Science Advances
AAAS

Supplementary Materials for

Liver cancer heterogeneity modeled by in situ genome editing of hepatocytes

Mei Tang *et al.*

Corresponding author: Tingbo Liang, liangtingbo@zju.edu.cn; Bin Zhao, binzhao@zju.edu.cn

Sci. Adv. **8**, eabn5683 (2022)
DOI: 10.1126/sciadv.abn5683

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

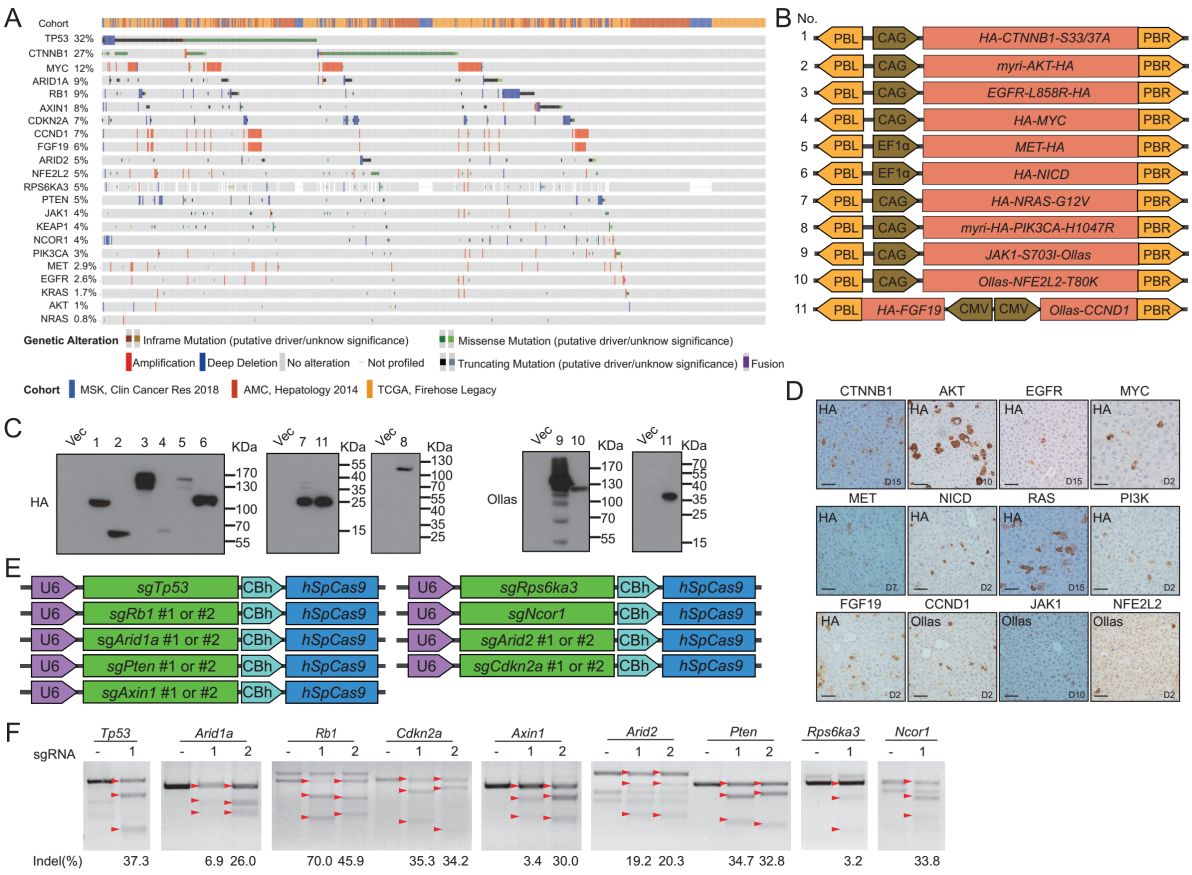

Fig. S1. Design of *in vivo* genome editing strategy for inducing mouse liver tumors. Related to Fig. 1. (**A**) Mutation frequencies of driver genes in human HCC of the MSK, AMC and TCGA cohorts. The heatmap shows genes (row) and tumors (column) with mutations. Genetic aberration types were indicated by different colors. Mutational frequency for each gene was indicated on the left. Source of each tumor was indicated on the top. (**B**) Illustration of *PB* transposon plasmids expressing transgenes. (**C**) Expression of transgenes from *PB* transposon plasmids is tissue culture. HEK293T cells were transfected and expression was examined by immunoblotting. (**D**) Expression of transgenes *in vivo*. Plasmids were hydrodynamically injected and mice were sacrificed at the indicated time. Expression of transgenes was determined by IHC of liver sections. Scale bars, 50 μm. (**E**) Illustration of pX330 plasmids expressing Cas9 and sgRNAs. (**F**) Efficiency of sgRNAs by T7E1 assay. Neuro-2a cells were transfected with sgRNAs in pEP-KO vector and selected with puromycin for 2 days. Target sequences were amplified from genomic DNA. T7E1 cleaved bands were quantified by ImageJ.

