Metformin decreases *Cyp26a1* to prevent hepatocarcinogenesis through down-regulating $CD8^+$ T cells

Weizhi He[1, 2, *], Miaomiao Chen[1, 2, *], Chong Li[3, *], Xicheng Wang[1, 2, *], Wenjian Chen[1, 2], Lili Pan[1, 2], Yangyang Cui[1, 4], Zhao Yu[1, 2], Guoxiu Wu[1, 2], Yang Yang[1, 2], Qinghe Tang[3], Jinghan Wang[3], Zhiying He[1, 2, #]

[1]Institute for Regenerative Medicine, Shanghai East Hospital, School of Life Sciences and Technology, Tongji University School of Medicine, Shanghai 200123, P. R. China.

[2]Shanghai Engineering Research Center of Stem Cells Translational Medicine, Shanghai 200335, P. R. China.

[3]Department of Hepatobiliary and Pancreatic Surgery, Shanghai East Hospital, Tongji University, Shanghai, China.

[4]Postgraduate Training Base of Shanghai East Hospital, Jinzhou Medical University, Jinzhou, Liaoning 121001, China.

*These authors have contributed equally to this work and shared first authorship. Author order was determined on the basis of seniority.

 # Correspondence:

Zhiying He, PhD, Professor, Executive Director, Institute for Regenerative Medicine, Shanghai East Hospital, Tongji University, 1800 Yuntai Rd., Shanghai 200123, P.R. China. Tel: +86-21-20334518-22241 Cel: +86-18049968868

Mail: zyhe@tongji.edu.cn

Abstract

Hepatocellular carcinoma (HCC) is a highly heterogeneous cancer, which limits the selectivity of prevention and treatment. Preclinical and clinical studies suggested that in patients with diabetes, prolonged use of metformin, the AMPK activator, was associated with a reduction of HCC incidence. This association promotes us to investigate the possible functions and mechanisms of metformin in HCC without diabetes backgrounds. Here, we found that several unique pathways that changed during chronic liver injury of *Fah*$^{-/-}$ mice, including glucose metabolic process and retinol metabolism. Further, metformin suppressed the tumor formation in chronic liver injury of *Fah*$^{-/-}$ mice. RNA sequencing, *in vivo* and *in vitro* experiments showed that metformin suppressed *Cyp26a1* gene expression of hepatocyte. Moreover, the down-regulation of Cyp26a1 leads to the increased level of all-trans-retinoic acid (atRA), which could suppress the tumor formation in our model. On the other hand, flow multicolor analysis showed that the cell number and proportion of cancer promoting (pro-tumor) CD8^{+} T cells increased significantly during chronic liver injury in *Fah*$^{-/-}$ mice, and both metformin and atRA treatment could reduce the number and proportion of pro-tumor CD8^{+} T cells. We also found metformin decreased the *Cyp26a1* expression through the AMPK/JNK/c-Jun pathway. In short, the association between the metformin and atRA may explain the commonness of their anti-tumor activities. Our findings highlight the importance of targeting the precancerous microenvironment for the prognosis, prevention and treatment of HCC.

Keywords: atRA; *Cyp26a1*; $CD8^+$ T cell; Metformin; Hepatocellular Carcinoma

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common neoplasm and the third leading cause of cancer death in the world (1). Although HCC treatment has made great advances over the past decades, tumor heterogeneity of the HCC still keeps its mortality rate high. HCC is a complex malignancy that can be triggered by various factors, including hepatic viral infections, environmental exposure to toxic substance (such as aflatoxin B1), alcohol abuse and non-alcoholic fatty liver disease (NAFLD) (2). Besides, growing evidence supports an association between HCC and metabolic syndrome, diabetes or obesity (3). In addition to these acquired HCC, patients with certain gene defects are bound to develop HCC (4-6). The heterogeneity of HCC, especially the diverseness in causes and progressions, complicates its therapeutic approach.

