Whole genome analysis for 163 guide RNAs in Cas9-edited mice reveals minimal off-target activity

Kevin A. Peterson, Sam Khalouei, Nour Hanafi, Joshua A. Wood, Denise G. Lanza, Lauri G. Lintott, Brandon Willis, John R. Seavitt, Robert E. Braun, Mary E. Dickinson, Jacqueline K. White, K.C. Kent Lloyd, Jason D. Heaney, Stephen A. Murray, Arun Ramani, and Lauryl M.J. Nutter*

*Corresponding author: lauryl.nutter@sickkids.ca

Supplementary Information

Supplementary Data 1. Sample and guide RNA information for whole genome sequence analysis.

Supplementary Data 2. Number of variants removed for each sample with each filtering step.

Supplementary Data 3. Variant calls for off-target sites detected in sequence analysis pipeline.

Supplementary Data 4. Intersection of target guide sites and variants identified in control samples.

Supplementary Data 5. Source data for Fig. 2 and Fig. 3b.

Supplementary Data 6. Provides methods details specific to each participating research centre.



Supplementary Figure 1. Summary of whole-genome sequencing data obtained from each KOMP2 production center showing base pair quality and depth. Lines within the boxes represent medians, lower edges the 1^{st} quartile, and upper edges the 3^{rd} quartile. The length of the box represents the interquartile range (IQR). Whiskers extend to the minimum or maximum values that fall within 1.5 X IQR of the 1^{st} or 3^{rd} quartiles, respectively. Each datapoint represents an individual sample, with the sample size per group as follows: n = 20 Baylor mice, n = 15 Jax mice, n = 24 TCP mice, and n = 19 UCD mice, all biologically independent.



chr6:118,087,940-118,091,073 (GRCm38/mm10)

Supplementary Figure 2. Comparison of structural variant callers for Rasgef1a deletion. The expected region is depicted with a red box and shows reduced sequence coverage consistent with Cas9 mediated deletion using a four-guide design strategy. The deletion was successfully called by lumpy and manta but filtered out by manta due to low confidence genotype filter indicted by light gray bar. This filtering resulted in the Rasgef1a deletion failing to meet the minimum threshold of being independently identified by at least two programs.



Supplementary Figure 3. Whole-genome sequencing evidence for detected Cas9 off-target activity for sites associated with guides containing 5 or fewer mismatches. (**a-c**) IGV screen shots for each region showing coverage, read data, refSeq gene, computationally detected off-target site and position of off-target guide with +/- strand indicated below relative to PAM sequence. Events confirmed by Sanger sequencing contain a green box with ICE results displayed below indicating the detected off-target mutation and estimated percentage of allele contribution. A light blue bar is drawn above the guide sequence and a red bar above the PAM with a vertical dash line indicating Cas9 cut site at the -3 position relative to PAM.



Supplementary Figure 4. Whole-genome sequencing evidence for detected Cas9 off-target activity for sites associated with guides containing 5 or fewer mismatches. (**a-c**) IGV screen shots for each region showing coverage, read data, refSeq gene, computationally detected off-target site and position of off-target guide with +/- strand indicated below relative to PAM sequence. Events confirmed by Sanger sequencing contain a green box with ICE results displayed below indicating the detected off-target mutation and estimated percentage of allele contribution. A light blue bar is drawn above the guide sequence and a red bar above the PAM with a vertical dash line indicating Cas9 cut site at the -3 position relative to PAM.



Supplementary Figure 5. Whole-genome sequencing evidence for detected Cas9 off-target activity for sites associated with guides containing 5 or fewer mismatches. (**a-c**) IGV screen shots for each region showing coverage, read data, refSeq gene, computationally detected off-target site and position of off-target guide with +/- strand indicated below relative to PAM sequence.

1	
	-

Mouse (mm10)	🔋 dn12 🔋 dn2277.665.973-27.067.108 💿 音 4 🕨 🖗 🚺 X 📮 l
	0.1 19 19 10 10 10 10 10 10 10 10 10 10 10 10 10
BCM_E12_Plwb1_off_target.bar vorage	
	с с с с с с с х т х с х
BCM_E12_Plwb1_off_target.bar	
Reforg genes	
Off-Target Mutation	B0(512/met_071
Off-Target Guide	· · · · · · · · · · · · · · · · · · ·
INDEL CONTRIB	BUTION V SEQUENCE
0	90% TTAGCTGCAGGTTCCTCAGCTCAGCCCCAGCCCCTGGAAACCCGTGTTCCGGTTAGTAATGCAATTGCATTAGCATTTGCTAATGCAATTGCAA
-11 💻	10% T T A G C T G C A G G T T C C T C A G C T C A G C T C A G C I T G T T T C C G T T C T G A T T G T T A A T G C A T T T G C T T A T T T C A A