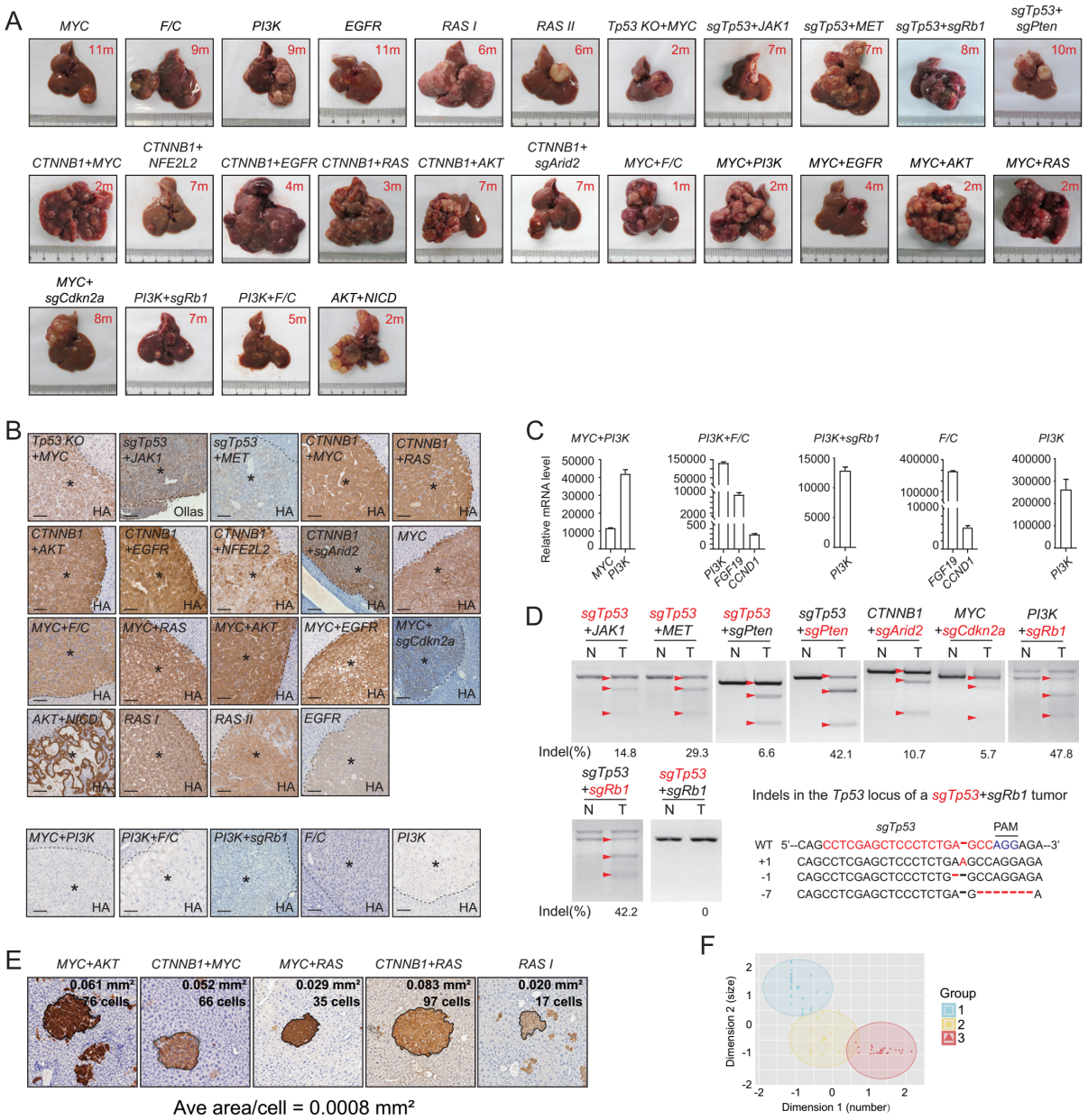

Fig. S2. Characterization of mouse liver tumors. Related to Fig. 1. (**A**) Representative mouse liver tumors in different genotypes. Time of harvest was indicated in red. (**B**) Expression of transgenes in tumors. Liver sections were stained with anti-HA or anti-Ollas tag antibodies. Scale bars, 100 μm. (**C**) mRNA expression levels of transgenes in 5 genotypes. RT-PCR was performed using primers specific to transgenes of human origin. Background level in normal liver was set as unit one. (**D**) Confirmation of target gene editing in tumors. Target sequences were amplified from tumor genomic DNA. T7E1 cleaved bands were quantified by ImageJ. Editing of the *Tp53* locus in *sgTp53+sgRb1* tumors was confirmed by sequencing of TA clones. Genes examined were in red. (**E**) Defining nodules by the area of 50 average tumor cells in fixed sections. Areas and cell numbers of 5 foci in different genotypes were quantified from HA stainings by ImageJ. (**F**) Patterns of liver tumorigenesis in mice were clustered into 3 groups by k-means analysis based on the number and size of nodules.

