YMTHE, Volume *28*

Supplemental Information

Doxycycline-Dependent Self-Inactivation of CRISPR-Cas9 to Temporally Regulate On- and Off-Target Editing

Anju Kelkar, Yuqi Zhu, Theodore Groth, Gino Stolfa, Aimee B. Stablewski, Naina Singhi, Michael Nemeth, and Sriram Neelamegham

Supplemental Methods

Vector construction: cDNA corresponding to the P2A self-cleaving peptide was formed by annealing oligos P2A-1 and P2A-2 (Supplemental Table S1A) containing 5'-NheI and 3'-MluI, a process that also introduces new HpaI and SpeI sites 3' of P2A. This product was inserted into LentiCRISPRV1 (Addgene, Plasmid# 49535) to replace the original P2A-Puromycin cassette, located between NheI and MluI. Next, dTomato was amplified from template vector using dT-F and dT-R primers with 5'-HpaI and 3'-SpeI overhangs (Supplemental Table S1A), and the product was cloned into the above vector to create 'LentiCRISPR-dTomato'. LentiCRISPR-dTomato was then digested with EcoRV and SacII creating a 3,606 bp fragment spanning the 3' end of Cas9 though the 5' end of the WPRE motif, and this fragment was cloned in LentiCRISPRV2 (Addgene plasmid 52961, [1]) digested with the same enzyme-pair. This final constructs is called 'SiC-V1'. Various guides targeting the genes ST3G4, EMX1 and VEGFA (Supplemental Table S1C) were selected from previous publications [2, 3]. These were derived by annealing complementary oligos with BsmBI sites and cloned into 'SiC-V1'. A scramble guide with no homology to humans served as control (Supplementary Tables S2, S3).

To create 'SiC-V2', the IRES-puro cassette in Tet-pLKO-puro (Addgene plasmid 21915, [4]) was excised using SmaI and KpnI, and this was replaced by a P2A encoding cDNA by annealing oligos P2A-3 and P2A-4 (Table S1A). In addition to preserving the previous enzyme sites, this cloning step also introduced HpaI immediately 5' of KpnI. Cerulean was then amplified using primers Cer-F and Cer-R from template plasmid (Addgene plasmid 64847, [5]) with 5'-HpaI and 3'-KpnI sites for insertion at this site. This final product is called 'SiC-V2'. CRISPR guides G1-G10 (Supplemental Table S1B) to edit the Cas9 gene were designed using CRISPR design tool (www.crispr.mit.edu/). Here, following annealing of complementary oligos for each target sequence, the product was gel extracted and cloned into the AgeI/EcoRI sites of 'SiC-V2' immediately following the H1/TetO promoter. All clones were confirmed by Sanger sequencing. Plasmids were prepared using Nucleobond Xtra Midi EF kit (Macherey Nagel, Bethlehem, PA) to obtain endotoxin-free DNA for cellular transfections.

Lentivirus preparation: HEK 293T cells were plated in 5-10, 15cm dishes 1-day prior to transfection. Lentiviral constructs along with pMD2.G and psPAX2 (Addgene plasmid #12259 and #12260) packaging plasmids was transfected into cells at 50-60% confluency, using calcium

phosphate method [6]. 6-8h post-transfection, medium was changed to Opti-MEM with GlutaMAX. The 1st virus batch was collected 18-20h thereafter, with additional growth medium being added along with 10mM sodium butyrate. A 2nd virus batch was then subsequently collected after 18-20h. Both batches were pooled, centrifuged at 1000g for 5 min to remove cell debris and the supernatant was passed through a 0.45μm filter. Following, ultracentrifugal concentration at 50,000xg for 2h, viral pellets were resuspended in 50-60μl IMDM medium to make a 500X virus. This reagent was aliquoted and stored at -80^{0}C until use. Transduction was done in the presence of 8μg/ml Polybrene as previously described [6]. Doxycycline (Clontech, Mountain View, CA) was added to cells at times and concentrations stated in Results.

