives and verified 1 real SNV.

Minor labeling discrepancies in the original Correspondence. There are labeling discrepancies in supplemental Figure 3 of the correspondence. In panel 3a, the top-10 off-target sites predicted by the Benchling software was originally performed when the guide was designed and used the mm9 build of the mouse genome. When the WGS analysis was performed later, the mm10 build was used, so Figure 3b-d used labels from the mm10 build. Regardless of the build, the sequences in panel 3a are the same, but for consistency, the chromosomal locations and gene names were relabeled using the mm10 build. For clarity, descriptive column titles were added, the first (*Pde6b*) sequence was removed in 3a, and the last five nucleotides of *Pcnt* in panel 3b were corrected (Revised Supplementary Figure 3).

a **Top 10-Predicted Off-Target Sites:**

Predicted Cut Site	Best homology within 100 nucleotides of SNV	Gene
gRNA	CCAACCTAAGTAGCAGAAAG	
chr18:86764777	CACCCCTAATTTAGCAGAAAG	No Associated Gene
chr13:40633714	CCTACCTCAGTACCAGAAAG	No Associated Gene
chr12:36835025	CAAACCTAAGGAGCAGAAAT	No Associated Gene
chr11:111097185	ACAAGCAGAGTAGCAGAAAG	No Associated Gene
chr10:125270806	GCAAGCTAAGTAGCAGAAAA	Slc16a7
chr10:112803271	CTTACCTAAGAAGCAGAAAG	No Associated Gene
chr8:75958718	AGACACTAAGTAGCAGAAAG	1700007B14Rik
chr5:49056372	CAAAGCCCAGTAGCAGAAAG	Kcnip4
chr5:148619553	CCTCCCTGATTAGCAGAAAG	No Associated Gene
chr2:145436435	ACAAACACAGTAGCAGAAAG	Slc24a3

b **10 True off-target site SNVs in Coding Regions**

SNV site	Best homology within 100 nucleotides of SNV	Gene
gRNA	CCAACCTAAGTAGCAGAAAG	
chr18:67834688	CCAAACATCTCTCATGTCTT	Cep192
chr9:66434164	AGTTATGTGGAACAGACACC	Herc1
chr7:63938170	CTCGCTCCACGCGGGGCTGG	Klf13
chr15:8324487	CCACACCTTCTCAGTTCAGC	Nipbl
chr19:12283753	GTAGGTCCAAGGCCTTCAAC	Olfr1433
chr10:76356251	CCACTCGAACGGCGGTGCGG	Pcnt
chr15:76176864	CCCCACTGTCCAGCAGCTCT	Plec
chr2:129609045	CCATTAGGTGGTACAGAGGA	Sirpa
chr5:124885592	CCATACAGCAGAGAAACCCC	Zfp664
chr2:167056200	CAACCCCAGAGGCTCTCAAA	Znfx1

c **10 True off-target site SNVs in Non-coding Regions**

SNV site	Best homology within 100 nucleotides of SNV	Gene
gRNA	CCAACCTAAGTAGCAGAAAG	
chr1:190479227	CTTTCAGCCTTGGTCAGAGA	Gm23153 (miscellaneous RNA)
chr2:92591648	GTTTCATGGTCCGTGGCAGG	Mir7221 (miRNA)
chr2:140439882	CTTTCTGCTAGTGTTGGTCC	Gm23846 (snRNA)
chr5:62132979	GAAAACTAAGATACAGAAAT	Gm22273 (snoRNA)
chr6:31721846	TTTTCTTTTATGGTGGCTGG	Gm13847 (lincRNA)
chr6:133758061	CCAATTCTAAGAGGGGCAC	Gm23375 (miRNA)
chr11:54214410	ATGTTTGTACCCTTGTTGCC	4933405E24Rik (lincRNA)
chr11:83274656	CCATAGCTACGAACCACAGA	Gm23444 (snRNA)
chr11:112711337	CTCCAGACATTGAATGAGCC	BC006965 (processed transcript)
chr18:54449411	CACACACAGAGAGAGAGAGA	Redrum (lincRNA)

