OMTM, Volume *27*

Supplemental information

Population-wide gene disruption in the murine lung epithelium via AAV-mediated delivery of CRISPR-Cas9 components

Honglin Chen, Steffen Durinck, Hetal Patel, Oded Foreman, Kathryn Mesh, Jeffrey Eastham, Roger Caothien, Robert J. Newman, Merone Roose-Girma, Spyros Darmanis, Soren Warming, Annalisa Lattanzi, Yuxin Liang, and Benjamin Haley

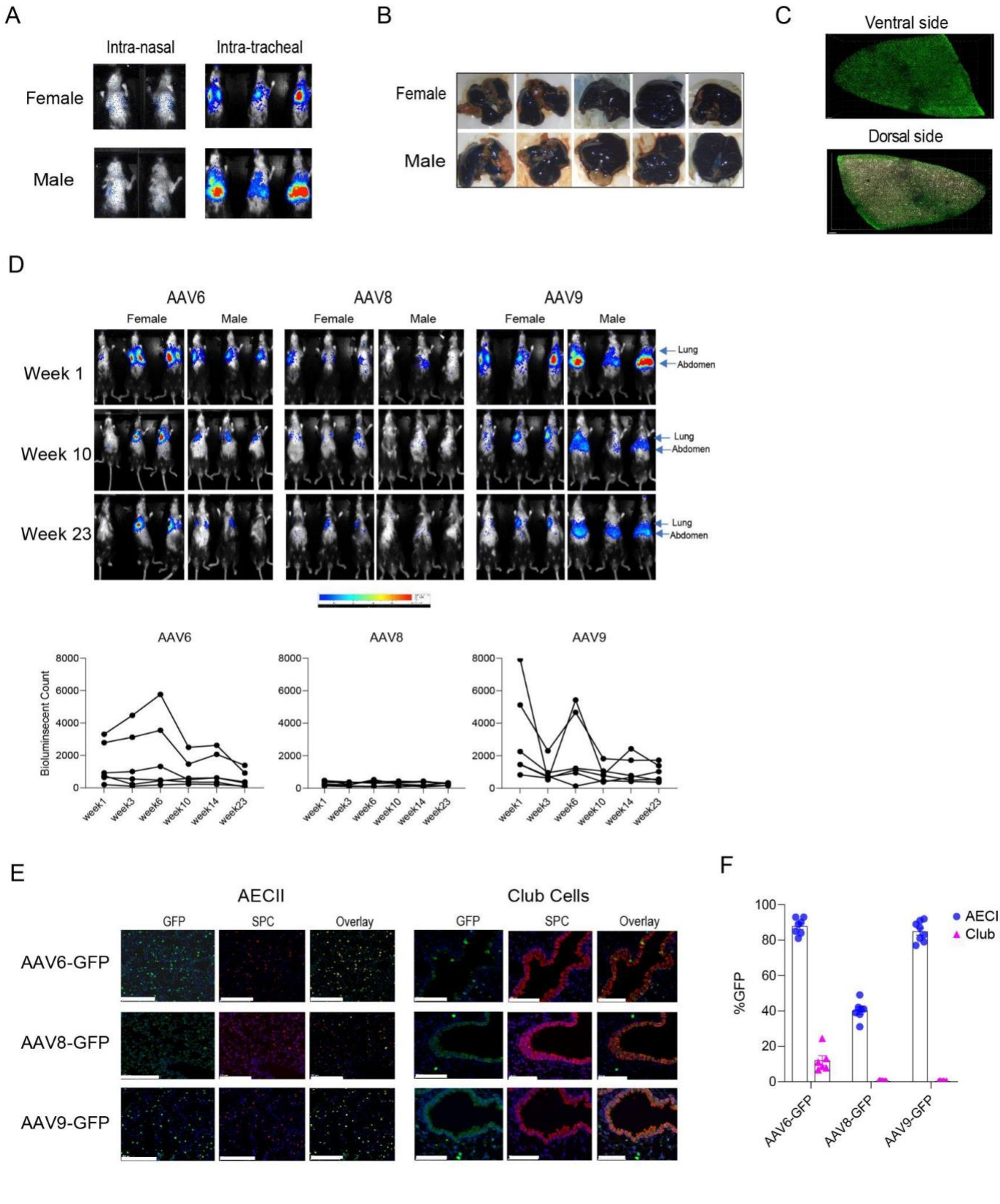

A
Intra-nasal
Intra-tracheal
Female
Male
B
Female
Male
C
Ventral side
Dorsal side
D
AAV6
AAV8
AAV9
Female
Male
Week 1
Week 10
Week 23
Lung
Abdomen
Bioluminsecent Count
8000
6000
4000
2000
0
week1
week3
week6
week10
week14
week23
E
AECII
Club Cells
GFP
SPC
Overlay
AAV6-GFP
AAV8-GFP
AAV9-GFP
F
%GFP
100
80
60
40
20
0
AECII cells
Club cells
AAV6-GFP
AAV8-GFP
AAV9-GFP

Figure S1. **Evaluation of delivery methods and transduction efficiency of AAV6, 8 and 9 in murine lung.**

A. Comparison of i.t. and i.n. delivery methods. AAV9 -CAG -luciferase virus at 1e11 vg/mouse was delivered by either method. *In vivo* bioluminescent images were taken at week 1 post-delivery. Bioluminescent signal was detected at the lung area post i.t. delivery. **B.** Representative images of lung post i.t. delivery of 4% trypan blue dye in 50ul PBS. Pictures of lungs were taken immediately after delivery. **C.** Spatial distribution of AAV transduced cells in the lung. Animals were delivered with AAV9-CAG-GFP at 2e11 vg/mouse by i.t method. 3D imaging of the lung lobe was done at week 8 post virus delivery using a Leica Sp8 confocal microscope. The laser can detect GFP signals from surface to 2mm under the surface. GFP expression was observed throughout alveolar space as well as airways. **D.** *in vivo* bioluminescent images of mice delivered with AAV-CAG-luciferase at 1e11 vg/mouse (upper panel). Male mice in the AAV9-CAG-luciferase group showed stronger bioluminescent signals at the abdomen area (indicated by the arrow). Quantification of the full body bioluminescent signal of each mouse (lower panel). **E.** Representative IF images of lung tissue sections stained with GFP and cell markers. Animals were sacrificed at week1 post AAV-CAG-GFP (1e11 vg) delivery. Scale bars indicate 200 μm (AECII) and 100μm (Club cells), respectively. **F.** Quantification of IF images as represented in **E**. %GFP was quantified in AECII and club cells, respectively. n=7. Data was obtained from two lung lobes per sample. Error bars represent SEM.