References

1. Sanjana, NE, Shalem, O, and Zhang, F (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* **11**: 783-784.
2. Fu, Y, Foden, JA, Khayter, C, Maeder, ML, Reyon, D, Joung, JK, *et al.* (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**: 822-826.
3. Mondal, N, Buffone, A, Jr., Stolfa, G, Antonopoulos, A, Lau, JT, Haslam, SM, *et al.* (2015). ST3Gal-4 is the primary sialyltransferase regulating the synthesis of E-, P-, and L-selectin ligands on human myeloid leukocytes. *Blood* **125**: 687-696.
4. Wiederschain, D, Wee, S, Chen, L, Loo, A, Yang, G, Huang, A, *et al.* (2009). Single-vector inducible lentiviral RNAi system for oncology target validation. *Cell Cycle* **8**: 498-504.
5. Dayananda, KM, Gogia, S, and Neelamegham, S (2011). Escherichia coli-derived von Willebrand factor-A2 domain fluorescence/Forster resonance energy transfer proteins that quantify ADAMTS13 activity. *Anal Biochem* **410**: 206-213.
6. Buffone, A, Jr., Mondal, N, Gupta, R, McHugh, KP, Lau, JT, and Neelamegham, S (2013). Silencing alpha1,3-fucosyltransferases in human leukocytes reveals a role for FUT9 enzyme during E-selectin-mediated cell adhesion. *J Biol Chem* **288**: 1620-1633.

Supplemental Table S1A. Primers used to create SiC vector system	
Oligo Name	Sequence
P2A-1	5'-CTAGCGGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGG CCGGCGACGTGGAGGAGAACCCCGGCCCCGTTAACACTAGTA-3'
P2A-2	5'-CGCGTACTAGTGTTAACGGGGCCGGGGTTCTCCTCCACGTCG CCGGCCTGCTTCAGCAGGCTGAAGTTGGTGGCGCCGCTGCCG-3'
P2A-3	5'-GGGAGTTCGCTAGCGGCAGCGGCGCCACCAACTTCAGCCTGC TGAAGCAGGCCGGCGACGTGGAGGAGAACCCCGGCCCCGTTAAC GGTAC-3'
P2A-4	5'-GGTACCGTTAACGGGGCCGGGGTTCTCCTCCACGTCGCCGGC CTGCTTCAGCAGGCTGAAGTTGGTGGCGCCGCTGCCGCTAGCGA ACTCCC-3'
dT-F	5'- GCATTAGTTAACATGGTGAGCAAGGGCGAGGAG-3'
dT-R	5'-GCATTAACTAGTTTACTTGTACAGCTCGTCCAT-3'
Cer-F	5'- GCATTAGTTAACATGGTGAGCAAGGGCGAGGAG-3'
Cer-R	5'-GCATTACCATGGAATGAACATGTCGAGCAGGTA-3'

Supplemental Table S1B: Target sequences for Cas9 gene and PCR primers used for Surveyor assay

Guide	Target sequence-**PAM** (5' – 3')	Target Strand	Forward primer * (5' – 3')	Reverse primer * (5' – 3')	% Editing in presence of Dox
G1	CTTGTACTCGTCGG TGATCA-**CGG**	Antisense	ACACAGGACCGG TTCTAGAGC	CAGATCCGGTTCTTCC GTCTG	8
G2	TACAGCATCGGCCT GGACAT-**CGG**	Sense	ACACAGGACCGG TTCTAGAGC	CAGATCCGGTTCTTCC GTCTG	16
G3	GACATCGGCACCAA CTCTGT-**GGG**	Sense	ACACAGGACCGG TTCTAGAGC	CAGATCCGGTTCTTCC GTCTG	10
G4	GCTGGGCACCTTGT ACTCGT-**CGG**	Antisense	ACACAGGACCGG TTCTAGAGC	CAGATCCGGTTCTTCC GTCTG	18
G5	CTTCTTGATGCTGTG CCGGT-**CGG**	Antisense	ACACAGGACCGG TTCTAGAGC	CAGATCCGGTTCTTCC GTCTG	10
G6	GCCCACATGATCAA GTTCCG-**GGG**	Sense	GTGGCCTACCAC GAGAAGTACCC	GCAGACAGGATGGCC TTGGCG	20
G7**	**TACGCCGGCTACAT TGACGG-CGG**	**Sense**	**ATACGACGAGC ACCACCAGG**	**GGGATGCTGCCGTTGT CGAAG**	**30**
G8	TTTGGTCAGCTCGTT ATACA-**CGG**	Antisense	TGGTGGACAAGG GCGCTTCC	TCCACGGAGTCGAAGC ACTCG	8
G9	GTTGAACCGATCTTC CACGC-**CGG**	Antisense	GCGAGCAGAAAA AGGCCATCGTG	CAGCCGTTCCTCGATC ATCTCTCT	10
G10	GAATGGGCGGGATA TGTACG-**TGG**	Sense	AGCGGATCGAAG AGGGCATCA	ACCTCTTCGGAGGGCA CGTT	20