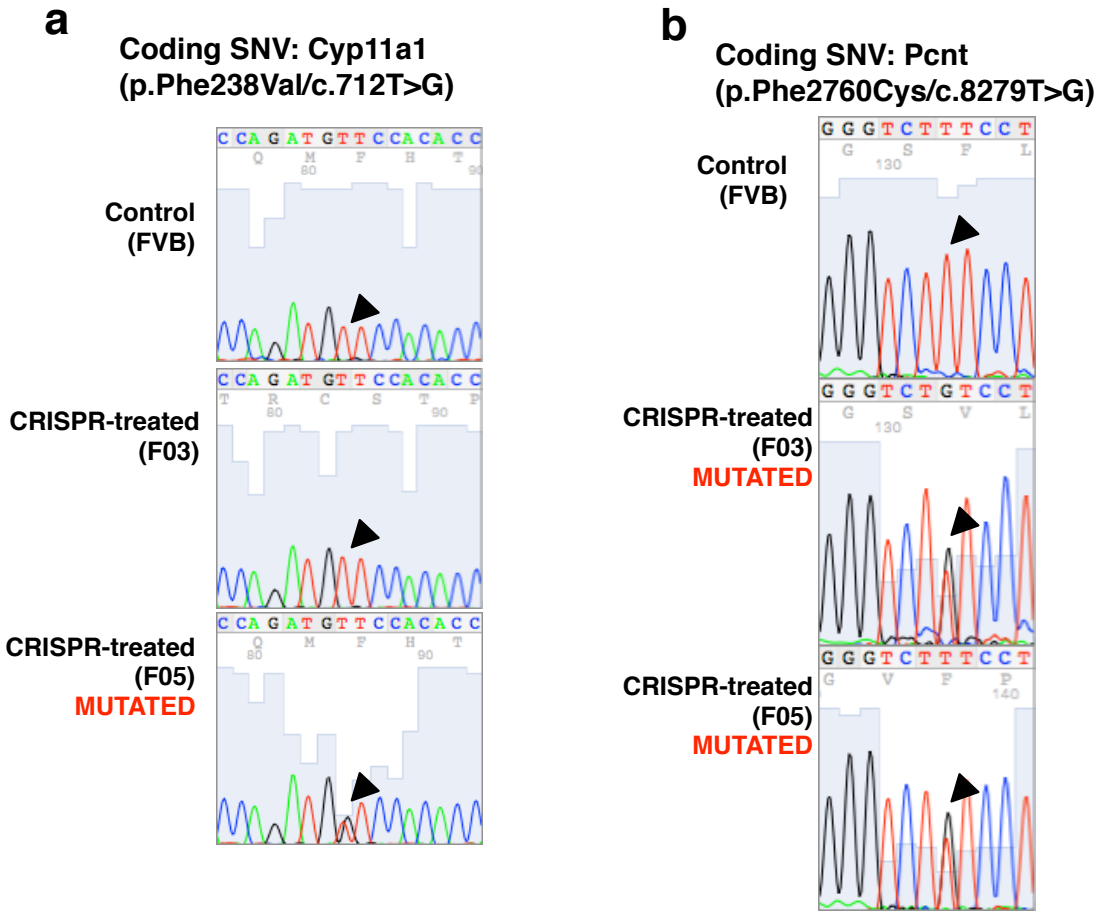
d **10 True off-target site Indels**

SNV site	Best homology within 100 nucleotides of SNV	Gene
gRNA	CCAACCTAAGTAGCAGAAAG	
chr4:106562094	CCCTCCTTTATATTCCTCTCT	2700068H02Rik
chr1:188016221	CCCAGCTAACTGTTTGCCCA	Esrrg
chr18:31470317	TCAACTAAGGTAGCAAAGAT	Gm25396
chr4:134618126	CAAAGGAGCTCAGGTTCAGA	Man1c1
chr17:31443854	TCAACGAGCTGAAAGCCGAA	Pde9a
chr8:107584736	CCAAAGCCTTTGGCAGTAAT	Psmd7
chr18:31910970	GCGGCCAGCAGCAGACACAG	Sft2d3
chr19:50287677	AAAACATTCAAATGCATAAA	Sorcs1
chr2:30020366	GCAACCTGAGCACAGAGTCC	Sptan1
chr4:132857590	CCAGCACCTACTCACCAGCT	Stx12

Revised Supplemental Figure 3: Sequence alignment of guide RNA to actual off-target regions does not show significant homology. a. The top 10 predicted off-target regions predicted *in silico,* by Benchling, aligned to the gRNA. Sequences are 80-95% homologous to the gRNA. b. Regions surrounding 10 randomly selected experimentally-observed SNVs in coding regions aligned to the gRNA. Sequences are 15-45% homologous to the gRNA. c. Regions surrounding 10 randomly selected experimentally observed SNVs in non-coding region aligned to the gRNA. Sequences are 5-65% homologous to the gRNA. d. Regions surrounding 10 randomly selected experimentally observed indels in both coding and non-coding region aligned to the gRNA. Sequences are 25-65% homologous to the gRNA.

