Supplemental Figure 1. Histology and diagnostic gene expression of HCC. (**A**) Normal liver, (**B**) advanced HCC, and (**C**) advanced HCC in three littermates treated with AAV8, age 20 months. (**D**) Microscopic view of a normal liver from untreated mouse with a wide distribution portal triads. (**E**) Liver of AAV8 treated mouse with advanced HCC showing a loss of normal hepatic architecture; portal triads are only visible in upper periphery. (**F**) Enlargement of portal triad from normal liver inset above. (**G**) Enlargement of inset form advanced HCC showing the boundary between HCC with dysplasia (lower half) and normal hepatocytes (**H**) Increased expression of *Afp* mRNA in HCC relative to expression in normal liver HCC from microarray results but not qPCR validated. *=P<0.01(ANOVA, Partek Genomics Suite 6.5) (See also Table S3).

Supplemental Table 1. Summary table of the unique integration loci with fragment read counts of ≥1 or ≥100 identified after integration capture and subsequent high throughput sequencing. (x)=AAV integrations that mapped to intergenic regions of the murine genome. The location of the AAV integrations and their relationship with annotated genes, as well as the software used to perform the analysis is included in Supplemental File 1.

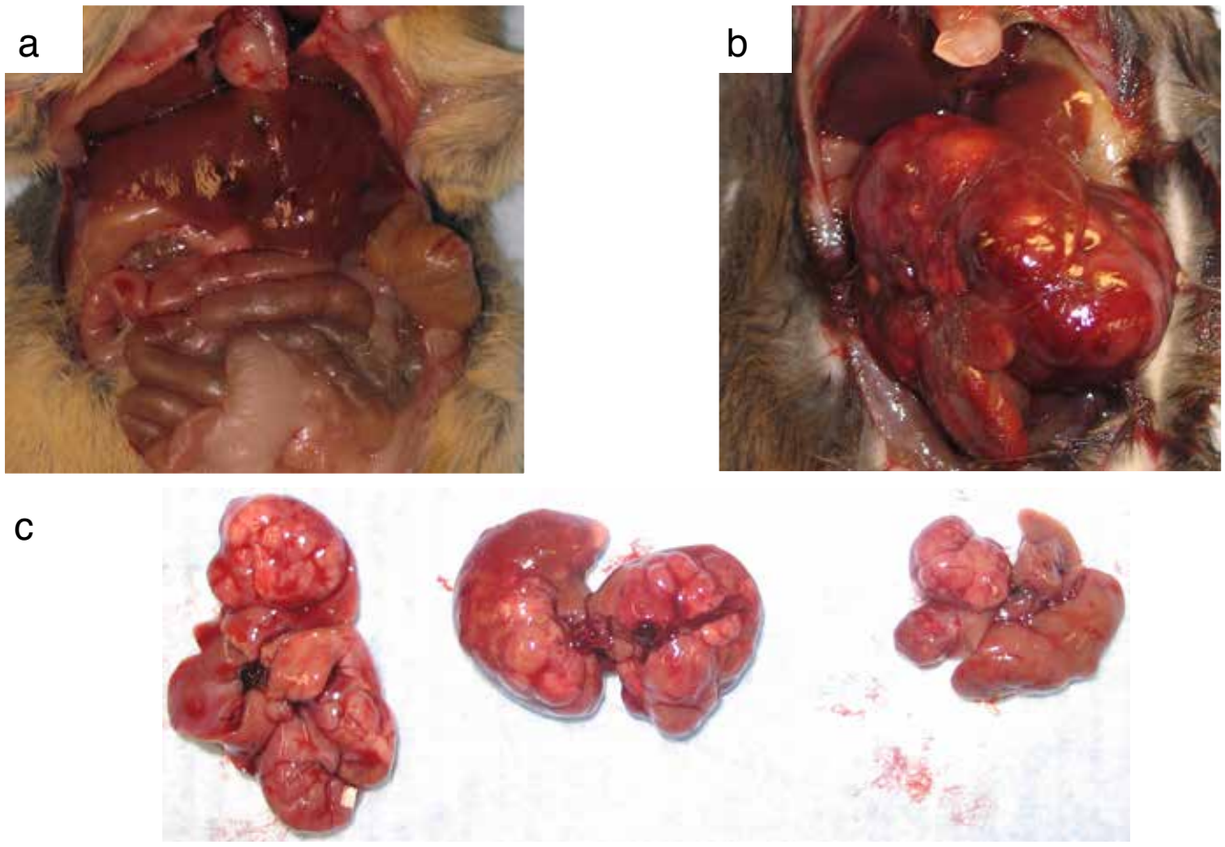
Supplemental Table 2. Genomic locations (Genome Reference Consortium Mouse Build 38) of AAV integrations in *Rian* verified by targeted PCR and sequencing.