* Same primer sets was used for G1-G5 as these sgRNA are proximal in Exon-1. Additional primers amplify other target sites.
**** sgRNA sequence used in this study and referred to as 'Cas9G7 sgRNA'.**

Supplemental Table S2: SgRNA binding sites and PCR primers used for fragment analysis or surveyor assay

Guide	Target sequence-**PAM** (5' – 3')	Target Strand	Forward primer with 5' universal primer overhang (5' – 3') *	Reverse primer (5' – 3')
EMX1-T	GAGTCCGAGCAGAAGAAGAA-**GGG**	Sense	AGCTGACCGCGACGAAAATTGGGCCCCTAACCCTATGTAGCCTCAG	CTGCCCTCGTGGGTTTGTGGTTG
EMX1-OT **	GAGTTAGAGCAGAAGAAGAA-**AGG**	Sense	AGCTGACCGCGACGAAAATTGCAAAGAAATGCCCAATCATTGATGC	CCCCGCTTGTCCATGTCTAGG
VEGFA-T	GACCCCCTCCACCCCGCCTC-**CGG**	Antisense	AGCTGACCGCGACGAAAATTGGTGAGTGACCTGCTTTTGGGGG	GCTAGCACTTCTCGCGGCTC
VEGFA-OT **	GACCCCCCCCACCCCGCCCC-**CGG**	Sense	AGCTGACCGCGACGAAAATTGGGCAAGATTCTGGAAGGGTCTTG	GCAGAAGCAACCCTCAGCACC
6-FAM universal forward primer	NA	NA	6-FAM-AGCTGACCGCGACGAAAATTG	NA
Scramble guide (Scr)	GCGTCGTCCGTCGCTACACA	NA	NA	NA
MGAT1	GTGGGGCGCTATCCTCTTTG-**TGG**	Sense	GCATCCCTAGGACTGCGGGC	CGCCATCGAGAGCGCTGACT
ST3Gal4	AGAAATAATCCTCAAGCCGC-**AGG**	Antisense	ATTAGTCATGGGCCTAACCCTAC	TGGACGGCAGAGAATAGGAAG
COSMC	GAGTCTTTGGGCTGCAGTAA-**AGG**	Sense	NA	NA

* Underlined sequences indicate universal primer overhang used to amplify 6-FAM PCR reaction
** Red, underlined text indicates nucleotides that differ between on-target and off-target sequences

Supplemental Table S3A: Off-target profile of Cas9G7 and Scr to reference human genome, as predicted by Cas Offinder		
No. of Mismatches	**Scr** GCGTCGTCCGTCGCTACACA-NGG	**Cas9G7** TACGCCGGCTACATTGACGG-NGG
1	0	0
2	0	0
3	0	0
4	4	5
5	83	91

Supplemental Table S3B: Cas9 G7 on-target (T) sequence and five top off-target (OT) sequences, with mismatched nucleotides red underlined.*				
Cas9G7 Target(T)/ Off Target (OT)	**Sequence**	**Chromosome**	**Position**	**Direction**
Cas9G7_T	TACGCCGGCTACATTGACGG	NA	NA	NA
Cas9G7_OT1	TACTCTGGCTACATTAGCGG	Chr12	54525245	+
Cas9G7_OT2	TAGGGCGTCTACATTCACGG	Chr7	1.52E+08	+
Cas9G7_OT3	TACTCAGGCCACATTGAGGG	Chr5	33639836	-
Cas9G7_OT4	TGAGCAGGCTACATTGAGGG	Chr10	37828244	-
Cas9G7_OT5	TACGCCGAGTAAATTGAAGG	Chr9	80632708	+
* Abbreviation NA: not applicable; additional primer information provided in Table S11				

Supplemental Table S4: SgRNA on-target sites for promiscuous sgRNA-library used for exome sequencing study