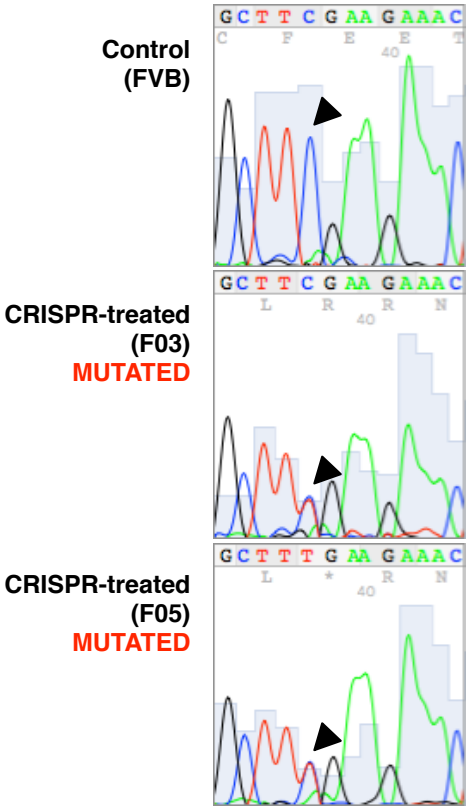
Questions regarding mutation rate. The total number of mutations detected in our Correspondence specifically excludes common germline variants, many of which were described by references such as the Uchimura et al manuscript,[22] in which C57BL/6J mice (JAX mice from Charles River) were used as wild-type control mice. Uchimura et al estimated 101.5 heterozygous SNVs per generation in one control mouse and 92.7 in the other (Uchimura et al Table 1), which is an order of magnitude less than the number of *heterozygous* mutations we found between F03 and F05 (Table), and this excludes known common germline mutations. Also, Uchimura et al performed WGS that included libraries amplified by PCR. Our correspondence included no PCR amplification in our WGS. It is unclear to what extent PCR amplification, which is itself known to introduce errors, could account for some of the mutation frequency observed in the Uchimura study.

Conclusion. 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, of course, that a range of WGS controls are needed that include parents, different gRNAs, different versions of Cas9, and comparisons of 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 *in vivo*. Overall, we are optimistic that some form of CRISPR therapy will be successfully engineered to treat blindness.