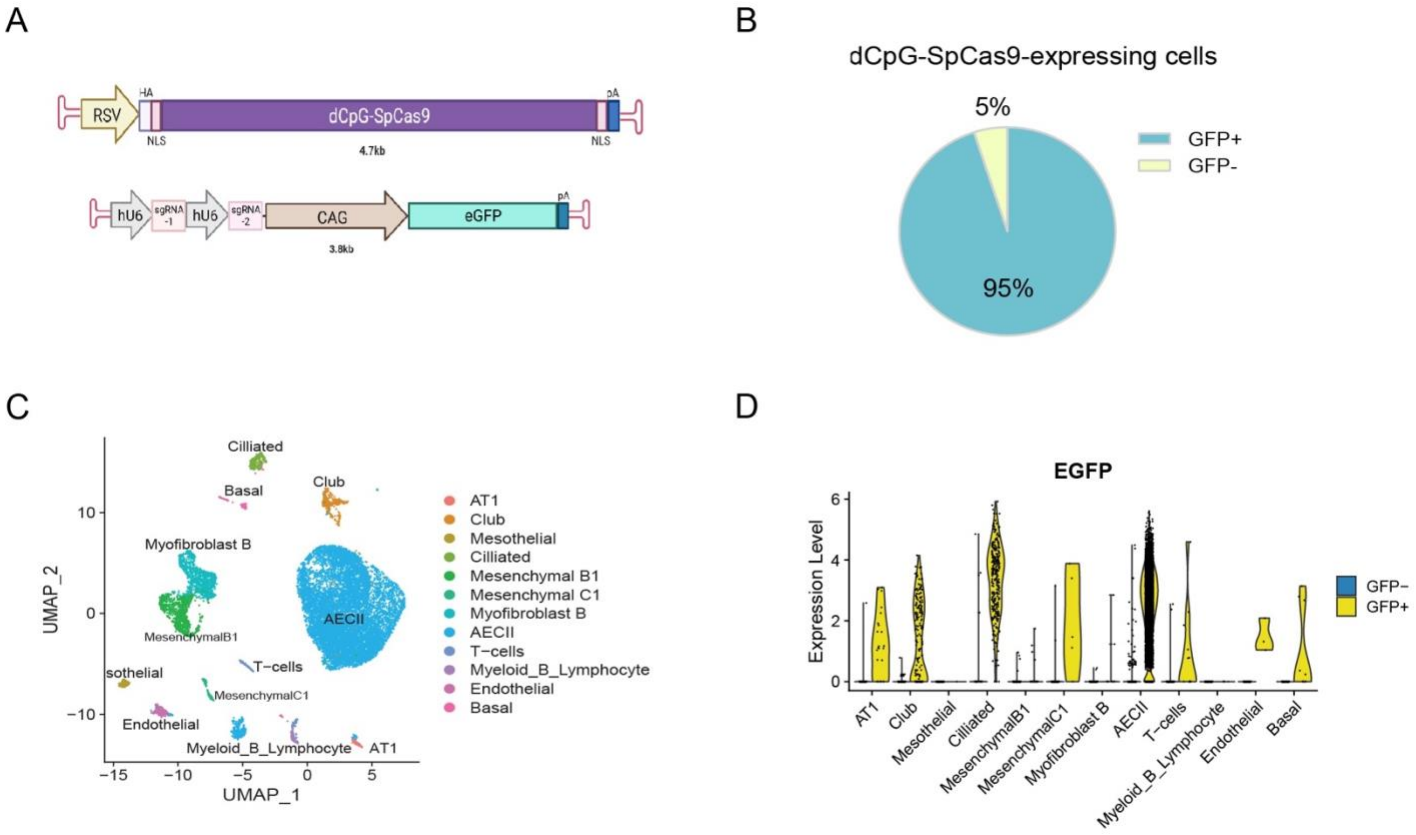


Figure S2. scRNA-seq analysis of AAV transduced cell types.
A. The diagram of AAV vectors used in this study. The dCpG-SpCas9 sequence was modified from SpCas9 with CpG islands replaced with non-CG dinucleotides. **B.** The pie chart shows the percentage of dCpG-SpCas9-expressing single cells that co-expressed GFP. **C.** UMAP clusters showing the overall cell types identified by scRNA-seq. **D.** The comparison of GFP expression levels in GFP- vs GFP^{+} samples as presented by the violin plot.