Metformin (Met), a synthetic analog of guanidine, is the most commonly prescribed drug in treating the type 2 diabetes mellitus (T2DM) (7). In addition to lowering glucose, metformin directly inhibits complex I (CI) of the electron transport chain, resulting in decreased complex I activity and oxidative phosphorylation (OXPHOS) level in cells (8-10). Consequently, the elevated AMP/ATP ratio activates adenosine monophosphate–activated protein kinase (AMPK) signaling pathway, which promotes cell cycle arrest and inhibits tumor cell proliferation (11). Recently, it has been demonstrated that clinical doses of metformin-bound PEN2 forms a complex with ATP6AP1, which leads to the inhibition of v-ATPase and the activation of

AMPK without effects on cellular AMP levels (12). Preclinical studies and retrospective population-based studies have suggested antitumor activity of metformin alone or in combination with conventional anticancer drugs (13-16). Preclinical and clinical studies also have suggested that in patients with diabetes, prolonged use of metformin could be associated with a reduction of the HCC incidence (17, 18). Several research groups also have demonstrated metformin inhibited non-diabetes induced HCC in a mouse HCC model (19, 20). Williams *et al.* showed metformin prevents HCC development via alleviating p21 overexpression and ameliorating pro-tumorigenic microenvironment (19). Shankaraiah *et al.* showed metformin prevents hepatocarcinogenesis by attenuating fibrosis in a transgenic mouse model of HCC (20). However, despite decades of research, the mechanism by which metformin inhibits HCC is still highly debated. Additionally, the heterogeneity of HCC also impels us to determine the function and mechanism of metformin in chronic liver injury models of HCC.

Due to a deficiency in fumarylacetoacetate hydrolase (FAH) (21), hereditary tyrosinemia type 1 (HT1) is the severe inherited disorder of the tyrosine degradation. The FAH deficiency leads to the accumulation of toxic metabolites, mainly in the liver (22). Untreated HT1 patients usually die before the age of two. HT1 patients suffer multiple chronic complications, including cirrhosis and a high risk of HCC (22). NTBC/Nitisinone (a drug blocking the pathway upstream of FAH) greatly delay the morbidity of HT1 patients, but patients receiving NTBC may still developing

HCC (23). The molecular basis of the pathogenic process in HT1 is still unclear. The murine model of *Fah*-deficiency (*Fah*$^{-/-}$) is a suitable animal model, featuring all phenotypic and biochemical expressions of the HT1 patients (24). Moreover, the chronic liver injury model of *Fah*$^{-/-}$ mice closely resembles human alcohol-induced and c-myc-altered HCC (25). Therefore, the study of FAH deficient chronically injured mice can serve as a good indicator of human HCC.

In this study, we demonstrate that several unique signal pathways are vastly changed in an HCC model induced by chronic liver injury. These pathways are characterized by the regulation of glucose metabolic process and retinol metabolism. We find that the metformin increased the AMPK activity and subsequently reduced the HCC incidence. Our data also shows the metformin suppressed the *Cyp26a1* gene expression, leading to the increase of the all-trans-retinoicacid (atRA) level, and then reduced the number and ratio of pro-tumor CD8$^+$ T cells. Moreover, metformin could suppress *Cyp26a1* gene expression through the AMPK/JNK/c-Jun pathway. Our results first disclose a link between the metformin and atRA, which may explain the commonness of their anti-tumor activities through inhibiting pro-tumor CD8$^+$ T cells.

Materials and methods

Patient Specimens

Patient specimens were obtained from sample information service platform of Shanghai East Hospital under protocols approved by Shanghai East Hospital. Final

diagnoses of the specimens were confirmed by experienced pathologists from Shanghai East Hospital. The study protocol conformed to the ethical guidelines approved by the Shanghai East Hospital Ethics Committee, and written informed consent was obtained from each patient.

Analysis of public clinical datasets

The gene expression data of 7858 samples of 30 normal tissues (including liver organ) were downloaded from GTEx (https://gtexportal.org/home/), and the gene expression data of 710 samples of 17 para-tumor tissues and 7801 samples of 34 cancer types (including HCC) were downloaded from TCGA (https://portal.gdc.cancer.gov/). Correlation analyses between genes and signatures were calculated "cor()" function in R and also conducted in GEPIA database (http://gepia.cancer-pku.cn).