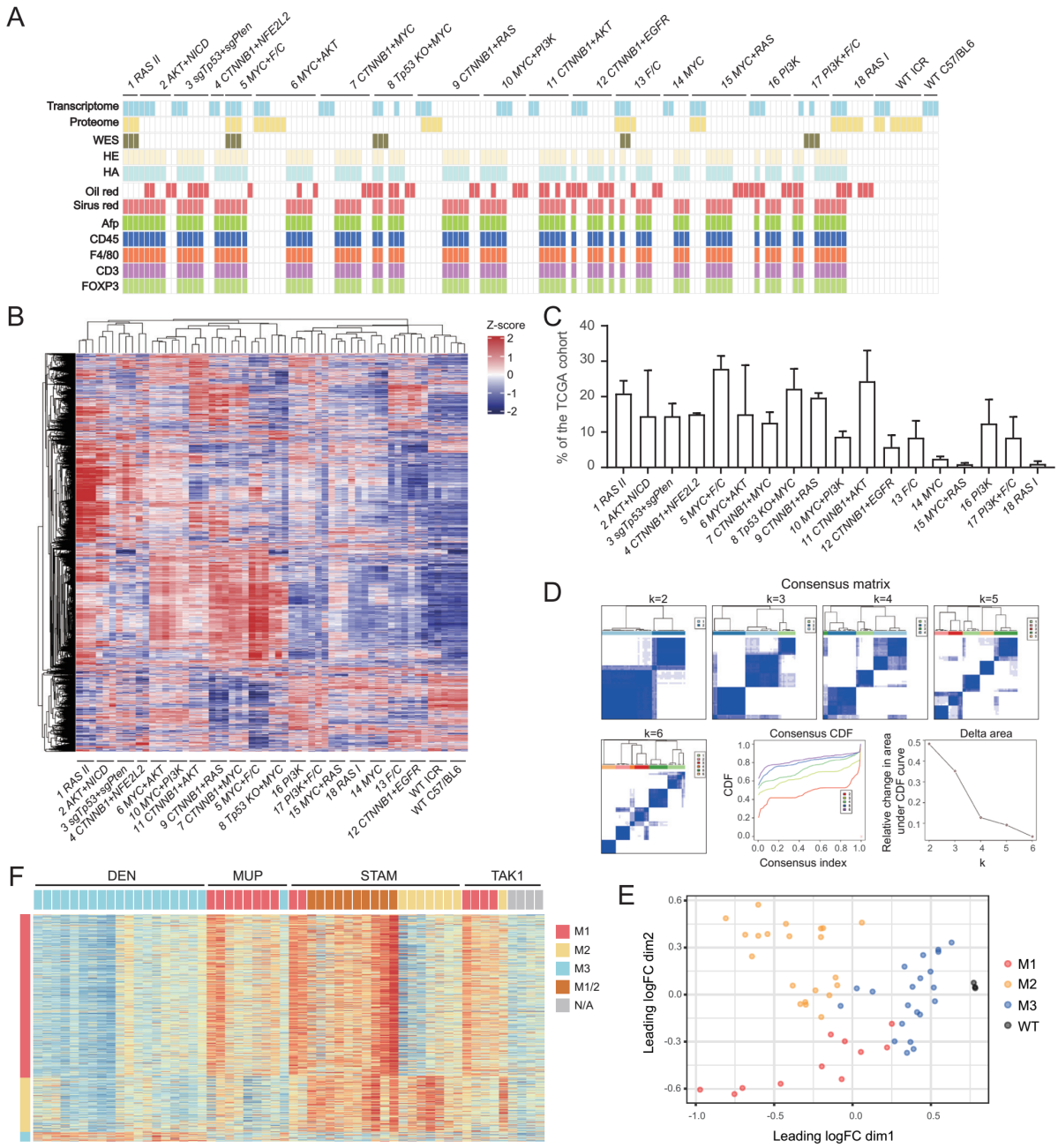

Fig. S3. Transcriptomic analysis. Related to Fig. 2. (**A**) Summary of mice used in phenotypical analysis. The heatmap shows analyses in rows and mice in columns. Multiple tumors from a mouse may be analyzed. (**B**) Unsupervised hierarchical clustering of mouse liver tumors by mRNA expression. The heatmap shows top 5,000 differential genes (row) expressed in 59 tumor and normal samples (column). (**C**) Transcriptomic similarity of human HCC to mouse liver tumor genotypes. Similarity was determined by the TROM method with $P < 0.05$. Data was presented as mean ± s.d. (**D**) Transcriptomic subtypes were identified by hierarchical consensus clustering. k was tested from 2 to 6. Consensus matrices, consensus cumulative distribution function (CDF) plot, and delta area (change in CDF area) plot were shown. (**E**) Separation of tumor subtypes and wild-type liver based on transcriptomic profiles. Analysis was done using multi-dimensional scaling. (**F**) Classification of reported mouse liver tumors using M signatures. The heatmap shows expression pattern of M signature genes (row) in tumors (column).