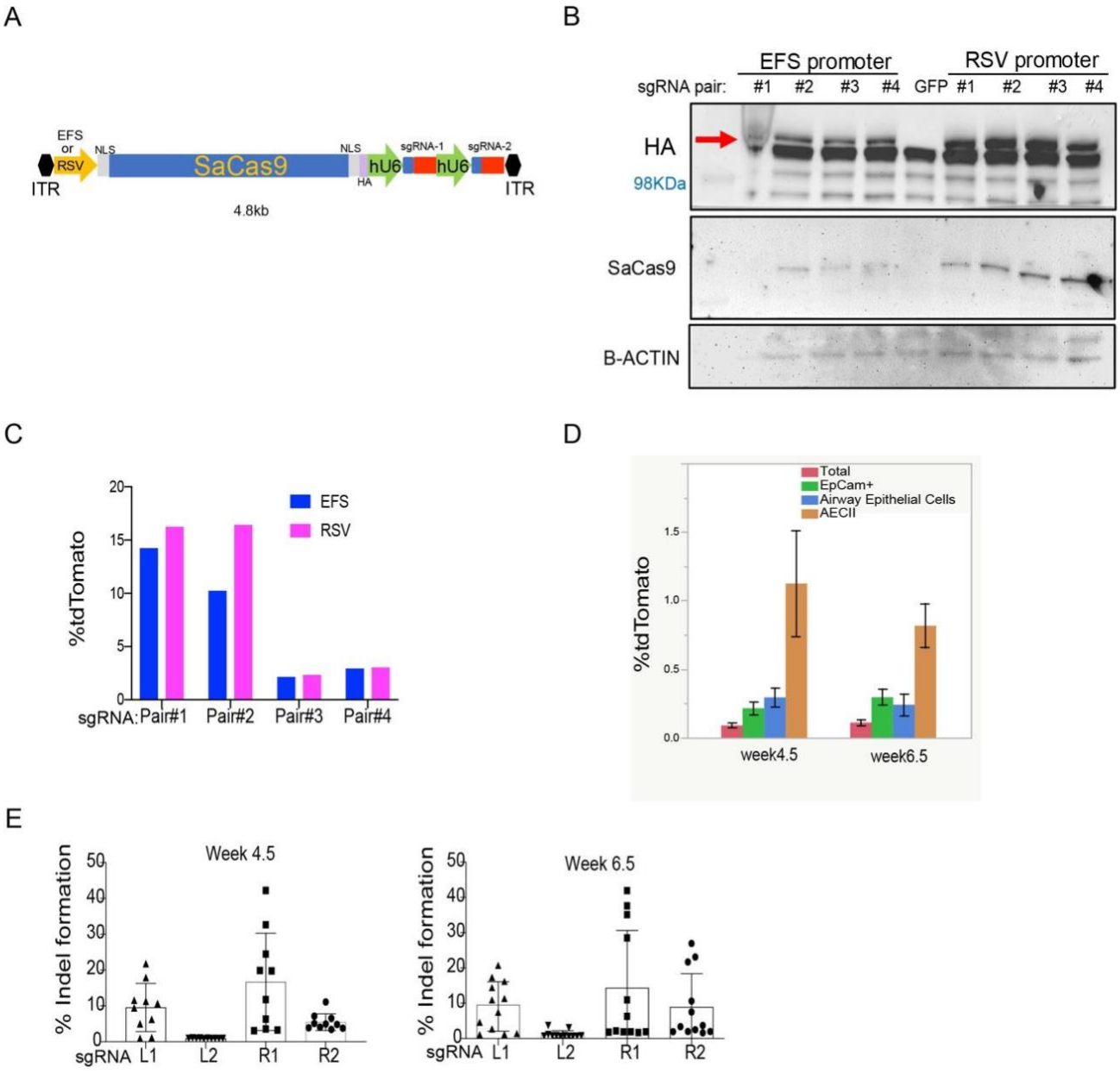


Figure S3. AAV delivery of SaCas9 CRISPR system to Rosa26-LSL-tdTomato mice.
A. The diagram of the pX601-all-in-one vector used in this study. **B.** Comparison of the two small promoters driving the SaCas9 expression. SaCas9 was tagged with HA. The constructs were nucleofected into the Rosa26-LSL-tdTomato MEFs cells. Proteins were extracted at day3 post nucleofection and subjected to Western blot analysis of the HA and SaCas9 expression. The red arrow indicated the HA band. **C.** tdTomato expression as a consequence of CRISPR/SaCas9-mediated deletion of the LSL cassette *in vitro*. The AAV constructs were nucleofected into the Rosa26-LSL-tdTomato MEFs cells and tdTomato expression was quantified by FACS at day7 post nucleofection. **D.** tdTomato expression as a consequence of CRISPR/SaCas9-mediated deletion of the LSL cassette *in vivo*. Three rounds of AAV were delivered to Rosa26-LSL-tdTomato mice with 7 days apart for each delivery. Animals were taken down at week 4.5 and week 6.5 post the 1st round of AAV delivery. The means of %tdTomato expression in different cell populations were shown. EpCam indicated the total lung epithelia; EpCam/B4 indicated distal airway epithelia; EpCam/SPC indicated AECII The error bar stands for SEM (n=10 per timepoint). **E.** NGS analysis of the % indel formation at the CRISPR/SaCas9 on-target sites. DNA was

extracted from the sorted AECII cells at week 4.5 and 6.5 timepoints. Each dot represents one animal in the scatter plots.

A

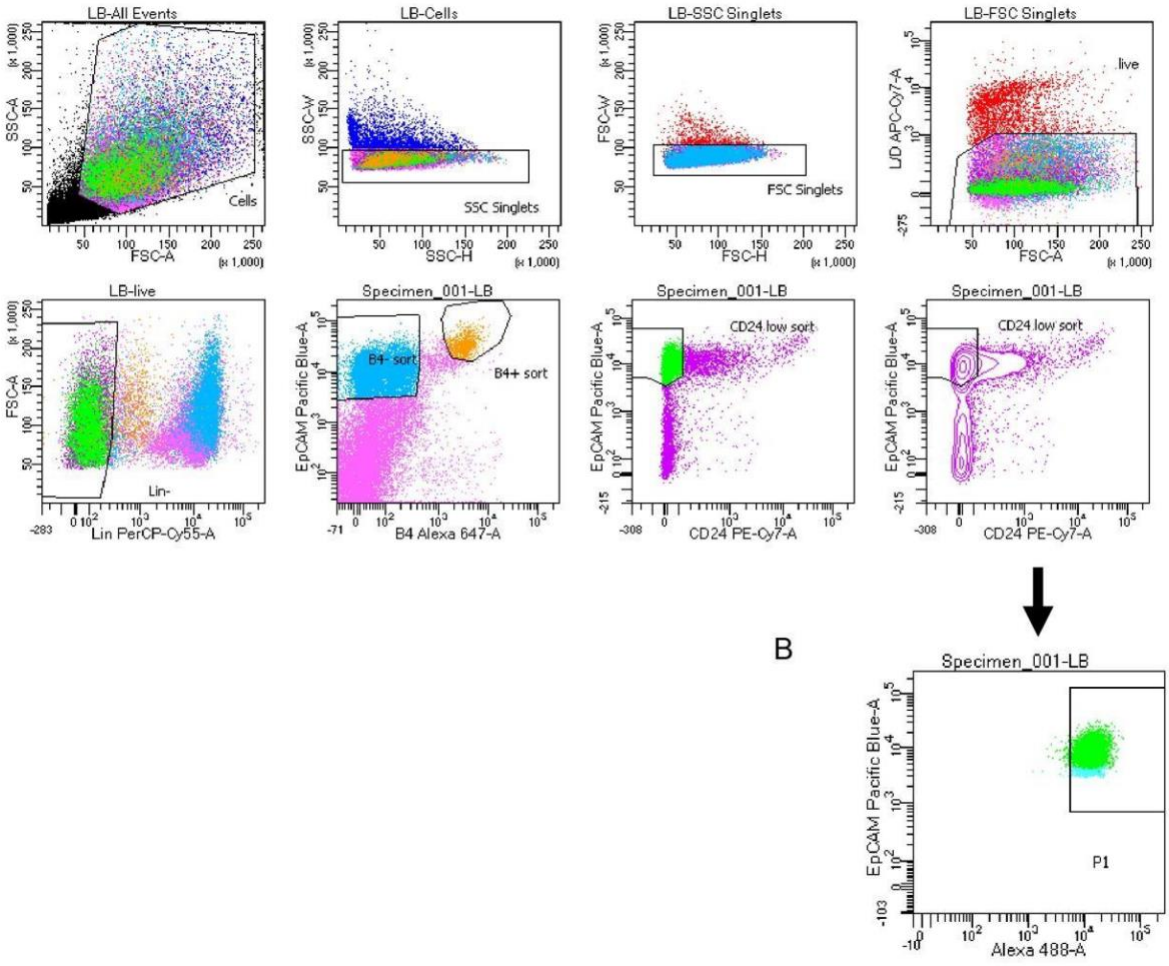


Figure S4. FACS gating schema to analyze and sort out AECII and distal airway epithelial cells.