Moreover, we obtained mRNA expression from chronic injury liver cases from GSE89632 and GSE148355. To understand whether there is a similar trend that AMPK signaling pathway is inhibited in human chronic liver injury as in mouse model, we downloaded RNA-seq dataset GSE148355 (26) and Microarray dataset GSE89632 (27) from the public database GEO for verification. GSE148355 dataset was composed of HCC samples, "premalignant" samples with different degrees of fibrosis, and non-tumor samples from patients who had undergone surgical resection. We mainly used the normalized data (FPKM) of 62 premalignant and non-tumor samples for analysis, including 15 non-tumor samples (Normal), 10 low fibrosis

samples (Fibrosis low, FL), 10 high fibrosis samples (Fibrosis high, FH), 10 cirrhosis samples (Cirrhosis, CS), 10 low degree of dysplasia nodule samples (Dysplastic nodule low, DL) and 7 high degree of dysplasia nodule samples (Dysplastic nodule high, DH). Likewise, we also analyzed microarray data from 63 NAFLD and healthy samples in GSE89632 dataset, including 20 simple steatosis samples (SS), 19 nonalcoholic steatohepatitis samples (NASH), and 24 healthy controls (HC).

In this study, we used the key genes in AMPK pathway to form a self-defined functional gene set, which was composed of PFKFB3, CPT1A, SIRT1, PPARGC1A, SLC2A4, PNPLA2, CRY1 and FOXO1. To infer whether AMPK pathway is inhibited during chronic liver injury, single-sample gene set enrichment analysis (ssGSEA) algorithm was used to calculate the enrichment score of this gene set in normal samples and samples with different degrees of chronic liver injury (28). Meanwhile, the expression of these eight single genes were also statistically analyzed in both healthy and chronic liver injury samples.

Mouse model and Treatments

As described in previous studies (29, 30), 129S4 *Fah*$^{-/-}$ mice undergo liver failure and death. All mice were bred and maintained under specific-pathogen free conditions at Shanghai East Hospital. Mice were housed with a light: dark cycle of 12h, ambient temperature of 24℃ and humidity of 55%. Female mice were purely randomly allocated to metformin/atRA treated or untreated group. Mice in metformin experiments received metformin (B1970, APExBIO) via intraperitoneal injection at a

dosage of approximately 250 mg/kg/day from 0 weeks until 12 weeks of chronic liver injury. 200 μg atRA (Sigma-Aldrich) dissolved in DMSO or DMSO alone was administered by intraperitoneal injection every other day from 0 weeks until 12 weeks of chronic liver injury. All experimental procedures on mice were approved by Institutional Animal Care and Use Committee of Shanghai East Hospital.

Real time quantitative PCR (qRT-PCR)

Quantitative PCR was performed as described previously (29). Briefly, total RNA was isolated from liver of mice using RNAiso plus (TaKaRa). RNA was reverse-transcribed using the PrimeScript RT reagent Kit (TaKaRa). Transcript expressions were determined by qRT-PCR using SYBR Premix Ex Taq II (TaKaRa) and the QuantStudio™ real-time PCR instrument (Applied Biosystems). The primers sequences used were as followed:

Cyp26a1: 5'-AAGCTCTGGGACCTGTACTGT-3' and 5'-CTCCGCTGAAGCACCATCT-3', Gapdh: 5'-AGGTCGGTGTGAACGGATTTG-3' and 5'-TGTAGACCATGTAGTTGAGGTCA-3', Il6: 5'-GTCCTTCCTACCCCAATTTCC-3' and 5'-TAACGCACTAGGTTTGCCGA-3', Il1β: 5'-GCAACTGTTCCTGAACTCAACT-3' and 5'-ATCTTTTGGGGTCCGTCAACT-3', Tnf: 5'-CCCTCACACTCAGATCATCTTCT-3' and 5'-GCTACGACGTGGGCTACAG-3', Ltβ: 5'-TGGCAGGAGCTACTTCCCT-3' and 5'-TCCAGTCTTTTCTGAGCCTGT-3'.

qRT-PCR calculations were performed using the comparative CT method. The data

172 are represented either as fold changes of the treated group compared to the control
173 group or as $2^{-\Delta\Delta CT}$ mRNA transcript abundance.