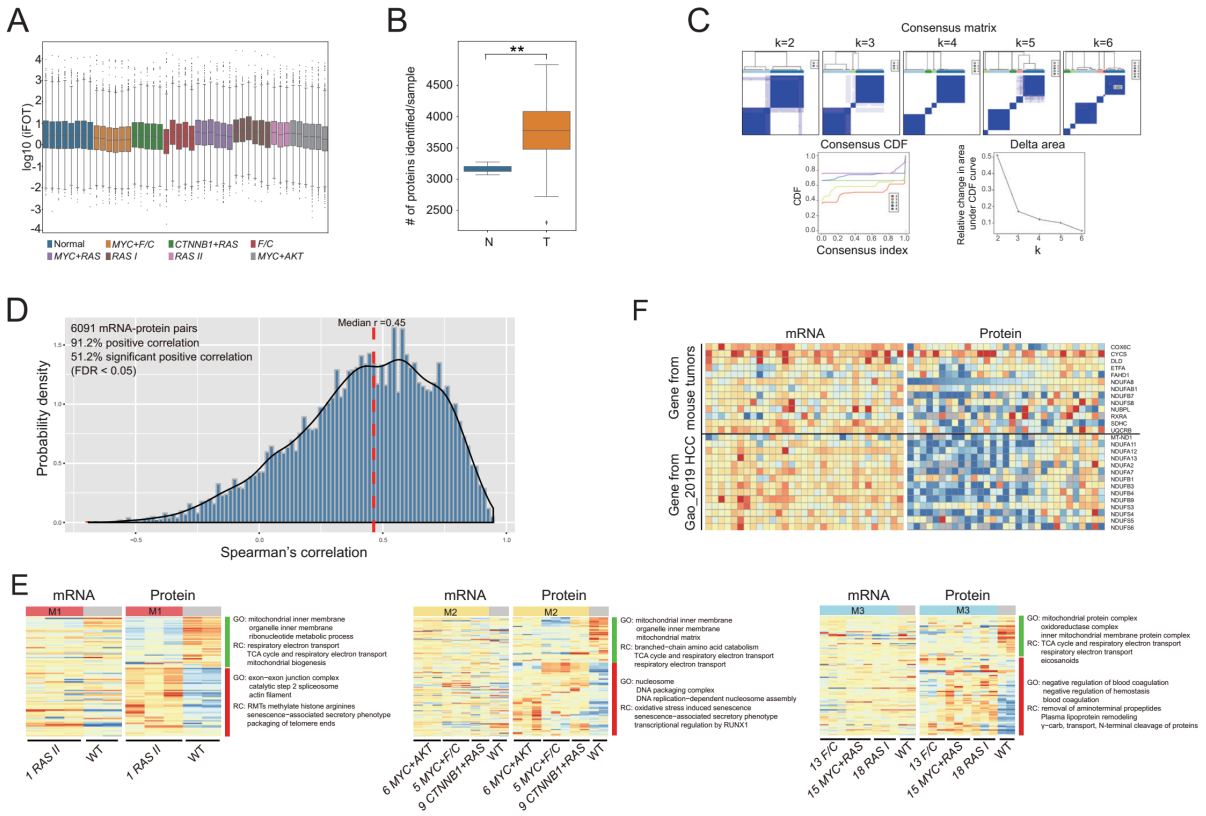

Fig. S4. Proteomic analysis. Related to Fig. 2. (**A**) Proteome detection quality by MS/MS. Distribution of log10-transformed iFOT abundance of identified proteins in 45 samples was shown in box plot. The middle line represents the median, the box represents the interquartile range, and bars extend to 1.5× the interquartile range. (**B**) More proteins were identified from tumor samples. N, normal; T, tumor. (**) P <0.01. (**C**) Proteomic subtypes were identified by hierarchical consensus clustering upon protein abundance. k was tested from 2 to 6. Consensus matrices, consensus cumulative distribution function (CDF) plot, and delta area (change in CDF area) plot were shown. (**D**) Histogram of mRNA-protein pairs. (**E**) Heatmap depicts non-correlated mRNA-protein pairs with differential expression on the protein level (FDR < 0.05 compared with normal tissue) in each subtype. Enrichment of MsigDB ontologies by protein groups were indicated. (**F**) Deregulation of mitochondrial respiratory chain proteins on protein but not mRNA levels was confirmed in human HCC. Genes from mouse tumors and from Gao_2019 cohort of human HCC were analyzed for protein abundance and mRNA expression in 31 cases of the CHNPP cohort with both RNA-seq and MS/MS data. The heatmap shows tumors in columns following the same order for mRNA (left) and protein (right) datasets.