A. Dissociated lung cells were stained with following markers: EpCam for epithelial marker; the AECII lineage markers including CD31, TER-119, CD11b, CD45 that termed as Lin, as well as the CD24 marker; the distal airway epithelial marker integrin β4. The FACS gating set-ups for sorting of AECII ($EpCam^+/Lin^-/CD24^-$) and distal airway epithelial ($EpCam^+$/ integrin $\beta4^+$) cell populations were shown. **B.** The sorted AECII population from A was permeabilized and stained with SPC (an AECII cell specific marker) antibody conjugated with Alexa-488. 98% of the population was stained positive for the SPC marker, indicating the highly enrichment of AECII in the sorted population.

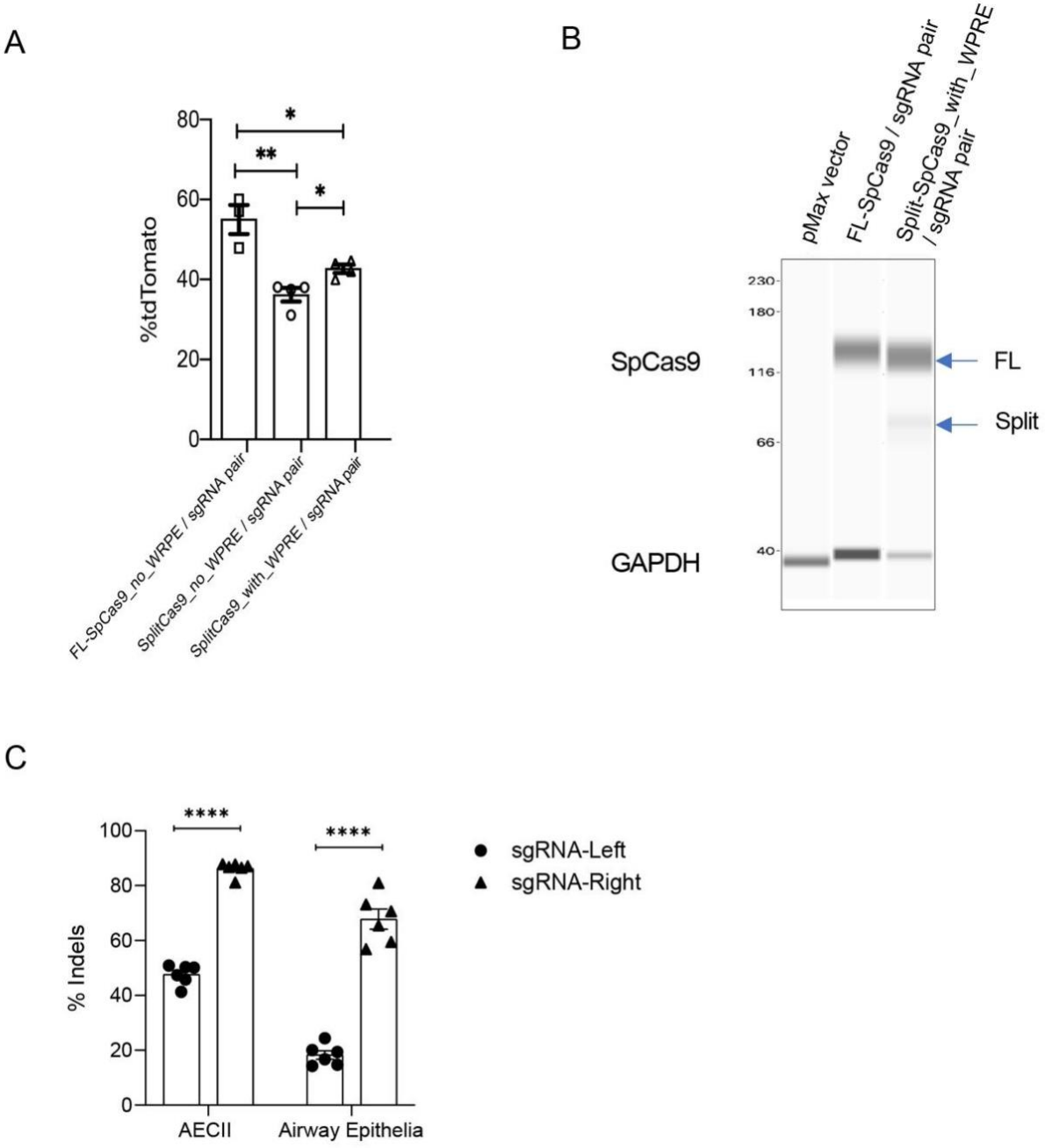


Figure S5. ***In vitro*** **assessment of AAV vector configuration#1 vs #2 in Rosa26-LSL-tdTomato MEFs.**

A. FACS analysis of %tdTomato expression as a consequence of CRISPR/SpCas9-mediated LSL cassette deletion in MEFs at day 4 post nucleofection of dual plasmids containing the CRISPR/SpCas9 components in configuration#1 (FL-SpCas9/sgRNA pair#2) or configuration#2 (Split-Cas9/sgRNA pair#2). Each dot represented the outcome from one experiment. Error bars represent SEM. * and ** denote a p-value <0.05 and <0.001, respectively. **B.** Capillary protein analysis of SpCas9 expression in MEFs. Protein lysates were extracted from MEFs at day 4 post nucleofection of dual plasmids containing the CRISPR/SpCas9 components in configuration#1 (FL-SpCas9/sgRNA pair#2) or configuration#2 (Split-Cas9/sgRNA pair#2). **C**. Indel analysis of sgRNA pair#1 in animals received one dose of AAV-SpCas9/sgRNA pair. %Indels formation at the two on-target sites was examined in the sorted AECII and airway epithelial cells at week 3 post-intubation of AAV at 2e12 vg/mouse (n=5-6). Each dot represents one animal. Error bar represents SEM. **** Denotes a p-value <0.0001.