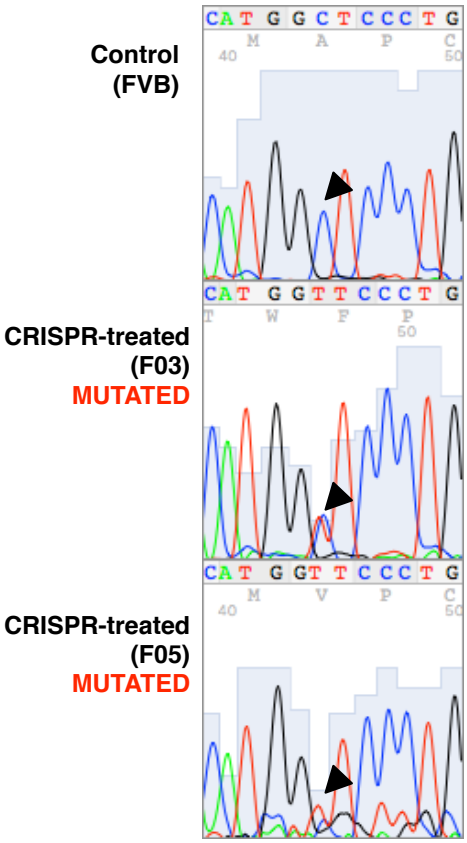
c

Coding SNV: Nipbl
(p.Arg1722Gln/c.5165G>A)



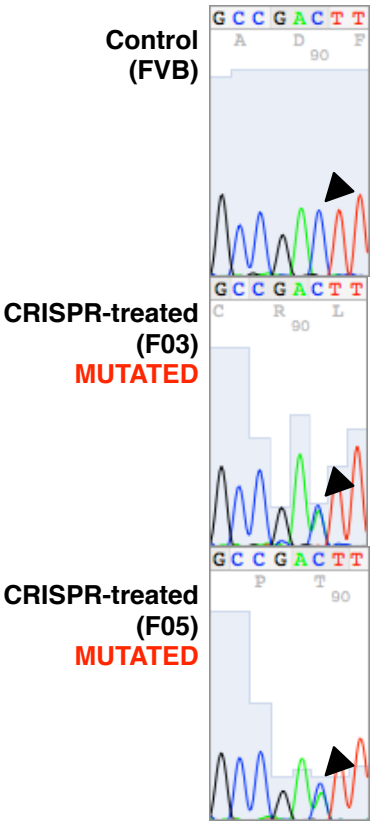
d

Coding SNV: Herc1
(p.Ala3845Val/c.11534C>T)



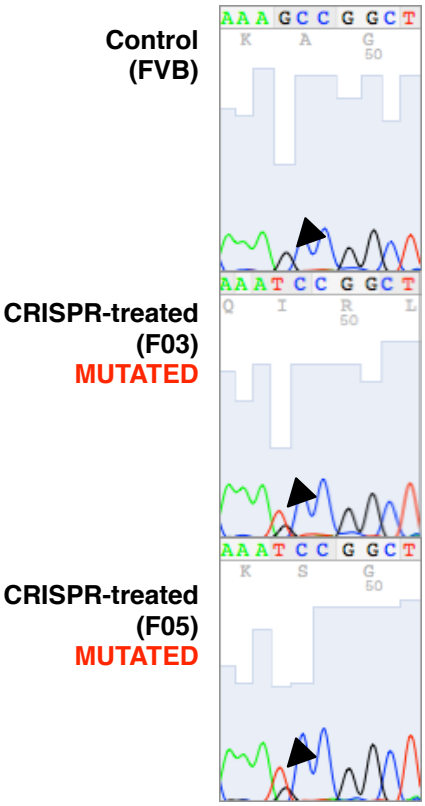
e

Coding SNV: Znfx1
(p.Thr268Asn/c.803C>A)