174 **RNA-Seq and Transcriptomic Analyses of Mouse Liver Tissues**

175 Total RNA (two liver tissues of C0 and C4, three liver tissues of C12 and Met12)
176 were extracted using the mirVana miRNA Isolation Kit (Ambion) following the
177 manufacturer's protocol. RNA integrity was evaluated using the Agilent 2100
178 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The samples with RNA
179 Integrity Number (RIN) ≥ 7 were subjected to the subsequent analysis. The libraries
180 were constructed using TruSeq Stranded mRNA LTSample Prep Kit (Illumina, San
181 Diego, CA, USA) according to the manufacturer's instructions. Then these libraries
182 were sequenced on the Illumina sequencing platform (HiSeqTM 2500 or Illumina
183 HiSeq X Ten) and 125 bp/150 bp paired-end reads were generated.

184 The transcriptome sequencing and analysis were conducted by OE biotech Co.,
185 Ltd. (Shanghai, China). Raw data (raw reads) were processed using Trimmomatic
186 (31). The reads containing ploy-N and the low-quality reads were removed to obtain
187 the clean reads. Then the clean reads were mapped to reference genome using hisat2
188 (32). FPKM value of each gene was calculated using cufflinks, and the read counts of
189 each gene were obtained by htseq-count (33). DEGs were identified using the DESeq
190 (2012) R package functions estimate Size Factors and nbinomTest. "P value < 0.05"
191 and "Fold Change > 2 or Fold Change < 0.5" was set as the threshold for significantly
192 differential expression. Hierarchical cluster analysis of DEGs was performed to

193 explore genes expression pattern. Gene Ontology (GO) enrichment and KEGG
194 pathway enrichment analysis of DEGs were respectively performed using DAVID
195 database(https://david.ncifcrf.gov/).

196 **Protein isolation and Western blot**

197 Livers were harvested from normal *Fah*$^{-/-}$ mice (100% NTBC), *Fah*$^{-/-}$ mice with
198 chronic liver injury (2.5% NTBC), or from different stages during the development of
199 HCC. Lysates were prepared from the liver tissues using RIPA lysis buffer. The
200 following primary antibodies were used for immunoblotting: anti-GAPDH (Protein
201 tech, HRP-60004), anti-AMPK (CST, 2532), anti-P-AMPK (CST, 2535), anti-
202 CYP26A1 (Santa Cruze, sc-53618), anti-P-JNK (CST, 9255), anti-JNK (CST, 9252),
203 anti-P- c-Jun (Abcam, ab40476), anti-c-Jun (Abcam, ab32385). Primary and
204 secondary HRP-labeled antibodies were used at 1:500 and 1:2000 dilutions
205 respectively. Detection was performed with SuperSignal West Femto Maximum
206 Sensitivity Substrate (Thermo Scientific).

207 **Metformin treatment of primary hepatocyte**

208 Liver of chronic injury of *Fah*$^{-/-}$ mice were perfused as before (34). 5×10^5 cells
209 of the primary hepatocyte were seeded in a 6-well plate (matrigel coated) with 2mL
210 Advanced DMEM/F-12 (Thermo Fisher, 12634010) supplemented with 10% fetal
211 calf serum. 24 hours later, the fresh medium replaces the old medium. The cells were
212 treated with different concentrations of metformin or AICAR for 6 hours. In the

213 presence of inhibitor, inhibitor was pre-treated to hepatocyte for 1 hour following the
214 adding of metformin.

215 **Luciferase assay**

216 Chronic injured liver of *Fah*$^{-/-}$ mice were perfused. About 2x10^4 cells of the
217 primary hepatocyte were seeded in a 24-well plate with 1mL Advanced DMEM/F-12
218 supplemented with 10% fetal calf serum. 2 hours later, 1 mL Advanced DMEM/F-12
219 without fetal calf serum replaces the old medium. The second day, cells were
220 transfected with a 6:1 ratio of a firefly luciferase reporter plasmid driven by a pGL3-
221 RARE-responsive promoter (Addgene) and a Renilla luciferase reporter plasmid
222 driven by a constitutive CMV promoter (Promega). 24 hours later, the culture
223 medium was replaced with fresh Advanced DMEM/F-12 medium and
224 metformin/atRA or vehicle. 24 hours later, activity of both reporters was measured
225 using the Dual-Luciferase Reporter kit (Promega) and read on a Tecan Infinite 200
226 PRO Reader. The firely luciferase to renilla luciferase ratio is reported as
227 “Firely/Renilla Luciferase Activity”.