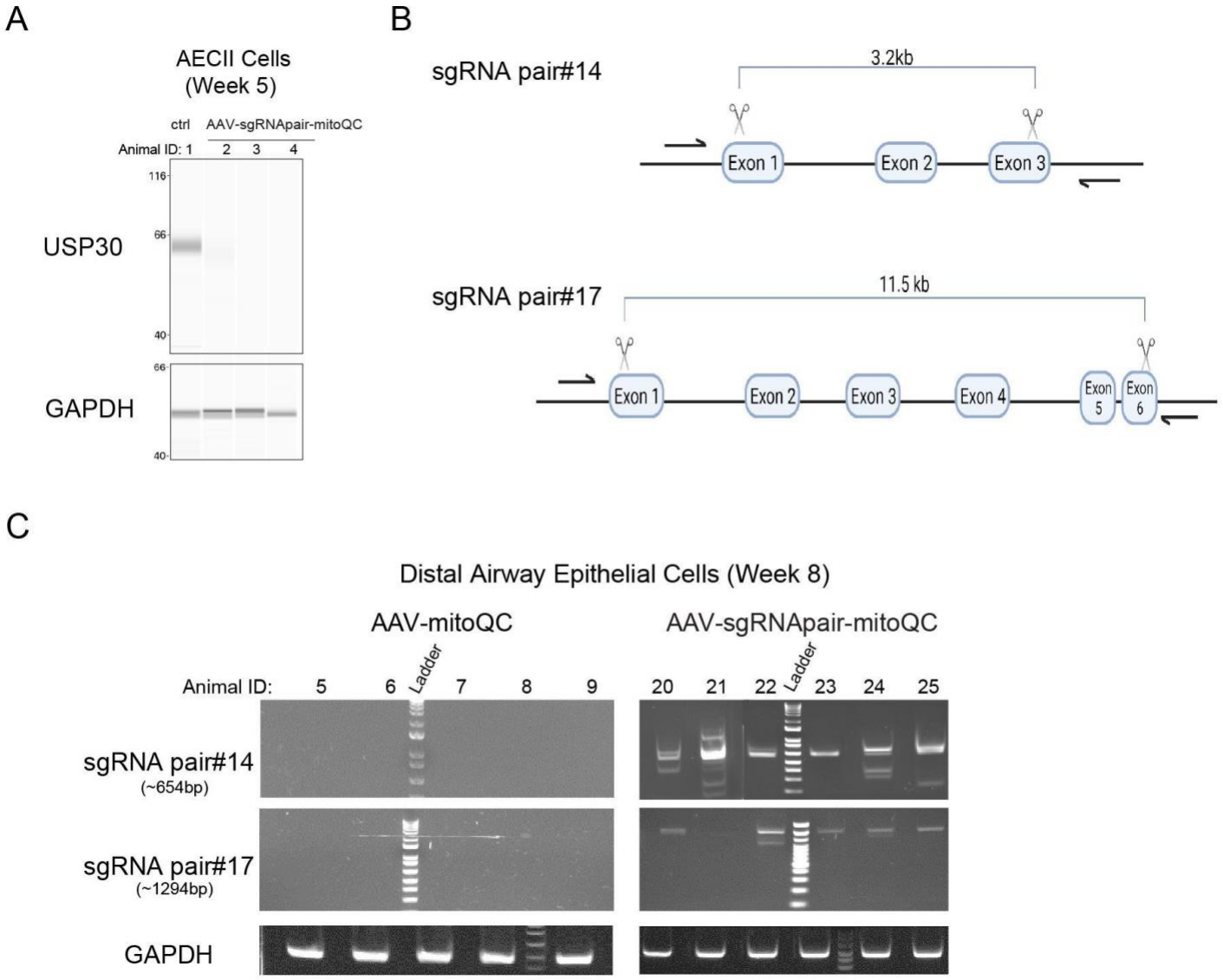
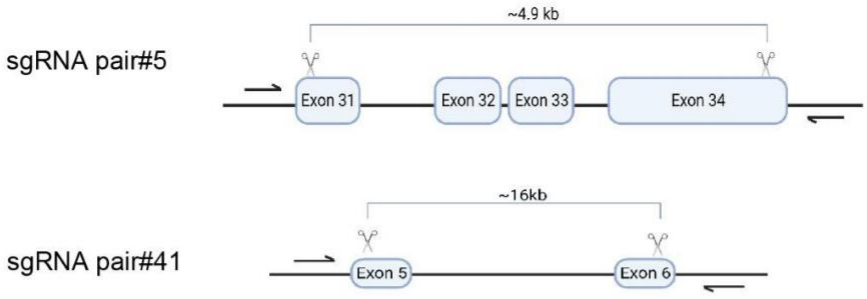
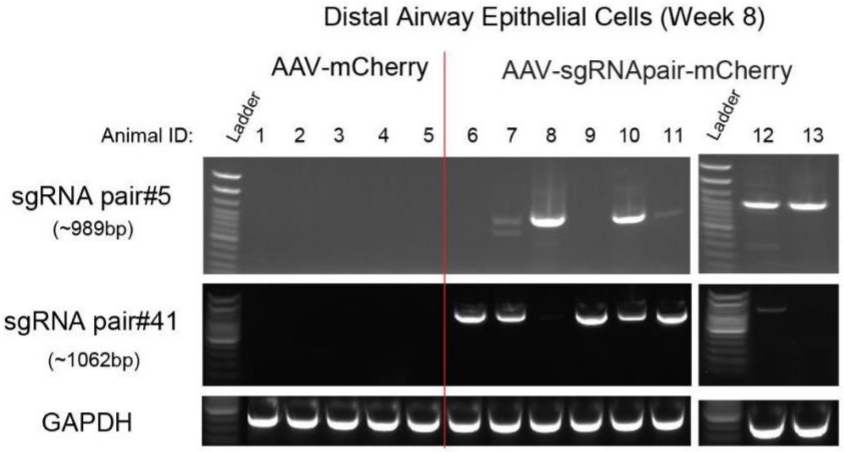


Figure S6. Analysis of USP30 protein level and genomic fragment deletion at *Usp30* locus. **A.** USP30 protein down-regulation in sorted AECII at week 5 post 1st round of AAV delivery. GAPDH served as the housekeeping control. The digital images of the capillary electrophoresis immunoblotting of USP30 and GAPDH were shown. **B.** PCR primer design to detect the deletion of USP30 exons by sgRNA pair#14 and 17. Arrows indicated where the primers were bound. The expected sizes of the PCR amplicons of the deleted genome were ~654bp and ~1294bp for sgRNA pair#14 and sgRNA pair#17, respectively. **C.** Representative gel images of PCR amplification of DNA extracted from sorted distal airway epithelial cells using primers shown in B. No PCR amplification was observed in the control group; PCR amplification was observed in the treatment group receiving either AAV-sgRNA pair#14 or AAV-sgRNA pair#17, indicating successful exon deletions in sorted airway epithelial cells. The TrackIt 100bp DNA ladder was used.

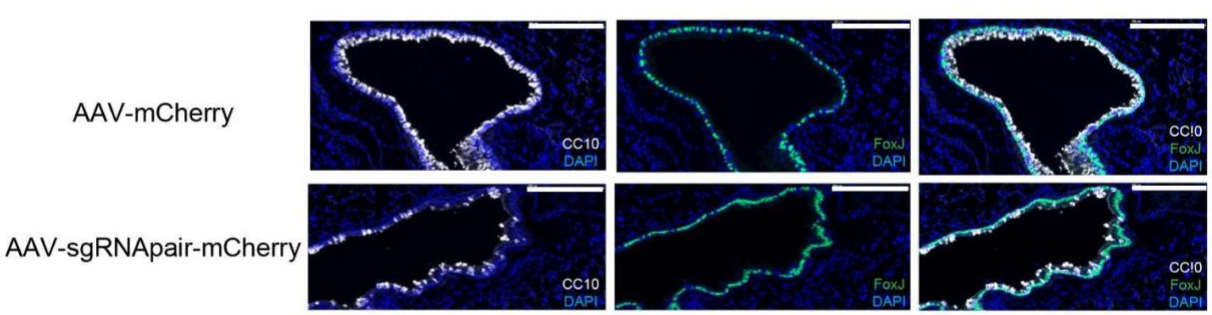
A


sgRNA pair#5
~4.9 kb
Exon 31
Exon 32
Exon 33
Exon 34
sgRNA pair#41
~16kb
Exon 5
Exon 6