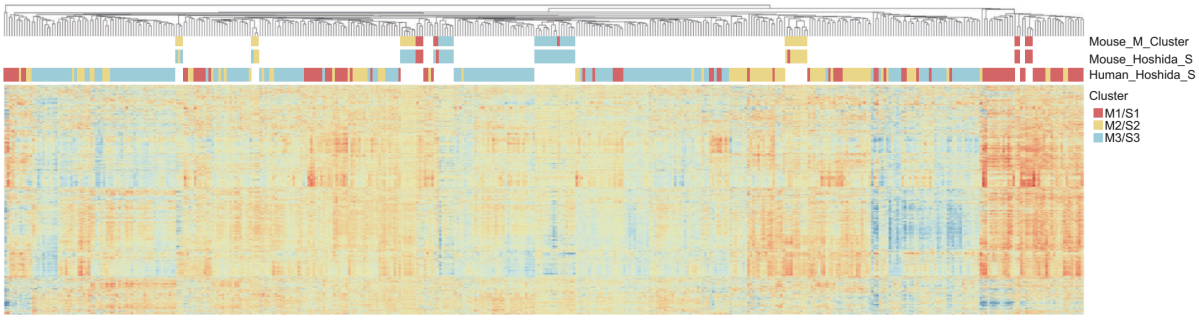

Fig. S5. Cross-species grouping of mouse and human tumors by transcriptomic profiles. Related to Fig. 2. Unsupervised hierarchical clustering was performed on mouse tumors and 374 cases of human HCC of the TCGA cohort (column) based on top 5,000 differential genes (row) of mouse tumors. Classification of tumors by Hoshida_S or mouse_M signatures were indicated on the top.