Supplementary Figure 6. Whole-genome sequencing evidence for detected Cas9 off-target activity for sites associated with guides containing 5 or fewer mismatches. (**a-b**) IGV screen shots for each region showing coverage, read data, refSeq gene, computationally detected off-target site and position of off-target guide with +/- strand indicated below relative to PAM sequence. Events confirmed by Sanger sequencing contain a green box with ICE results displayed below indicating the detected off-target mutation and estimated percentage of allele contribution. A light blue bar is drawn above the guide sequence and a red bar above the PAM with a vertical dash line indicating Cas9 cut site at the -3 position relative to PAM.



Supplementary Figure 7. Whole-genome sequencing evidence for detected Cas9 off-target activity for sites associated with guides containing 5 or fewer mismatches. (**a-c**) IGV screen shots for each region showing coverage, read data, refSeq gene, computationally detected off-target site and position of off-target guide with +/- strand indicated below relative to PAM sequence. Events confirmed by Sanger sequencing contain a green box with ICE results displayed below indicating the detected off-target mutation and estimated percentage of allele contribution. A light blue bar is drawn above the guide sequence and a red bar above the PAM with a vertical dash line indicating Cas9 cut site at the -3 position relative to PAM.



Mouse (mm10)	Chr1	chr1:57,957,562-57,957,672	Co 🗂 🔹 Þ 🥥 🗖 🗴	< 🖵		
	qA1 qA2 qA3 qA4	4 qA5 qB qC1.1	1 qC1.2 qC2 qC3 qC4	qCS qD qEL1 qE2.1 qE	E2.2 qE2.3 qE3 qE4 qF	qG1 qG3 qH1 qH2.2 qH3 qH4 qH5 qH6
	* *	57,957,580 bp	57,957,400 bp	111 bp 57,557,520 bp	57, 557, 640 bp	57,937,990 kp
JAX_E05_Hspb6_off_target.barr enage	(F- 52)					
	тттс	A T		c	т т	Ť
MY ENE Works of torout house	G G	G		—		
WA_E05_Hepde_on_argecoart		c				
Sequence 🗕		C C A C A G T G C T G G C T G C C		A G T C C A C A C T A G T C C T T G G	GAACAGAATCATAGTCTAGC	T T C A C A G A G C C T T C C A C A C A T A G C A A A
Off-Target Mutation				JAX_ED5_Hep16_OT1		
INDEL CONTRI	BUTION - SEQUENCE			•		
0		G C A G C A C A G T C C	CACACTAGT CC	T T G G G A A C A G A A T C /	A T A G T C T A G C T T C A C	A G A G C C T T C C A C A C A T A G C
		and the second second				
b						
Mouse (mm10)	😴 chr1	chr1:22,335,602-22,335,912	Go 音 < 🕨 🏟 🖪 🗴	< 🗭 I		
Mouse (mm10)	gA1 qA2 qA3 qA	chr1:22,335,602-22,335,912	Co 音 < > 🖗 🖪 >	<	12.3 qt2.3 qt3 qt4 q ²	- 111111111111111111111111111111111111
Mouse (mm10)	chr1 4A5 0A2 0A3 4A 9 1	chr1.22,335,602-22,335,912	Co 🗃 < > 🕸 🚺 >	4 0 0 011 023 0 053 00 011 023 0 1115p 1115p 213,000p	2.2 q12.3 q13 q44 q* 1 20.334,000 be	्रामामामामामामा वटा वटा वटा वरा वरा वरा वरा वरा स्टा वटा वरा वरा वरा वरा वरा वरा वरा वरा वरा वर
Mouse (mm10)	chr1	chr122,35,802-22,335,912	Go 🚔 ← → @p 🖪 > 1 q012 q02 q03 q04 1 223148 bp	<	12.2 qf2.3 qf.3 qf.4 q.4 1 X2.154.00 ba	- 111111111111111111111111111111111111
Mouse (mm10)	chc1	chr1.22,335,602-22,335,912	Co 🕋 < > 🖗 🚺 >	4 p	2.2 q2.3 q2.3 q44 q7 1 20.144.00 ba	
Mouse (mm10) avx, st0, Treen(71, eff, Seeper avxrage	chri dui	c chr122,335,802-22,335,912	Go ∰ ← ⊳ 420 [] > 1 ≪12 ∉27 ∉07 €04 1 2331,48 №	< C (C (C	с с , <u>уши</u> ве 3. 453 ба 44 б.	с с т т т т т т т т т т т т т
Mouse (mm10)	chr1 40.1 40.2 40.3 40.5 4 - - - 0 - - - 0 - - -	c c	Co 🔐 +	4 20 401 402 40 923 90 401 402 9 	22 4/23 4/3 4/4 4/ 1 2334/6 b 6 C C	40. 40.
Mouse (mm10) dXX_510_Them171_64[Jaquti dXX_510_Them171_64[Jaquti dXX_510_Them171_64[Jaquti	C chri qAL qAZ qAL qA 	chr122,335,602-22,335,912 chr122,335,912 chr122,355,912 chr122,355,912 chr122,355,912 chr122,355,912 chr122,355,9		4 5 00 ett. (22) 0	с с 1 <u>хинае</u> 1 <u>хинае</u>	1 1
Mouse (mm10) dox_610_1mem171_set_texpet www.p dox_610_1mem171_set_texpet dox_610_1mem171_set_texpet dox_610_1mem171_set_texpet	chri dui	C chr122,335,802-22,335,912		4 (2) (2) (2) (2) (2) (2) (2) (2) (2) (2)	C C C C	45. 45. 45. 45. 45. 45. 45. 45. 1 22. 5. 5. 4. 4. 1 22. 5. 5. 4. 1 22. 5. 5. 4. 1 22. 5. 5. 5. 1 22. 5. 5. 5. 2 2 5. 5. 5. 3 3 5. 5. 5. 4 4. 4. 5. 5.
Mouse (mm10) AM. 512, "mm171, off, langel: ownys AM, 510, Trent 71, off, langel: Geograms Review genes PT Tauet Manhou	chr1 q41 q42 q43 q4 q41 q42 q41 q42 q41 q41	C chri 22,335,802-22,335,912				461 461
Mouse (mm10) AAX_510_meet11_set_leaged average daX_510_meet11_set_leaged daX_510_meet11_set_leaged daX_510_meet11_set_leaged Contract Modelen	C A T T T C A T T T T C T T T	C T T T T T A C A G A A A G A G A G A		C C I A C A C A C A C A C A C A C C C C	22 4/23 4/3 4/4 4/ 1 2314.66 b C C C C C C	40.1 40.1 <td< th=""></td<>