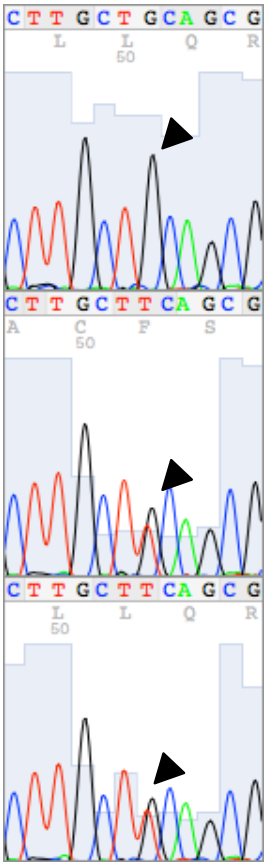


f

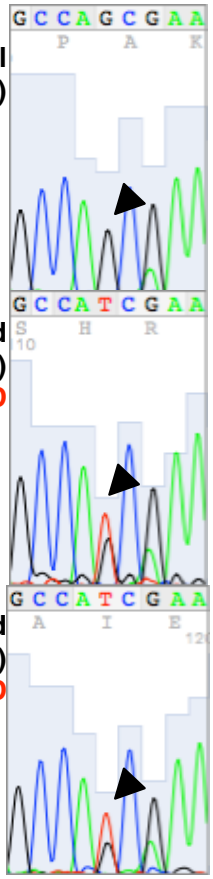
Coding SNV: Sirpa
(p.Ser75Ile/c.224G>T)



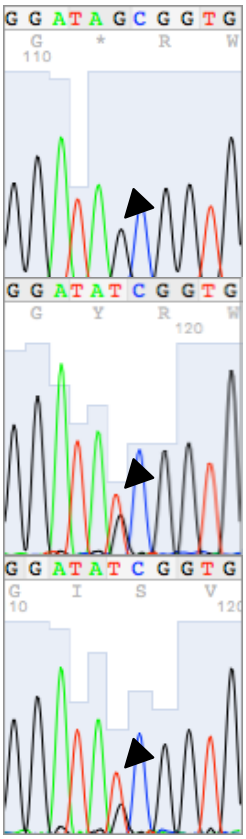
g Coding SNV: Pabpc4 (p.Ala577Ser/c.1729G>T)



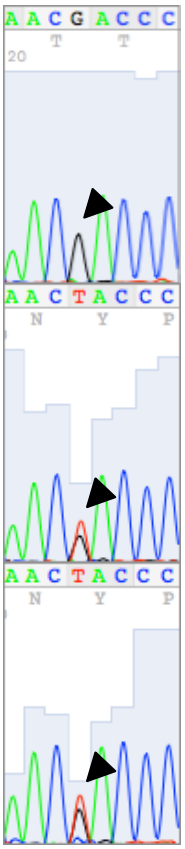
h Intronic SNV: Alg9 (chr9:50820025, G>T)



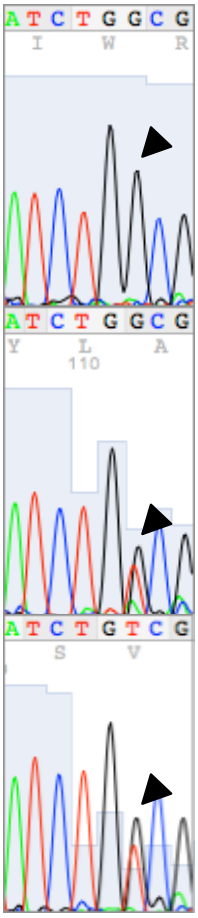
i Non-coding SNV: Ankrd10 (chr8:11640093, G>T)



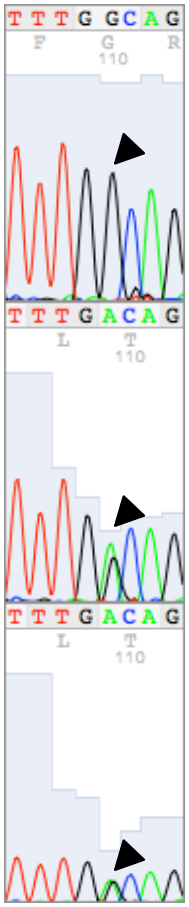
j Intronic SNV: Arhgap42 (chr9:9042738, G>T)