B


Distal Airway Epithelial Cells (Week 8)
AAV-mCherry
AAV-sgRNApair-mCherry
Animal ID:
Ladder
1 2 3 4 5
6 7 8 9 10 11
Ladder
12 13
sgRNA pair#5
(~989bp)
sgRNA pair#41
(~1062bp)
GAPDH

C


AAV-mCherry
AAV-sgRNApair-mCherry
CC10
DAPI
FoxJ
DAPI
CC!0
FoxJ
DAPI

D

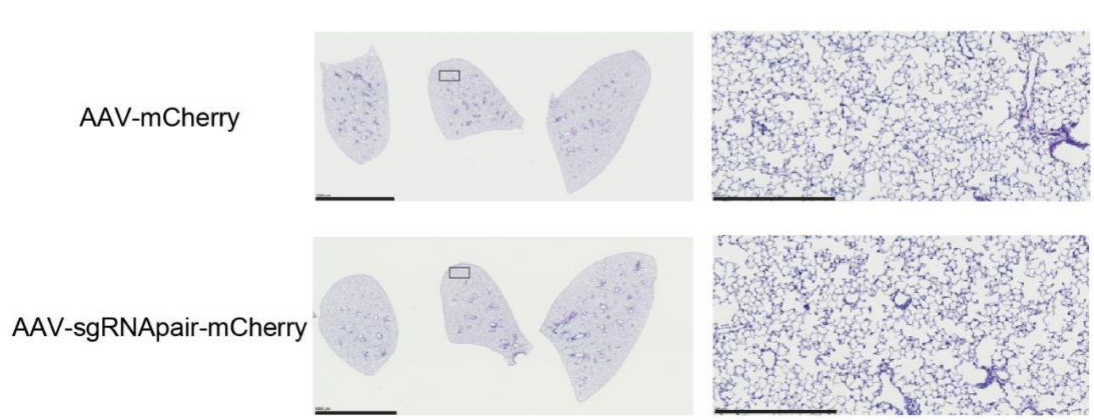

AAV-mCherry
AAV-sgRNApair-mCherry

Figure S7. Effect of CRISPR/SpCas9-mediated NOTCH2 down-regulation.
A. PCR primer design to detect the deletion of Notch2 exons by sgRNA pair#5 and 41. Arrows indicated where the primers were bound. The expected sizes of the PCR amplicons of the deleted genome were ~989bp and ~1062bp for sgRNA pair#5 and sgRNA pair#41, respectively. **B.** Representative gel images of PCR amplification of DNA extracted from sorted distal airway epithelial cells using primers shown in B. No PCR amplification was observed in the control group; PCR amplification was observed in the treatment group receiving either AAV-sgRNA pair#5 or AAV-sgRNA pair#41, indicating successful exon deletions in sorted airway epithelial cells. The TrackIt 100bp DNA ladder was used. **C.** Representative IF images of club and ciliated cells at week 8 timepoint. Club cells stained in white color (CC10: a cell marker for club cells); ciliated cells stained in green color (FoxJ1: a cell marker for ciliated cells). Cell nuclei stained by DAPI in blue. Scale bars = 200μm. **D.** Representative H&E staining of lung sections at week 4 timepoint. No morphological changes observed in the alveolar area. The right images presented the lung tissue area indicated in the black box in the left images. Scale bars =5000μm and 500μm in the left and right side of images, respectively.