228 **Serological Analyses**

229 For Serum indicators, blood sample was collected from the retro-orbital sinus of
230 test animals. Plasma was prepared using Microtainer plasma separator tubes (BD) and
231 stored at −80 °C. Analysis the biochemical indicators of serum was according to the
232 previously established protocol (29).

233 **Hematoxylin-Eosin staining (H&E), Immunohistochemistry (IHC)**

Hematoxylin-Eosin staining (H&E) and Immunohistochemistry (IHC) are showed as before (29). For hematoxylin-eosin staining, fresh liver tissues were fixed with 4% paraformaldehyde (PFA), then were routinely embedded in paraffin and sectioned into slices (2 μm). The slices were processed for roasting, dewaxing and rehydration, and were stained with hematoxylin (Beyotime, Shanghai, China) for 5–10 min, rinsed with water for 15 min, sliced into 95% alcohol (Sinoreagent, Shanghai, China) for 30 seconds, and then stained with eosin (Beyotime, Shanghai, China) for an appropriate amount of time (0.5–2 min). Finally, the slice is dehydrated rapidly and mounted with neutral resin (MXB, Fujian, China).

For immunohistochemistry staining, IHC steps are the same as H&E before rehydration. The slices were soaked in 0.01 M citric acid buffer (pH 6.0) or EDTA (pH 9.0) and placed in a pressure cooker for 2–4 min at 121 °C/100 kpa. Cool to room temperature. 3% H_2O_2solution blocked endogenous peroxidase and 1% BSA blocked nonspecific loci for 30 min at room temperature. The slices were incubated with primary antibodies for at 4 °C overnight and secondary antibodies conjugated with HRP at 37 °C for 30 min. Staining with DAB (Vector Laboratories, Burlingame, CA) was applied to the sections. The sections were stained with hematoxylin (Beyotime, Shanghai, CHN), dehydrated rapidly and mounted with neutral resin (MXB, Fujian, CHN). The used primary antibodies are summarized as below: CD8 (Abcam, ab209775), CD45 (CST, 70257).

Flow Cytometry

Fresh liver tissue samples were minced, then digested in mouse Liver Dissociation Kit (MACS 130-105-807) for 0.5h at 37 ℃. Cell suspensions were generated through a 70μm nylon mesh. After filtration and washing, cells were suspended in a 40% Percoll solution following by centrifugation at 600 g for 20 min. Cells were prepared as a single-cell suspension for FACS staining. For surface staining, the following antibodies were used: APC-CY7-LIVE/DEAD, FITC-CD45, V450-CD3, AF700-CD4, V500-CD8, BV605-CD49B, PE-F4/80, BB700-CD11B. The stained cells were acquired for analysis using the Cytoflex (BECHMAN). Flow cytometry data was analyzed with FlowJo software (Tree Star Inc.).

Blood Glucose Measurement

Blood was sampled in mice by nicking the tail vein and blood glucose levels were measured using ACCU-CHEK Active test strips read by an ACCU-CHEK Active meter (Roche Diagnostics, Indianapolis, IN) following the manufacturer's instructions.

Single cell RNA sequencing (scRNA-seq) analysis

The scRNA-seq data of *Fah*$^{-/-}$ mouse at 0, 12 and 18 weeks after the withdrawal of NTBC were downloaded from GEO databases with the GEO accession number GSE130880 (35). The dimension reduction analysis of Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP) were calculated by "RunTSNE()" and "RunUMAP()" function of Seurat package (R-

version 4.0.5). Correlation analyses between signatures were calculated "cor()" function in R. In addition, we used ssGSEA algorithm (28) to calculate the scores of pro-tumor $CD8^+$ T cell signature in CLI *Fah*$^{-/-}$ mouse at 0, 12 and 18 weeks.