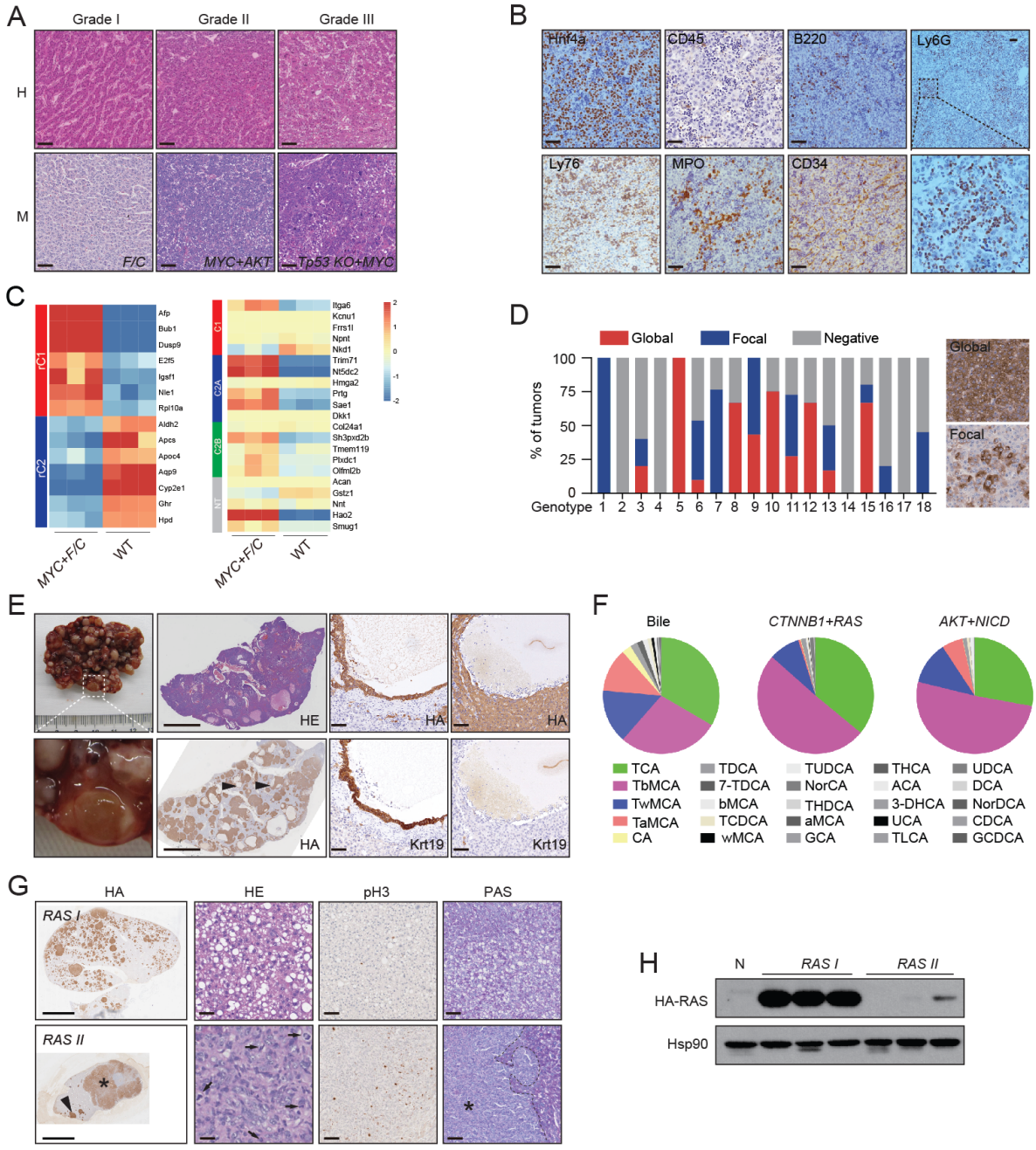

Fig. S6. Histopathological analysis of mouse liver tumors. Related to Fig. 3. (**A**) Typical histopathological features of tumors in different grades. Scale bars, 80 μm. (**B**) Extramedullary hematopoiesis in *MYC+F/C* tumors as indicated by IHC. Scale bars, 40 μm. (**C**) Molecular classification of *MYC+F/C* tumors by two hepatoblastoma signature. Genes are in rows and samples in columns. (**D**) Expression of AFP in mouse tumors quantified from IHC stainings. Two typical patterns were shown on the right. (**E**) Intratumor pseudoglands of *CTNNB1+RAS* tumors. Two areas representing pseudoglands lined by Krt19^{+} cells or not were shown. Scale bars of whole liver sections, 5 mm; other scale bars, 80 μm. (**F**) Bile acids in cystic fluid of *CTNNB1+RAS* and *AKT+NICD* tumors. Fluid was withdrawn from cysts of tumors or from the gallbladder. Bile acid compositions were determined by MS analysis. (**G**) Distinct histopathological features of *RAS I* and *RAS II* tumors. HA scale bars, 5 mm; HE scale bars, 40 μm; pH3 and PAS scale bars, 80 μm. Asterisks denote *RAS II* tumors, and arrow head indicates a *RAS I* nodule in the same liver. Arrows indicate cells during mitosis. (**H**) Expression of RAS in *RAS I* and *RAS II* tumors. Normal liver (N) or tumors were lysed and analyzed by immunoblotting using anti-HA antibody. Hsp90 was used as a loading control.

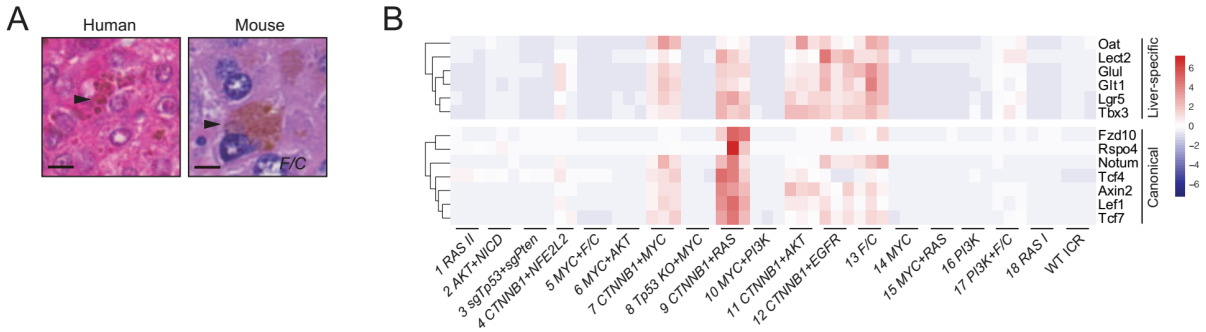

Fig. S7. *Ctnnb1*-related phenotypes in *F/C* and *PI3K*+*F/C* tumors. Related to Fig. 4. (**A**) Representative cholestasis in *F/C* tumors and in human HCC as indicated by arrowheads on HE stained liver sections. Scale bars, 20 μm. (**B**) Heatmap depicted the mRNA levels of liver-specific (upper) and canonical (lower) Wnt/β-catenin target genes. Genotypes were indicated on the bottom.