Supplemental Table 3. RNA and miRNA specific microarray results showing differential gene expression of 2 fold or greater in HCCs versus normal liver. Microarray data have been accessioned with the Gene Expression Omnibus (GEO) under series: GSE57597 http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=eranyesanvebhsz&acc=GSE57597 and GSE61632 http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=qnkfiwiezrghdyl&acc=GSE61632

Supplemental Table 4. *Rtl1* Expression is not increased in livers treated with AAV8-hAAT-synMUT despite AAV integrations in *Rian*. Relative Quantification (RQ) of *Rtl1* normalized to *Gapdh* in normal livers treated with AAV8-CBA-MUT and AAV8-hAAT-synMUT, and HCCs treated with AAV8-CBA-MUT as determined by quantitative PCR.

Supplemental Table 5. *Mir543* Expression is not increased in livers treated with AAV8-hAAT-synMUT despite AAV integrations in *Rian*. Relative Quantification (RQ) of *Mir543* normalized to *Gapdh* in normal livers treated with AAV8-CBA-MUT and AAV8-hAAT-synMUT, and HCCs treated with AAV8-CBA-MUT as determined by quantitative PCR.

Supplemental Table 6. Unique integrations in the livers of AAV8-hAAT-synMUT treated mice (n=5) with fragment read counts of ≥1 or ≥100. (x)=AAV integrations that mapped to intergenic regions of the murine genome. The location of the AAV integrations and their relationship with annotated genes, as well as the software used to perform the analysis is available in Supplemental File 1.


a
b
c
d
e
f
g
h
mRNA Afp Expression (fold change)
4
3.5
3
2.5
2
1.5
1
0.5
0
*
Normal Liver
HCC

Sample	AAV-Vector	*Rian Chr12 location*
HCC-1	AAV8-CB7-GFP	109,609,692
HCC-2	AAV8-CB7-mMut	109,610,965
HCC-3	AAV8-CB7-mMut	109,611,522
HCC-4	AAV8-CB7-mMut	109,611,533
HCC-5	AAV8-CB7-hMUT	109,611,545
HCC-6	AAV8-CB7-hMUT	109,611,550
HCC-7	AAV8-CB7-mMut	109,611,559
HCC-8	AAV8-TBG-mMut	109,611,566
HCC-9	AAV8-CB7-mMut	109,611,567
HCC-10	AAV8-TBG-mMut	109,611,571
HCC-11	AAV8-CB7-hMUT	109,611,571
HCC-12	AAV8-CB7-GFP	109,611,571
HCC-13	AAV8-CB7-mMut	109,611,596
HCC-14	AAV8-CB7-hMUT	109,615,496
HCC-15	AAV8-CB7-hMUT	109,615,496

Supplemental Table 2. Genomic locations (Genome Reference Consortium Mouse Build 38) of AAV integrations in *Rian* verified by targeted PCR and Sanger sequencing.

Biological Group	Target Name	RQ	RQ Min	RQ Max	C_T Mean	ΔC_T Mean	ΔC_T SE	$\Delta\Delta C_T$
Normal liver treated with AAV8-CBA-MUT (n=4)	*Rtl1*	1.00	0.44	2.27	31.07	14.45	0.57	0.00
	Gapdh				16.62			
Normal Liver treated with AAV8-hAAT-synMUT (n=5)	*Rtl1*	1.26	0.71	2.22	31.11	14.12	0.40	-0.33
	Gapdh				16.99			
HCC treated with AAV8-CBA-MUT (n=6)	*Rtl1*	2,612.90	1,627.56	4,194.79	18.31	3.10	0.34	-11.35
	Gapdh				15.21			

Supplemental Table 4. *Rtl1* Expression is not increased in livers treated with AAV8-hAAT-synMUT despite AAV integrations in *Rian*. Relative Quantification (RQ) of *Rtl1* normalized to *Gapdh* in normal livers treated with AAV8-CBA-MUT and AAV8-hAAT-synMUT, and HCCs treated with AAV8-CBA-MUT as determined by quantitative PCR.