Guide	Target sequence-PAM (5' – 3')	Target Strand
EMX1	GCCCTCGCAGCTGCTGCGGCTGG	Sense
C1GalT1C	GTAGGTGATGATGCTCATGGTGG	Sense
ZSCAN2	ATACATCAGCGAATCCACACTGG	Sense
MGAT1	CGGCTGGGCGGGAGGGGCCGCGG	Sense
ST3Gal4	GGGGATGGAGGAGCTGGTGATGG	Antisense

Provided in separate excel workbook (all related to experiment in Figure 2)

Table S5-S9 legend: These tables contain the target gene sequence (human EMX1, ZSCAN2, MGAT1, ST3Gal4 or C1GalT1C) and a defined target length of 20 bases as predicted by CCTop software. The core/seed length is specified to include 12 bases from the PAM sequence, and these are flanked by square brackets in the alignment field ('[ ]'). Potential off-targets were identified by parsing the human genome (GRCh37/hg19) to identify intronic and exonic regions with 0-2 base mismatches in the core/seed region and up to 4 total mismatches in the 20 base target sequence. For each of these putative off-targets, the Table provides: i. gene name, gene id, strand information and chromosome location including start and end positions; ii. Number of mismatches, with the symbol '|' identifying matched bases and '-' mismatched bases. iii. "distance" specifies how far the predicted off-target event is from the nearest gene, in the case of intergenic and intronic off-target predictions. The on-target and selected off-target sites from these tables were also amplified using PCR, and amplicons analyzed using deep sequencing. The on- and off-target sequences selected for such analysis are marked using bold fonts in Tables S5-S9, and listed together in Table S10. Full PCR amplicons are listed in Table S11.

Table S10 legend: On-target and off-target gene sequences and summary data for human EMX1, ZSCAN2, MGAT1, ST3Gal4 and C1GalT1C, investigated in the deep sequencing NGS experiment. This table is filtered from Tables S5-S9. On- and off-target editing at these sites was quantified using multiplex PCR amplification followed by NGS.

Table S11 legend: Table lists all forward and reverse primers, target sequences and other data used to generate amplicons that were multiplexed and analyzed using NGS.