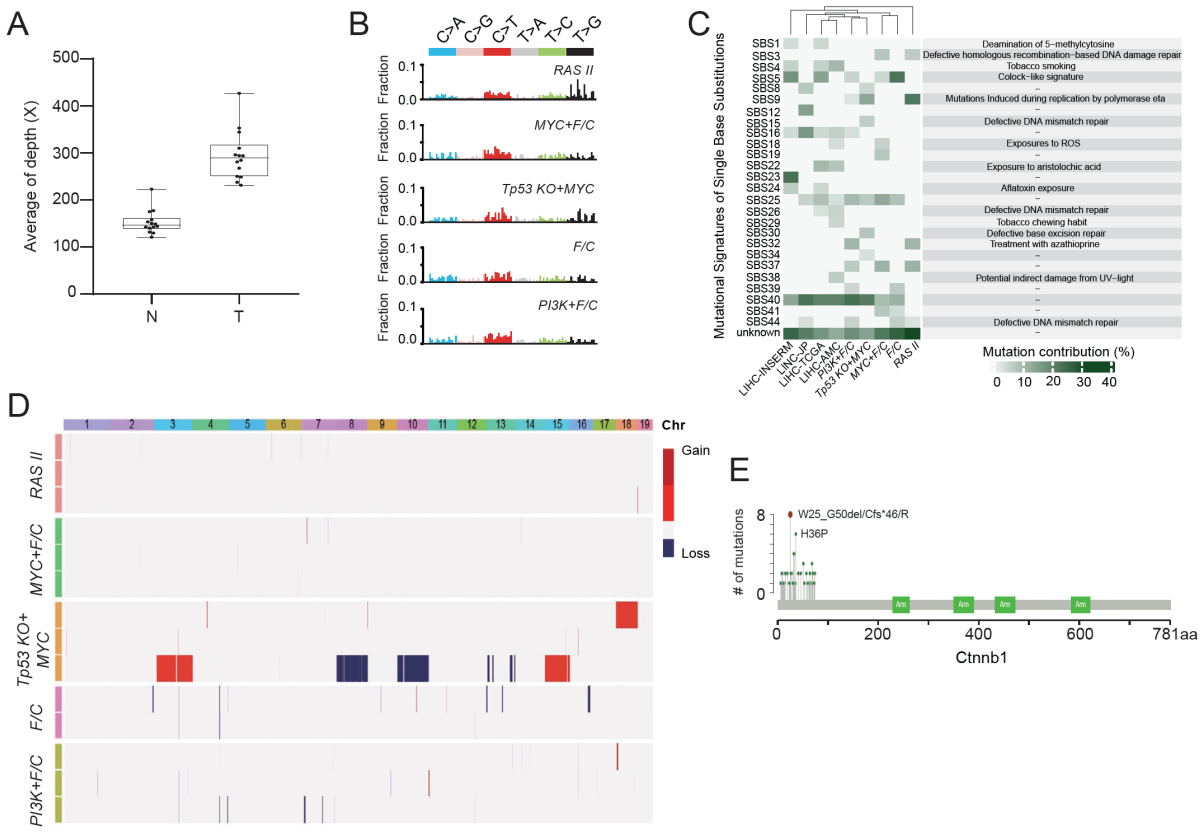

Fig. S8. Additional mutational features of mouse tumors revealed by WES. Related to Fig. 4. (**A**) Average depth of WES for normal and tumor samples. Box plots showing the median and interquartile range. (**B-C**) Nucleotide substitutions with 5' and 3' flanking bases following the standard/alphabetic order of trinucleotides (**B**) and mutation contributions for the identified signatures (**C**) in mouse liver tumors. Signatures account for >5% of mutations in at least one of the 9 groups were displayed. (**D**) Frequency of CNVs along the genome. Copy number loss (blue) and copy number gain (red) regions were highlighted. Genotypes were indicated on the left and chromosomes were indicated on the top. (**E**) Nonsynonymous mutations of Ctnnb1 found in mouse liver tumors.

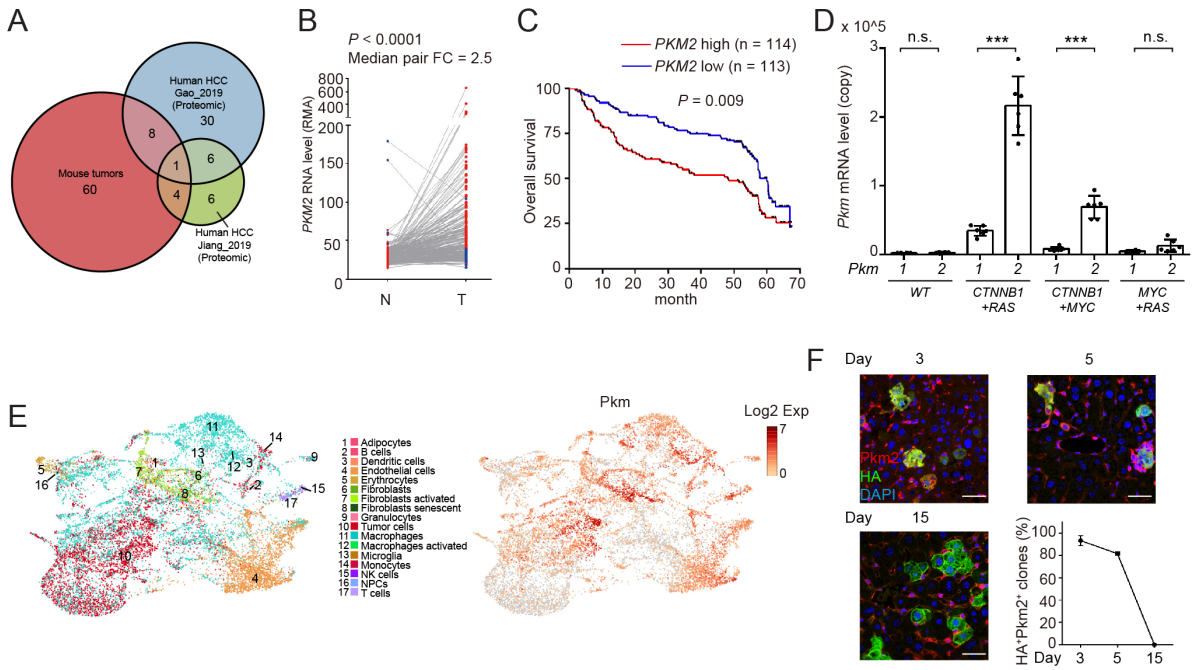

Fig. S9. Dysregulation of *Pkm2* expression in cancer. Related to Fig. 5. (**A**) Venn diagram of drug targets deregulated in mouse liver tumors and in human HCC. (**B**) Aberrant expression of *PKM2* is human HCC. N, non-tumor; T, tumor. 232 pairs of the LCI cohort (GSE14520) were analyzed. (**C**) High expression of *PKM2* predicts worse survival. Kaplan-Meier plot of human HCC of the LCI cohort (GSE14520) with median separation by expression of *PKM2*. (**D**) Expression of *Pkm* isoforms in mouse liver tumors by RT-PCR-based absolute quantification of mRNA. Data was presented as mean ± s.d. of six biological repeats. (**E**) The UMAP analysis of all single cells color-coded for their assigned cell type (left) and for the expression of *Pkm* (right). (**F**) Dynamic expression of Pkm2 during *RAS*-induced tumorigenesis. Expression of Pkm2 in lesions 3-15 days after injection was visualized by multiplex IHC. Scale bars, 50 μm. Portions of $Pkm2^+$ lesions were quantified from 3 mice at each time point. Data was presented as mean ± s.d. n.s., not significant; (***) P <0.001.