Statistics

All results were presented as "mean ± SEM" as indicated. The differences between groups were analyzed with one-way or two-way ANOVA, Student's t-test or Mann-Whitney U test with two-tailed. The overall survival (OS) and disease-free survival (DFS) were analysis by Log-Rank test using GraphPad Prism 7. ** $P < 0.01$, * $P < 0.05$, NS = not significant.

Data Availability

The RNA sequencing raw data for this study were generated at OE biotech Co., Ltd. All data generated or analysed during this study are included in the manuscript and supporting file; Source Data files have been provided for Figures 1 and 3.

Results

Characters of *Fah*$^{-/-}$ mice with chronic Liver injury

Using the *Fah*$^{-/-}$ mouse model as a surrogate for chronic liver disease, we first described the liver phenotype of *Fah*$^{-/-}$ mice under different conditions with our and other published works (29, 36). As detailedly described in Figure S1, under 100% concentration of NTBC (7.5 mg/L), the liver of *Fah*$^{-/-}$ mice was preserved normally; *Fah*$^{-/-}$ mice developed acute liver injury and died at 3-6 weeks without any NTBC; *Fah*$^{-/-}$ mice survived under 2.5% concentration of NTBC (0.2 mg/L), but suffered chronic liver injury and formed HCC after 12 weeks.

To gain further insight into the molecular mechanisms that contribute to HCC development in chronic liver injury, we first aimed to characterize the dynamic transcriptomic changes in chronic injured liver compared with normal liver. After performing high throughput RNA Sequencing, the differentially-expressed genes between normal livers of *Fah*$^{-/-}$ mice (100% NTBC, named as C0) and chronic injured livers of *Fah*$^{-/-}$ mice (2.5% NTBC for 4 weeks, named as C4) were identified (Fig. 1A-B, Table S1). Two thousand genes with significantly altered expression were identified (log2 Fold Change > 1, P value < 0.05). Genes with expression level altered by more than two-fold were further analyzed by gene enrichment analyses, including Gene Ontology (GO) and KEGG. Several pathways had been enriched in *Fah*$^{-/-}$ mice with Chronic Liver Injury (CLI *Fah*$^{-/-}$ mice) (Fig. 1C, Table S2). HCC represents a classic instance of inflammation-related cancer, and chemically or genetically induced

HCC is highly dependent on inflammatory signaling (3). We found many of these differentially expressed genes are involved in inflammatory processes, and they are also implicated in HCC development (Fig. 1C). Consistently, several KEGG pathway terms, especially the p53 signaling pathway, were altered in *Fah*$^{-/-}$ mice livers (Fig. 1C). The process of retinol metabolism was also significantly altered both in the CLI and normal *Fah*$^{-/-}$ mice (Fig. 1C, blue arrow). Of note, the glucose metabolic process was also enriched in the CLI *Fah*$^{-/-}$ mice according to GO and KEGG analyses (Fig. 1C, red arrow). Body weight and blood glucose level in the CLI *Fah*$^{-/-}$ mice were measured. We found there is similar body weight and significantly lower blood glucose level in the CLI *Fah*$^{-/-}$ mice compared with normal *Fah*$^{-/-}$ mice, indicating that HCC model on the CLI *Fah*$^{-/-}$ mice represents non-diabetes induced HCC (Fig. 1D-E). The function of metformin depends on the activation of AMPK and thereby we wanted to know whether the AMPK activity was changed in our chronic liver injury model of *Fah*$^{-/-}$ mice. We detected the activity of AMPK in *Fah*$^{-/-}$ mice livers of C0 (Uninjured Control, 0 week), C4 (CLI, 4 weeks) and C12 (CLI, 12 weeks), respectively. And we found that the activity of AMPK in liver decreased significantly when the *Fah*$^{-/-}$ mice suffered CLI (Fig. 1F). We tested human patients with normal or cirrhotic liver tissues and found that the AMPK activity is also repressed in chronic injured human liver (Fig. 1G). To understand whether there is a similar trend on the inhibition of AMPK signaling between the human chronic liver injury and the mouse model, we obtained two human chronic liver injury related datasets from GEO