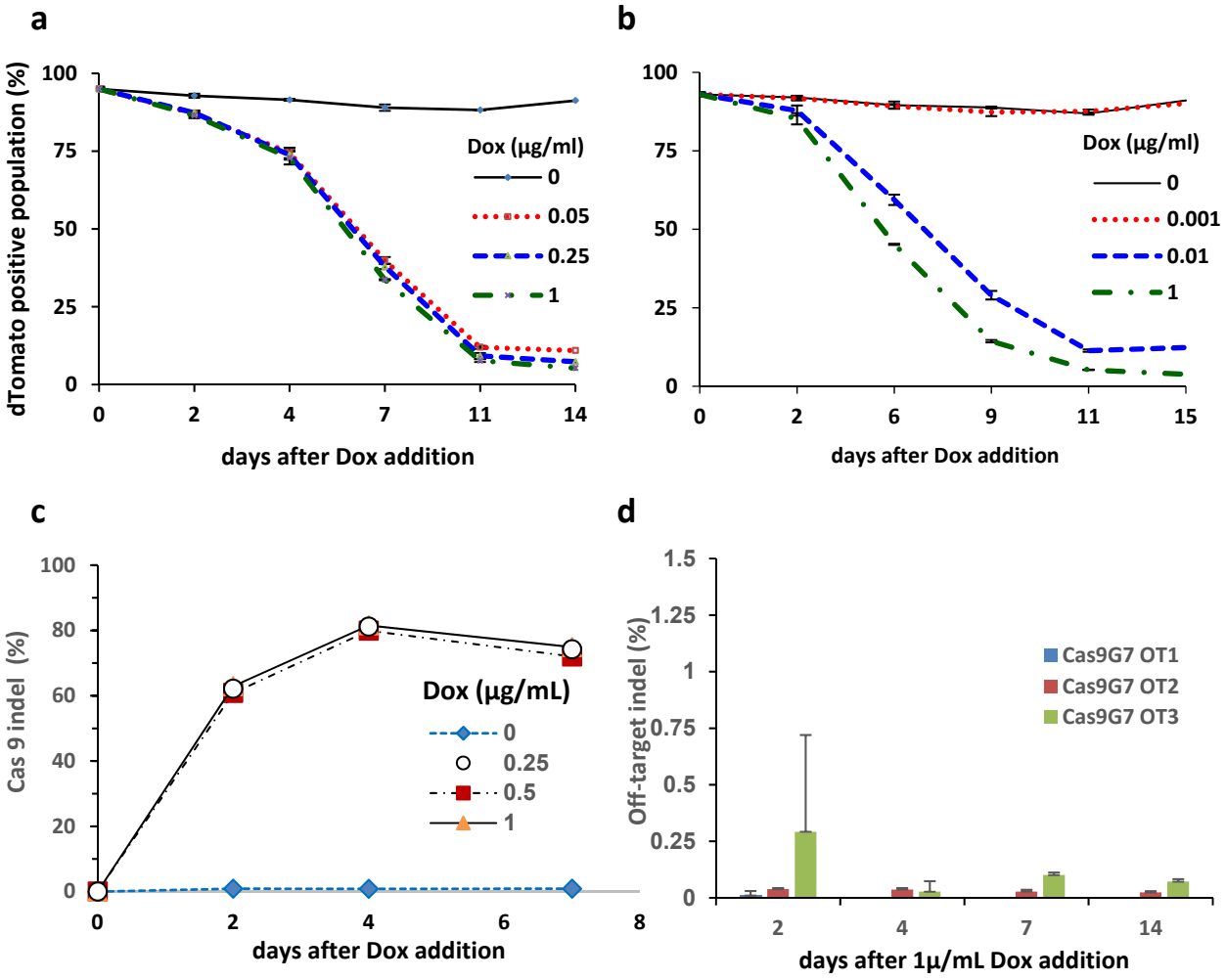


Figure S1 (Related to Fig 1). Cas9 on- and off-target editing. a-b. 'HEK-Cas9dTomato-Cas9G7' cells from Fig. 1 were treated with different concentrations of doxycycline (Dox) in two independent experiments shown in panels **a**-**b**. 90% of the cells did not express dTomato fluorescence by day 11 when Dox was >0.01µg/ml. Dox dose dependence is noted in panel **b** between 1-10ng/mL. **c-d.** Genomic DNA was isolated from study in panel **a**. PCR was performed to amplify the region surrounding the on-target editing site in Cas9 (panel **c**) and also three off-target sites (panel **d**). The target sequences are listed in Supplemental Table S3B and details about primers are provided in Table S11. 150bp paired-end sequencing of PCR amplicons was performed to detect indels. >80% Cas9 gene editing was measured by 4-days. Off target editing was <0.5% in all cases. Data in panels c, d are from triplicate runs, with error bars being too small to be visible in many instances.