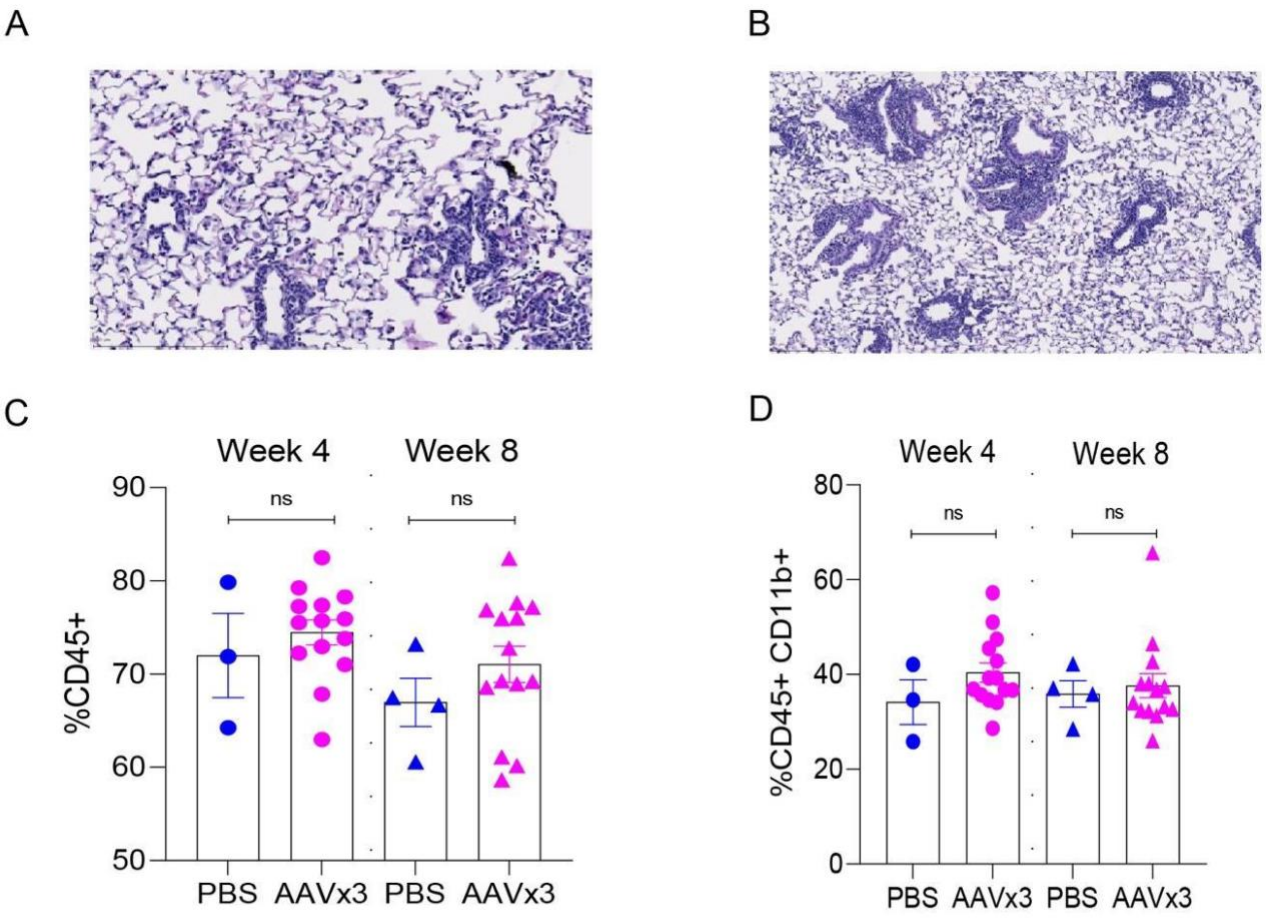


Figure S8. Minimal host inflammatory response to AAV transduction.
A. Representative H&E staining of lung tissue showing free histiocytes within the alveoli. **B.** Representative H&E staining of lung tissue showing peribronchiolar inflammation. **C and D.** FACS analysis of dissociated lung cells from SpCas9 transgenic mice received three doses of AAV6-*Notch2*-sgRNApair-mCherry at 4e12vg/dose/mouse. Cells were stained with anti-CD45 **(C and D)** and anti-CD11b antibodies **(D)**. Two timepoints were examined (Week 4 and week 8 post the 1st dosing of AAV). Each dot represents one animal. ns denotes not significant.

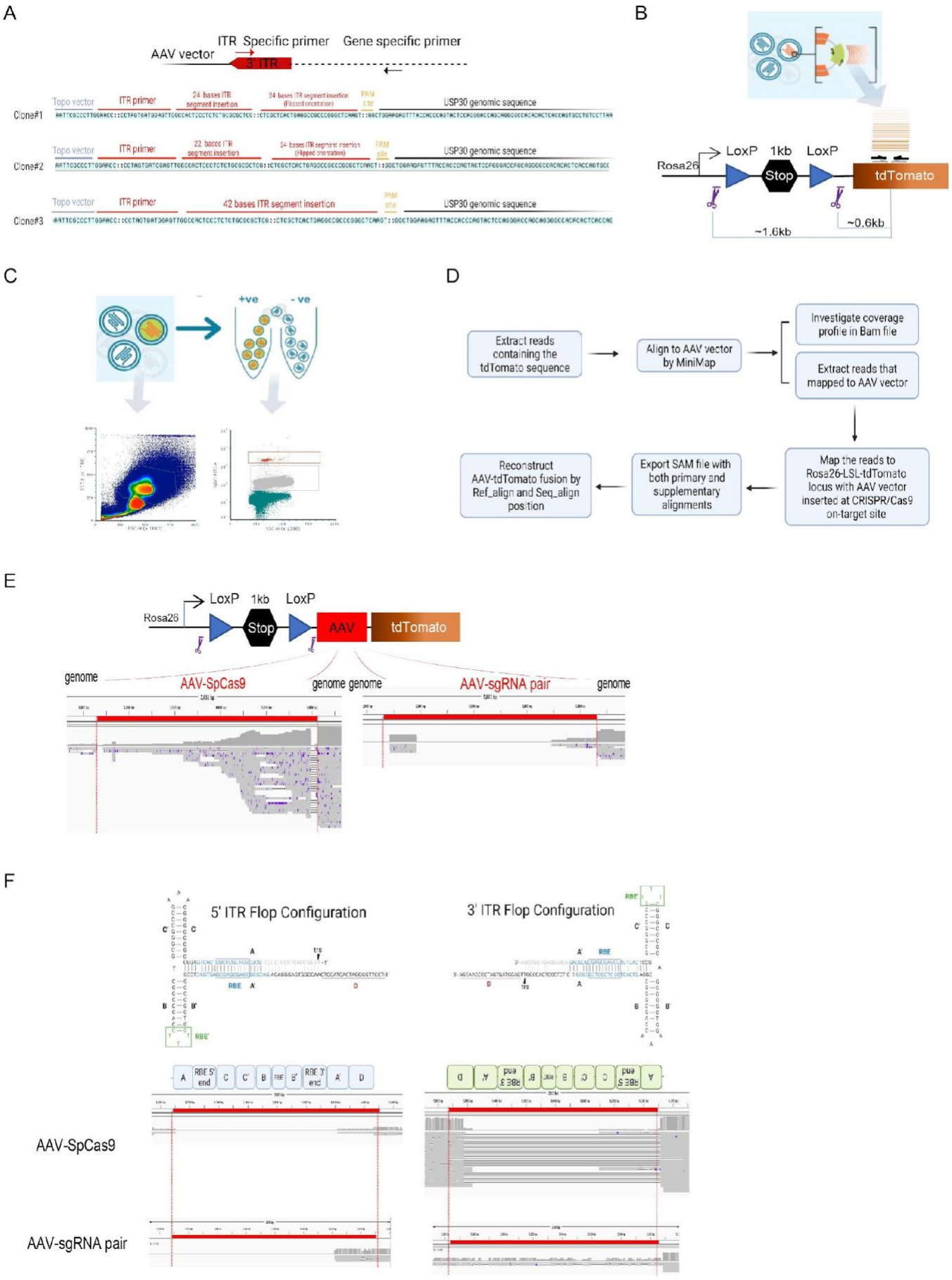

A
ITR Specific primer
Gene specific primer
AAV vector
5' ITR
Topo vector
ITR primer
24 bases ITR segment insertion
24 bases ITR segment insertion (Flipped orientation)
PAM site
USP30 genomic sequence
Clone#1
Clone#2
22 bases ITR segment insertion
24 bases ITR segment insertion (Flipped orientation)
Clone#3
42 bases ITR segment insertion
B
Rosa26
LoxP
1kb
Stop
tdTomato
~1.6kb
~0.6kb
C
+ve
- ve
D
Extract reads containing the tdTomato sequence
Align to AAV vector by MiniMap
Investigate coverage profile in Bam file
Extract reads that mapped to AAV vector
Map the reads to Rosa26-LSL-tdTomato locus with AAV vector inserted at CRISPR/Cas9 on-target site
Export SAM file with both primary and supplementary alignments
Reconstruct AAV-tdTomato fusion by Ref_align and Seq_align position
E
AAV
genome
AAV-SpCas9
genome genome
AAV-sgRNA pair
genome
F
5' ITR Flop Configuration
3' ITR Flop Configuration
RBE
AAV-SpCas9
AAV-sgRNA pair

Figure S9. Integration of AAV vector in CRISPR/SpCas9 on-target cut sites *in vitro* and *in vivo.*