334 database, and ssGSEA algorithm was applied to evaluate whether the AMPK pathway
335 was hampered under chronic liver injury. It was found that with the aggravation of
336 liver fibrosis, the enrichment score of AMPK pathway gradually declined, and the
337 AMPK pathway tended to be suppressed in all status of chronic injured liver
338 compared with normal tissues (Fig. 1H). Similarly, the enrichment score of AMPK
339 pathway in Non-Alcoholic Steatosis Hepatitis (NASH) and simple steatosis groups
340 were significantly lower than those in the healthy control group (Fig. 1I). In addition,
341 at the single gene level, we observed that the expression level of FOXO1, CRY1 and
342 other AMPK pathway core genes were significantly lower in chronic injured liver
343 than those in healthy control (Fig. S2). Beside of these data, preclinical and clinical
344 studies have suggested the activity of AMPK in human liver decreased significantly
345 with chronic liver injury and lower activity of AMPK indicates poorer survival (37-
346 40). All the data above showed that the CLI *Fah*$^{-/-}$ mice have signal pathway altering
347 patterns mimicking human chronic liver diseases, and is a non-diabetic induced HCC
348 mouse model serving as a good indicator of human HCC.

349 **Metformin prevents hepatocarcinogenesis in *Fah*$^{-/-}$ mice with chronic liver injury**

350 Previous studies suggest that metformin treatment inhibits HCC development in
351 several animal models of HCC (19, 20). We tried to figure out whether metformin
352 inhibited HCC formation of the CLI *Fah*$^{-/-}$ mice. To investigate whether the
353 metformin affects HCC formation of the CLI *Fah*$^{-/-}$ mice, we switched from 100% to
354 2.5% NTBC administration at the experimental duration of 8 weeks, at the same time,

355 animals were treated with daily intraperitoneal injections of either metformin (Met12,
356 250 mg/kg body weight) or an equal volume of saline solution (C12) for 12 weeks
357 (Fig. 2A). The main pharmacologic mechanism of metformin is activating the AMPK
358 pathway (11). The results of western blot showed that in our chronic liver injury
359 model, AMPK was activated in the metformin treatment group, indicating that the
360 metformin worked (Fig. 2B). HCC were obviously visible in both macroscopic and
361 histological examination in all *Fah*$^{-/-}$ mice (n=12). Metformin treatment significantly
362 delayed tumor formation; only 50% of the *Fah*$^{-/-}$ mice developed HCC after 12-week
363 2.5% NTBC treatment (n=12) (Fig. 2C). Furthermore, *Fah*$^{-/-}$ livers without metformin
364 treatment displayed a significantly greater number and size of tumors than those
365 treated with metformin (Figs. 2D-E). In addition, analyses on serum liver function
366 indicators were used to confirm the degree of liver injury in both groups. We found
367 the level of serum aspartate transaminase (AST), serum alanine transaminase (ALT),
368 and the AST/ALT ratio were accordingly not significantly changed in both groups
369 (Fig. 2F). Taken together, these data indicate that metformin prevents
370 hepatocarcinogenesis in *Fah*$^{-/-}$ mice with chronic liver injury, even without alleviating
371 liver functions.

372 **Metformin suppressed *Cyp26a1* expression of hepatocyte *in vivo* and *in vitro***

373 The major function of metformin depends on the activation of the AMPK pathway
374 (11). To identify potential AMPK-targeted gene that mediating the metformin-
375 induced hindering of HCC development in *Fah*$^{-/-}$ mice, RNA-Seq and corresponding

differential gene expression analyses on the livers of *Fah*$^{-/-}$ mice, with or without metformin treatment, was performed. The results revealed that the mRNA levels of 200 genes were significantly changed with metformin treatment (log2 Fold Change >1) (Fig. 3A, Table S3). From these 200 altered genes, we presumed the gene *Cyp26a1* was the target of metformin in the hindering of hepatocarcinogenesis for four reasons: (i) Williams *et al.* showed metformin prevents HCC development in the chronic liver injury of *Ncoa5*$^{+/-}$ mice (19). We comparatively re-analyzed our and Williams's sequencing data (GSE110524). There are 200 significantly changed genes in our RNA-Seq data (127 upregulated genes, 73 downregulated genes), 157 significantly differential genes in GSE110524 data (53 upregulated genes, 104 downregulated genes) (19). We conducted intersection analyses of the up-regulated and down-regulated genes between these two data, respectively. There was no up-regulated overlapped genes. Interestingly, only one intersected gene, *Cyp26a1*, was obtained in the down-regulated genes by metformin (Fig. 3B). (ii) The expression of *Cyp26a1* is significantly up-regulated in C4 compared to C0, and the decrease of its expression significantly affects the physiological and biochemical function of liver (Table S1). (iii) *Cyp26a1* encodes one of the cytochrome P450 superfamily enzymes which metabolize all-*trans*-retinoic acid (atRA), 9-*cis*RA and 13-*cis*RA (41). The decrease of *Cyp26a1* leads to the augmentation of atRA level, while *at*RA has been demonstrated as anti-tumor metabolite in several tumor types (42-45). (iv) In the process of chronic liver injury of *Fah*$^{-/-}$ mice, the metabolic signal pathway of retinol