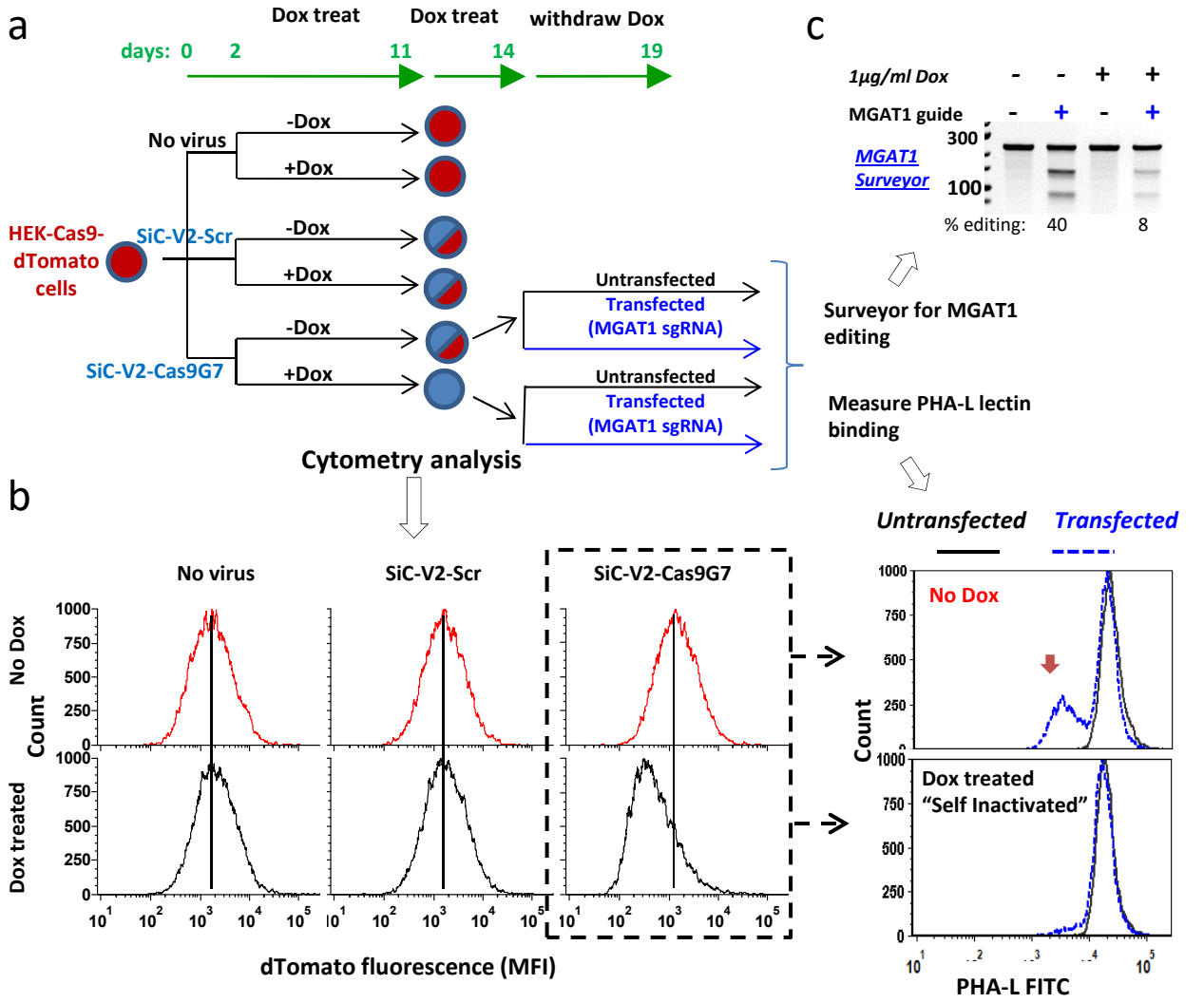
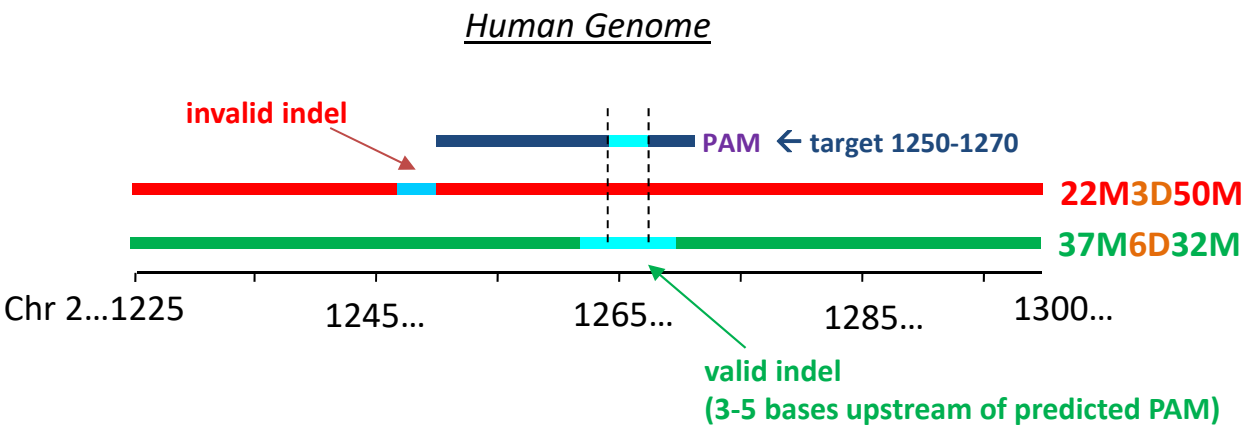