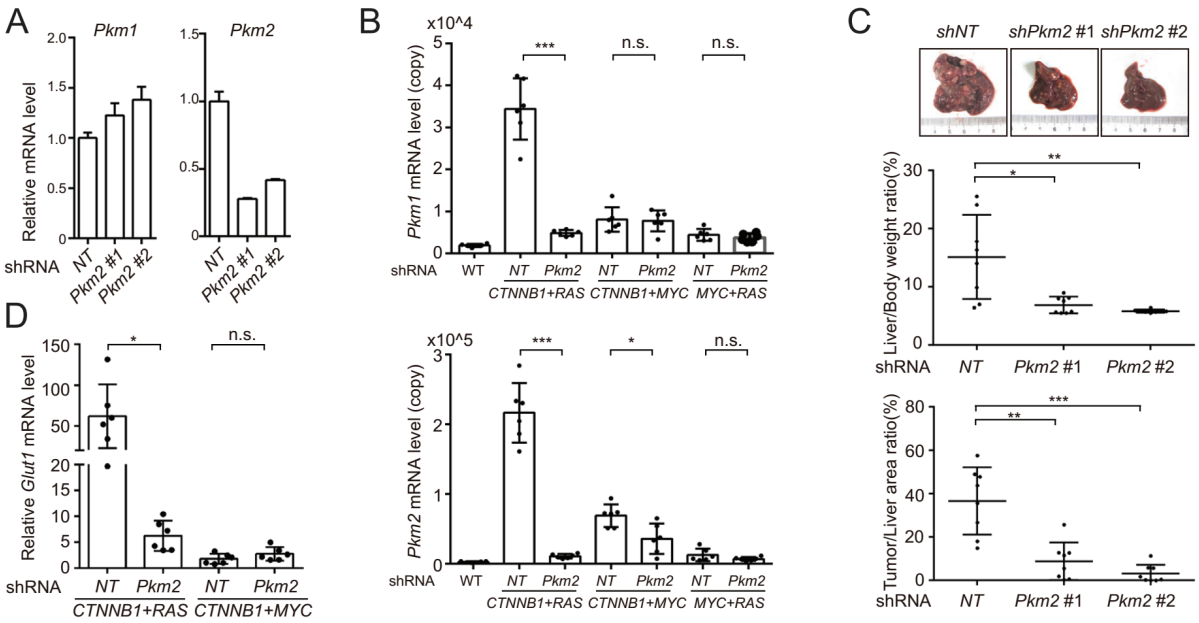

Fig. S10. Effect of *Pkm2* knockdown in tumorigenesis. Related to Fig. 6. (**A**) Efficiency of *Pkm2* shRNAs in H2.35 cells as determined by RT-PCR. Data was presented as mean ± s.d. of two technical repeats. (**B**) Effect of *Pkm2* knockdown on expression levels of *Pkm* isoforms in tumors by RT-PCR-based absolute quantification of mRNA. Data was presented as mean ± s.d. of six biological repeats. (**C**) Knockdown of *Pkm2* by individual shRNAs suppressed tumorigenesis induced by *CTNNB1+RAS*. Experiments were similar to that in (Fig. 6B-D). (**D**) Knockdown of *Pkm2* normalized *Glut1* expression in *CTNNB1+RAS* tumors as determined by RT-PCR. Fold change to normal liver was presented as mean ± s.d. of six biological repeats. *P* values were calculated by Student's t-test. n.s., not significant; (*) $P < 0.05$; (**) $P < 0.01$; (***) $P < 0.001$.

Supplementary Tables

Table S1. Driver gene mutation frequencies in 7 datasets. Related to Figure 1. (See Excel file)

Table S2. Summary of liver tumorigenesis induced by *in situ* genome editing. Related to Figure 1. (See Excel file)

Table S3. Transcriptomic profile of mouse liver tumors. Related to Figure 2. (See Excel file)

Table S4. Proteomic profile of mouse liver tumors. Related to Figure 2. (See Excel file)

Table S5. Histopathological profile of mouse liver tumors. Related to Figure 2 and 3. (See Excel file)

Table S6. Mutational profile of mouse liver tumors. Related to Figure 4. (See Excel file)

Table S7. Dysregulation of drug targets in mouse liver tumors. Related to Figure 5. (See Excel file)

Table S8. Detailed information for antibodies and oligonucleotides used in the study. (See Excel file)