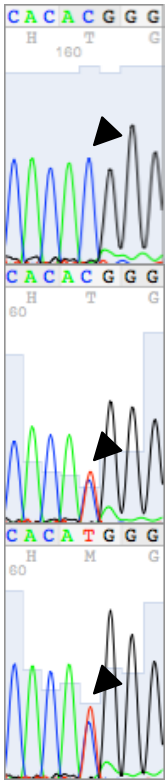
k Non-coding SNV: Fgfr2 (chr7:131642208, G>T)



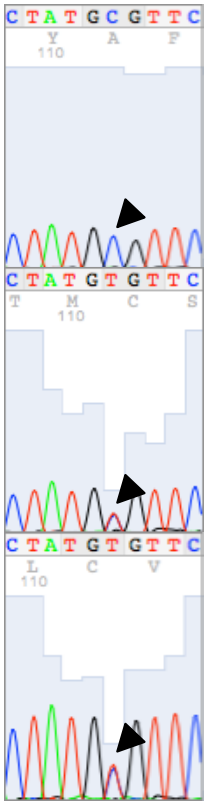
l Non-coding SNV: BEGAIN (chr12:109072362, G>A)



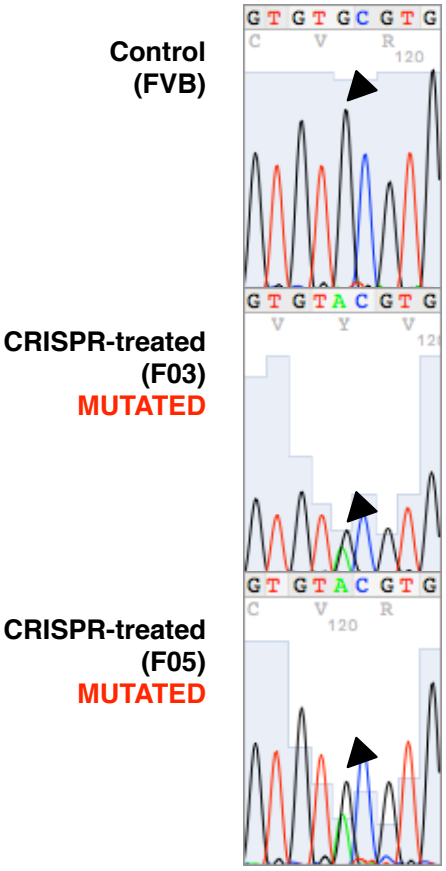
m Intronic SNV: BC034090 (chr1:155223484, C>T)



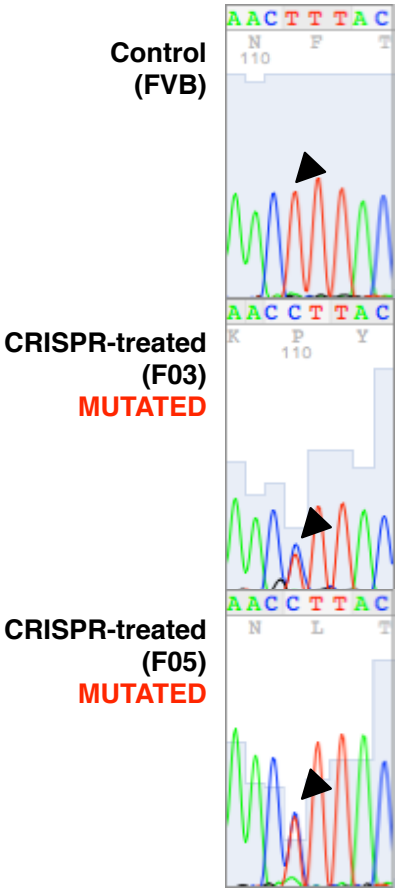
n Intronic SNV: Gria1 (chr11:57215794, C>T)