Biological Group	Target Name	RQ	RQ Min	RQ Max	C_T Mean	ΔC_T Mean	ΔC_T SE	$\Delta\Delta C_T$
Normal liver treated with AAV8-CBA-MUT (n=4)	*Mir543*	1.00	0.61	1.63	32.59	16.25	0.34	0.00
	Gapdh				16.35			
Normal Liver treated with AAV8-hAAT-synMUT (n=5)	*Mir543*	1.46	0.95	2.24	32.57	15.70	0.30	-0.55
	Gapdh				16.87			
HCC treated with AAV8-CBA-MUT (n=6)	*Mir543*	69.47	33.37	144.61	25.50	10.13	0.52	-6.12
	Gapdh				15.37			

Supplemental Table 5. *Mir543* Expression is not increased in livers treated with AAV8-hAAT-synMUT despite AAV integrations in *Rian*. Relative Quantification (RQ) of *Mir543* normalized to *Gapdh* in normal livers treated with AAV8-CBA-MUT and AAV8-hAAT-synMUT, and HCCs treated with AAV8-CBA-MUT as determined by quantitative PCR.

File Descriptions
==========

AAV_GeIST.tgz (191KB): This package provides the custom scripts used to identify the integrations (aav_geist.sh), and annotate them (aav_annotation.sh). The aav_geist.sh workflow is a modified version of the GeIST workflow used to detect MLV integrations in human cells (http://research.nhgri.nih.gov/software/GeIST/).

The labeling of the samples in the following files is based on the 6 bp barcodes used for multiplexing. We analyzed two 96-well plates of samples, and each sample well received two barcodes. The first sample plate got barcodes from barcode plates 5 and 6, and the second sample plate received barcodes from barcode plates 8 and 9.

The data below have undergone filtering. We removed 2,466 putative integrations in exons of Mut because the sequence reads contained sequence from the AAV vector genome, rather than sequence that was unique to the murine genome. In addition, the samples analyzed in this paper were sequenced with other, non-liver samples. Results from these non-liver tissues have been removed.

liver_aav_location_Chandler_et_al.bed.gz (225KB): Indicates the position of each integration detected. Note that this file contains every integration in liver, including those with low fragment count. Information on fragment count can be found in liver_aav_annotation_Chandler_et_al.txt.gz.

Integrations are organized by the plate and well designation of the barcode with which they were detected. Note that, for the reasons explained above, integrations in which both barcodes were detected appear twice. For example, the integration at position chrX:10718223-10718224 is listed with designations "plate5_E5" and "plate6_E5". These two entries represent a single event that was in well E5 of the first plate of samples, and which was detected with both barcodes assigned to that well.

Because AAV integrates in between bases, the bases indicated in this file are those immediately 3' to the integration. So, if the sequence has + orientation, the integration is on the left side of that base; if it's -, the integration is on the right.

liver_aav_annotation_Chandler_et_al.txt.gz (1.1MB): Provides information about each integration and the Ensembl transcripts (release 70) with which they intersect. Integrations that intersect multiple transcripts show up once for each transcript. The annotation software was originally designed for gene traps, so integrations are annotated with respect to the nearest downstream exon. This means that, if an integration is in the intron of one gene and the exon of another, the integration will only be annotated for the exon hit.

liver_aav_tissues_Chandler_et_al.txt.gz (4KB): This tab-delimited file indicates what type of tissue sample each well contained. All samples are from the mouse liver. "HCC" indicates a sample taken from hepatocellular carcinoma, and "Normal Liver-adj HCC" indicates a sample taken from normal liver adjacent to an HCC. The third column indicates the AAV vector used, and the fourth column contains the Mut genotype.

Five additional mice were treated with AAV8-hAAT-synMUT, and samples of their livers subjected to the same analysis as above. These data have also undergone filtering: 197 putative integrations in Hexb and Gfm2 were removed because it was determined that they arose from a vector artifact.

AAV8-hAAT-synMUT_location_Chandler_et_al.bed.gz (366KB)
AAV8-hAAT-synMUT_annotation_Chandler_et_al.txt.gz (1.9MB)
AAV8-hAAT-synMUT_tissues_Chandler_et_al.txt.gz (<1KB)