Figure S2 (Related to Fig 1). Dox-dependent editing by SiC-Cas9G7 abrogates CRISPR-based editing. a. HEK-Cas9-dTomato^{+} cells (from Fig. 1) were either untransduced, transduced with SiC-V2 with scramble sgRNA (Sic-V2-Scr) or Cas9G7 sgRNA (Sic-V2-Cas9G7). On day 2, the cells were split into two groups with 1µg/ml Dox being added to half of them for another 9 days. **b.** Flow cytometer analysis on day 11 confirmed dTomato-loss/Cas9-editing in Sic-V2-Cas9G7 transduced cells (dashed rectangle) upon Dox addition. **c**. Sic-V2-Cas9G7 transduced cells with and without Dox were transfected on day 14 with plasmid fragment containing a U6 promoter that drives a sgRNA previously shown to target *MGAT1* (Stolfa *et al*. *Sci Rep.* 6:30392, 2016, Addgene 80009). Untransfected cells served as control. Five days post-transfection (day 19), Mgat1 editing was measured based on surveyor (top right) and reduction in PHA-L-FITC (*Phytohemagglutinin-L-FITC*) lectin binding (bottom right). In the absence of Dox, 40% MGAT1 editing was measured using the surveryor assay and 30% of the cells displayed reduced PHA-L-FITC binding (red arrow). Dox-treatment reduced MGAT1 activity (Surveyor assay) and prevented loss of PHA-L binding (flow cytometry) to 8-10%. (***Notes***: i. PHA-L binds branched N-glycans that are absent upon knocking out Mgat1; ii. plasmid fragment for above study was obtained by gel purification of BglII digested plasmid Addgene# 80009).



Supplemental Fig S3 (Related to Fig 2). Counting indels during exome analysis. To detect for indels, we parsed for "D" (deletion) or "I" (insertion) in the predicted on- and off-target regions of the exome CIGAR string. Valid indels (shown in green), that were counted, had insertions/deletions 3-5 bases upstream of the PAM sequence, in the computational predicted on/off-target site. This is shown using the overlap of blue bars. Indels that were further away (shown in red) were not included in the analysis.

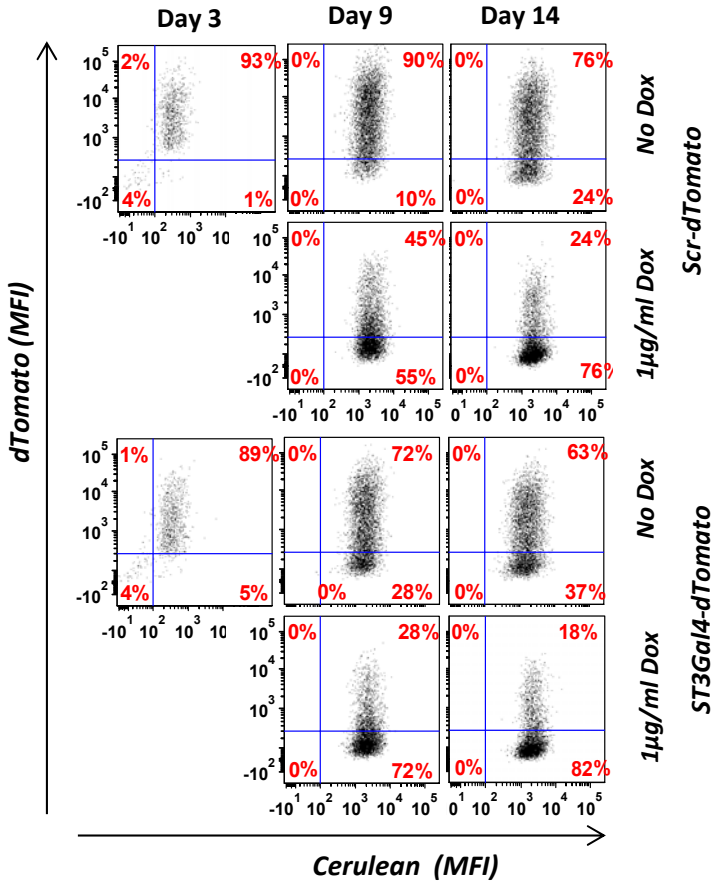


Figure S4 (Related to Fig. 4). SiC results in specific loss of dTomato reporters in HL60 cells treated with Dox. HL-60 cells were transduced with SiC-V2-Cas9G7 and either Sic-V1-Scr (top half) or SiC-V1-ST3Gal4 sgRNA lentivirus (bottom half). Cells expressing both constructs (dTomato^{+} CFP $^{+}$) were sorted on day 3 (left column). Dox was added to half the sorted cells on day 4 and both population were evaluated by flow cytometer for CFP and dTomato signal at day 9 (middle column) and day 14 (right). Loss of dTomato signal is observed in the presence of Dox in both cases. Some dTomato negative cells are seen even without Dox, since bulk flow sorting results in a heterogenous population of cells with some cells expressing relatively low dTomato expression. Cas9 editing is however absent in these cells as confirmed using the surveyor assay (Fig. 4b, main manuscript).