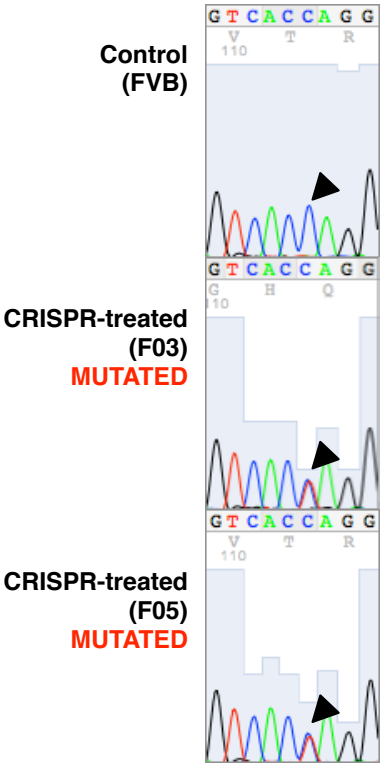
o Non-coding SNV: Foxred2 (chr15:77961298, G>A)



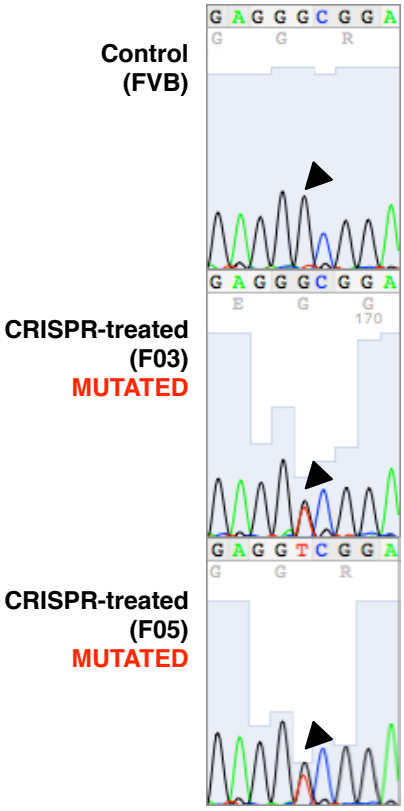
p Intronic SNV: Inpp5a (chr7:139416700, T>C)



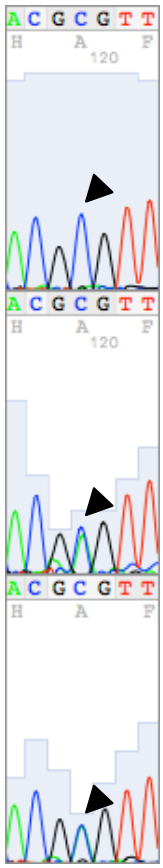
q Intronic SNV: Palmd (chr3:116949130, C>T)



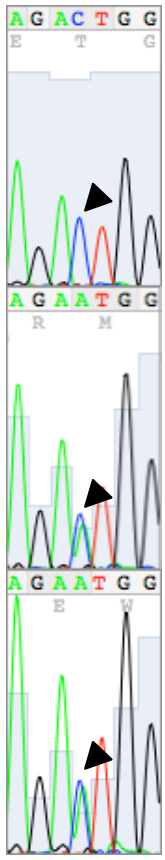
r Non-coding SNV: S1pr2 (chr9:20979401, G>T)



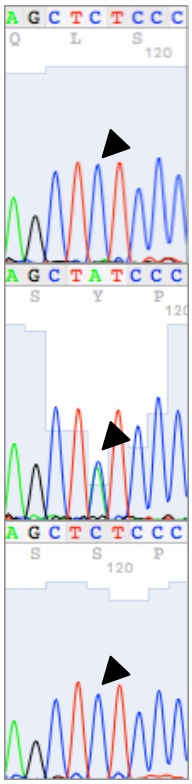
s Intronic SNV: Ppap2b (chr4:105194244, C>A)



t Intronic SNV: Rbms (chr2:60807016, C>A)



u Non-coding SNV: Plekha7 (chr7:116122976, C>A)



v Intronic SNV: Dzip1 (chr14:118925099, G>A)