Supplementary Figure 8. Whole-genome sequencing evidence for detected Cas9 off-target activity for sites associated with guides containing 5 or fewer mismatches. (**a-b**) IGV screen shots for each region showing coverage, read data, refSeq gene, computationally detected off-target site and position of off-target guide with +/- strand indicated below relative to PAM sequence. Events confirmed by Sanger sequencing contain a green box with ICE results displayed below indicating the detected off-target mutation and estimated percentage of allele contribution. A light blue bar is drawn above the guide sequence and a red bar above the PAM with a vertical dash line indicating Cas9 cut site at the -3 position relative to PAM.



Supplementary Figure 9. Whole-genome sequencing evidence for detected Cas9 off-target activity for sites associated with guides containing 5 or fewer mismatches. (**a-c**) IGV screen shots for each region showing coverage, read data, refSeq gene, computationally detected off-target site and position of off-target guide with +/- strand indicated below relative to PAM sequence. Events confirmed by Sanger sequencing contain a green box with ICE results displayed below indicating the detected off-target mutation and estimated percentage of allele contribution. A light blue bar is drawn above the guide sequence and a red bar above the PAM with a vertical dash line indicating Cas9 cut site at the -3 position relative to PAM.



Supplementary Figure 10. Whole-genome sequencing evidence for detected Cas9 off-target activity for sites associated with guides containing 5 or fewer mismatches. (**a-c**) IGV screen shots for each region showing coverage, read data, refSeq gene, computationally detected off-target site and position of off-target guide with +/- strand indicated below relative to PAM sequence. Note, off-target mutationn shown in panel c was confirmed via CIRCLE-seq.



Supplementary Figure 11. Whole-genome sequencing evidence for detected Cas9 off-target activity for sites associated with guides containing 5 or fewer mismatches. (**a-c**) IGV screen shots for each region showing coverage, read data, refSeq gene, computationally detected off-target site and position of off-target guide with +/- strand indicated below relative to PAM sequence. Note, off-target mutations shown in panels b,c were identified via CIRCLE-seq.



Supplementary Figure 12. Comparison of predicted off-target mismatch allowance, variant class and positional dependence. (**a**) The proportion of variants overlapping a predicted off-target guide site in control animals compared to Cas9 gene-edited (+Cas9). All of the variants detected in control samples were associated with guides with 5 or 6 mismatches. (**b**) Percentage of mismatches identified in off-target guides allowing up to 5 or 6 mismatches and one DNA/RNA bulge. Positions are numbered with -1 adjacent to NRG PAM sequence and -20 most distal.