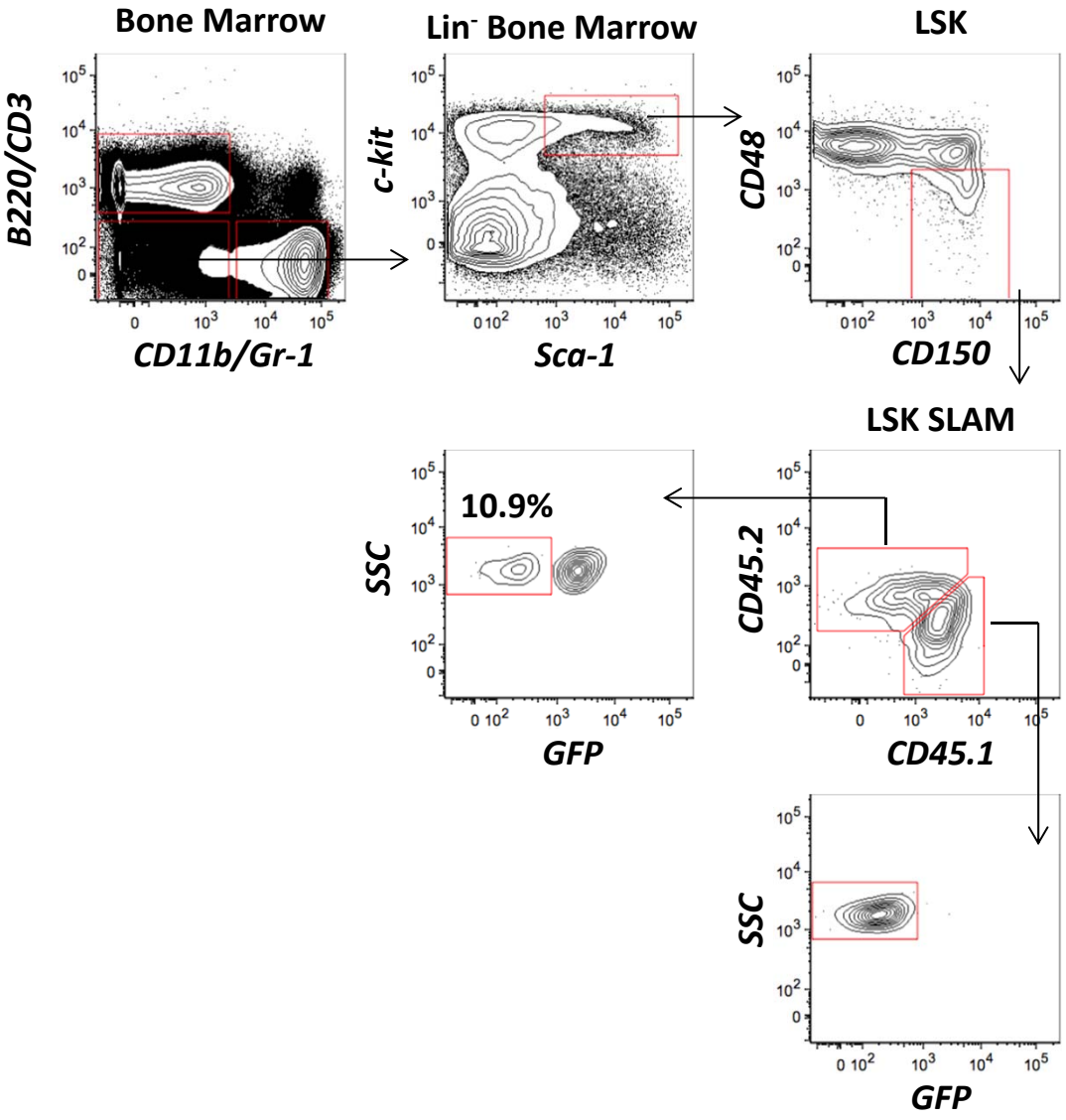


Figure S5 (Related to Fig. 6). Gating strategy for analysis of hematopoietic stem cells (HSCs; LSK SLAM) and terminally-differentiated populations in mouse bone marrow.