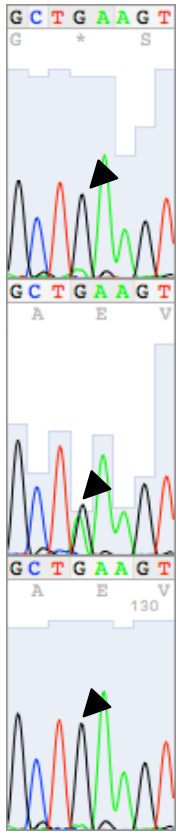


Figure 10. Sanger sequencing confirms heterozygous mutants detected by WGS in CRISPR-treated mice.

Methods

Topo Cloning and Sanger Sequencing

Mutated regions were amplified using primers (Integrated DNA Technologies), Biolase DNA polymerase (Bioline) and dNTP mix (New England Biolabs), and subsequently TOPO cloned using TOPO-TA cloning kit (ThermoFisher). Colonies containing the insert were expanded, and PCR amplification of the insert was performed using M13 primers. Crude PCR products were sent for Sanger sequencing (Functional Biosciences).

Primers:
M13 Forward: GTAAAACGACGGCCAGT
M13 Reverse: CAGGAAACAGCTATGAC

Indel chrX:123734364 Forward: CCCTTCACGTTAAACATATTGGA
Indel chrX:123734364 Reverse: TTGACTTACTTTTATATCCAGCCACTT

Indel chr4:66453492 Forward: TTTGGGATGATGGAGGAGAG
Indel chr4:66453492 Reverse: TCATTGTGCCACCAAGAAAC

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

2. Jinek, M. *et al.* Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science.* **343,** 6176 (2014).

3. Naiman, K. *et al.* A defect in homologous recombination leads to increased translesion synthesis in E. coli. *Nucleic Acids Res.* **44,** 7691-7699 (2016).

4. Deem, A. *et al*. Break-induced replication is highly inaccurate. *PLoS Biol*. **9**, e1000594 (2011).

5. Donnianni, R.A. and Symington L.S. Break-induced replication occurs by conservative DNA synthesis. *Proc. Natl. Acad. Sci. USA*. **110**, 13475-13480 (2013).

6. Carvalho, C.M. *et al*. Replicative mechanisms for CNV formation are error prone. *Nat Genet*. **45**, 1319-1326 (2013).

7. Clark, M.B. *et al.* The dark matter rises: the expanding world of regulatory RNAs. *Essays Biochem.* **54**, 1-16 (2013).

8. Mattick, J.S. and Makunin I.V. Non-coding RNA. *Hum Mol Genet.* **1**, R17-29 (2006).

9. Pauli, A. *et al.* Non-coding RNAs as regulators of embryogenesis. *Nat Rev Genet.* **12**, 136-149 (2011).

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

11. 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).

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

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

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

15. E. Dolgin, E.S. *et al.* Inbreeding and outbreeding depression in Caenorhabditis nematodes. *Evolution.* **61**, 1339-1352 (2007).

16. 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).

17. 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).

18. 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).

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

20. Nakajima, K. *et al.* Exome sequencing in the knockin mice generated using the CRISPR/Cas system. *Scientific reports* **6**, 34703 (2016).

21. Mianne, J. *et al.* Correction of the auditory phenotype in C57BL/6N mice via CRISPR/Cas9-mediated homology directed repair. *Genome Medicine.* **8**, 16 (2016).

22. Uchimura, A. *et. al*. Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type laboratory mice and mutator mice. *Genome Res*. 2015 Aug;25(8):1125-34.