A. PCR analysis of AAV vector integration *in vitro*. MLE-12-SpCas9 cells were treated with AAV-sgRNA pair targeting *Usp30* at MOI = 1e5 vg and DNA was extracted at day 6 post transduction. Representative individual AAV ITR integration clones by TOPO cloning of the amplified PCR products were shown. The integrated ITR segments were underscored by red lines. The *Usp30* genomic sequence was underscored by black lines. The CRISPR/Cas9 PAM sites were underscored by orange lines. The Topo vector sequence was underscored by blue lines. **B.** Samplix Xdrop enrichment primer design. Upper panel: Samplix workflow step 3 - Identification of region of interest by target amplicon (~150bp, placed within the region of interest); lower panel: the PCR primers design for the generation and evaluation of the 150bp target amplicon. The genomic region of interest was outlined in the diagram. The detection and evaluation primer pairs were indicated by black and gray lines, respectively. Both primer pairs bind to the tdTomato genomic locus. The scissors icons represented the CRISPR/Cas9 on-target cut sites (LoxP sgRNA pair #1). The distance between the primer pairs and the LoxP sgRNA left and right on-target sites were shown. **C.** Samplix Xdrop enrichment sorting. Upper panel: Samplix workflow step 4-5, droplet staining and sorting on flow cytometer; Lower panel: Example of the gating setup for droplets. The positive droplet population was highlighted in red box. **D.** The workflow for identification of AAV integration in the genome of interest after Xdrop enrichment and ONT long read sequencing. **E.** IGV view of the sequence alignments to the reference genomic locus with AAV vector genome inserted in the on-target site targeted by the right-sgRNA of the sgRNA pair#1. The upper track showed the genetic composition of the reference genome. The scissors icons represent the CRISPR/Cas9 on-target cut sites; Lower track showed the IGV alignment. The AAV vector was underlined by a red line in the IGV. **F.** Zoomed-in IGV view of sequences aligned to the 5' and 3'- ITR regions in the AAV vector (same reference genome as shown in E). The drawings of 5' ITR and 3' ITR in the flop configuration were shown in the top panel. The regions in the ITR sequence were annotated, including the A/A' stem, B/B' arm, C/C' arm, D sequence (underlined in red color), the Rep-binding element (RBE) (in blue color) and the second Rep-binding element (RBE') (in green color). The blue box indicated the RBE core sequence. Arrowhead, terminal resolution site (trs). Gray letters: a region absent in vector ITR. The ITR sequence was indicated as a red line in the IGV view.

Table S1. sgRNA sequences.

Target	CRISPR/Cas9 system	sgRNA name	Left sgRNA sequence (5' - 3')	Left PAM	Right sgRNA sequence (5' - 3')	Right PAM
LSL cassette	SaCas9	LoxP sgRNA pair#1	TGCTATACGAAGTTATTCGCG	ATGAAT	TACGAAGTTATATTAAGGGTT	CCGGAT
		LoxP sgRNA pair#2	TGCTATACGAAGTTATTCGCG	ATGAAT	CGACCTGCAGCCCAAGCTAGA	TCGAAT
		LoxP sgRNA pair#3	TGCTGTCTCATCATTTTGGCA	AAGAAT	TACGAAGTTATATTAAGGGTT	CCGGAT
		LoxP sgRNA pair#4	TGCTGTCTCATCATTTTGGCA	AAGAAT	CGACCTGCAGCCCAAGCTAGA	TCGAAT
	SpCas9	LoxP sgRNA pair#1	TAGGAACTTCTTAGGGCCCG	CGG	TAGCTTGGGCTGCAGGTCGA	GGG
		LoxP sgRNA pair#2	AAAGAATTGATTTGATACCG	CGG	CAAGCTAGATCGAATTCGGC	CGG
Usp30	SpCas9	sgRNA pair#14	GCTAAGTTCCCGGGCGCAGG	CGG	CCTGCCCTGCGTTTGTCAAG	TGG
		sgRNA pair#17	CCGCCCGCCGCAGCGGAATG	AGG	GAAAACGCCTACCTCTCGTG	TGG
		sgRNA pair#22	TGCGCACCGGGGCGGCCGTC	AGG	ACACACGCGAGAGCGCGCTA	AGG
Notch2	SpCas9	sgRNA pair#3	AAGATGACCCCATCGATCGA	CGG	TCACGTTATACTCGTCCAGG	AGG
		sgRNA pair#5	CAGCATCAGCTCTCGAATAG	CGG	TCGCTTATGACACACTGACG	GGG
		sgRNA pair#41	CCCAACCACAAGTGTCAGAA	TGG	CCGCCTTTCCCTCCGGACAC	AGG

Table S2. PCR primer sequences.

Sequence ID	Sequence (5'- 3')
Detection of genomic deletion in *Usp30* locus	
sgRNA pair#14_F	AGCCTTTCCACATGAACCAA
sgRNA pair#14_R	CAGACAGTAGATGAGAACCAACG
sgRNA pair#17_F	GCAAATGCGAAGGCTCATAG
sgRNA pair#17_R	TACCTGGTGTTCACAGTGTTT
Detection of genomic deletion in *Notch2* locus	
sgRNA pair#5_F	GCTTCCCTACACTCTTCACATAG
sgRNA pair#5_R	CCAGGCAAGTTCTGTCCTAAA
sgRNA pair#41_F	CCCAACCTGCCATTTCTTCT
sgRNA pair#41_R	GAGCTATCTTCTTCTAGTGGGTTTG
Detection of ITR-*Usp30* fusion events	
ITR primer_F	GGAACCCCTAGTGATGGAGTT
Usp30 primer_R	GCGTGCACACACAACAAATA
Samplix Xdrop Detection primers (generate 150bp target amplicon)	
WO100018_tdT_11F	GAAGGGCGAGATCCACCAG
WO100018_tdT_11R	TCGTTGTGGGAGGTGATGTC
Samplix Xdrop Evaluation primers	
WO100018_tdT_10F	GATCCACCAGGCCCTGAAG
WO100018_tdT_10R	CCAGCTTGGTGTCCACGTAG

Table S3. Enrichment of DNA materials via Samplix Xdrop multiple displacement amplification in droplets (dMDA).

Sample	PCR assay (detection primer pair)	Starting DNA amount (ng)	Positive droplets sorted	DNA amount post dMDA (ng)
Control	WO100018_tdT_ 11	10	331	696
AAV-SpCas9/sgRNA pair	WO100018_tdT_ 11	10	253	432

Table S4. Statistics of ONT reads.

A	B	C	D	E	F	G
Sample ID	Total reads	% Mapped reads to mouse genome (GRCm39)	Number of reads containing 30bp sequence* in tdTomato locus, adjacent to where detection primers bind	Median mapped read length from column D (bp)	Number of reads from column D mapped to AAV vector	Median mapped read length from column F (bp)
Control	305391	99%	309	4078	0	N/A
AAV-SpCas9/sgRNA pair	230788	98.05%	326	4404	39	4169

*: 30bp sequence - CTCCACCGAGCGCCTGTACCCCCGCGACGG