(a precursor of atRA) was enriched by GO and KEGG analysis (Fig. 1C blue arrow), which implies retinol pathway is important in the HCC formation. For all these reasons, we presume metformin may suppressed the expression of *Cyp26a1* to prevent hepatocarcinogenesis in *Fah*$^{-/-}$ mice with chronic liver injury.

CYP26A1 functions primarily to eliminate bioactive atRA (46). RNA-seq data showed *Cyp26a1* gene expression in metformin treatment group was reduced (Fig. 3C). In line with the results from the RNA-Seq analysis, *Cyp26a1* mRNA levels were down-regulated in metformin-treated liver tissues of *Fah*$^{-/-}$ mice according to the qRT-PCR assay (Fig. 3D). Western blot also showed CYP26A1 protein expression was reduced in metformin-treated liver tissues (Fig. 3E). In the adult, *Cyp26a1* is mostly highly expressed in the liver (47). We examined the cellular source of *Cyp26a1* in the mouse liver. Perfused mouse liver was separated to hepatocytes and non-parenchymal cells by low-speed centrifugation (50 g for 2 min). qRT-PCR revealed that *Cyp26a1* was mainly expressed by hepatocyte in the mouse liver (Fig. 3F). The scRNA-seq data of immune cells of another HBV-induced HCC model in Fah$^{-/-}$ mouse (35) also showed that only 23 of 27058 immune cells expressed *Cyp26a1* (data not shown). Next, we examined whether metformin directly suppressed the *Cyp26a1* expression in hepatocytes. Perfused hepatocytes were plated on a plate containing matrigel. qRT-PCR and western blot showed that the expression of *Cyp26a1* in hepatocytes decreased with the increase of metformin concentration (Fig. 3G-H). Protein CYP26A1 plays a major role in atRA clearance which means down-regulation of

418 *Cyp26a1* leads to the elevation of atRA level (41). The pGL3-RARE-luciferase is a
419 plasmid contains retinoic acid receptor (RARE) response element in front of
420 luciferase (firefly) reporter gene, which can reflect the level of atRA (45, 48). We
421 transfected this plasmid and control plasmid (pRL-CMV-renilla-luciferase) to primary
422 hepatocytes of the *Fah*$^{-/-}$ mice with chronic liver injury. The Firely/renilla luciferase
423 activity of pGL3-RARE-luciferase is increased with atRA treatment while pGL3-
424 promoter-luciferase is unchanged (Fig. 3I), which means pGL3-RARE-luciferase
425 activity could reflect the level of atRA in hepatocytes. We then tested pGL3-RARE-
426 luciferase and pGL3-promoter-luciferase activity on the metformin treatment. We
427 found that the Firely/renilla luciferase activity of pGL3-RARE-luciferase is increased
428 while pGL3-promoter-luciferase is unchanged with metformin treatment, which
429 means metformin increased the level of atRA (Fig. 3J). Together, these data
430 demonstrated metformin directly suppresses the *Cyp26a1* gene expression of
431 hepatocytes and thereby leads to the elevated level of atRA, thereafter exerting anti-
432 tumor effects.

433 **Metformin decreased the expression of *Cyp26a1* through the AMPK/JNK/c-Jun**
434 **signaling**

435 The AMPK is a major effector of metformin (49), and acadesine (AICAR) is
436 another AMPK activator (50). We treated perfused primary adherent hepatocytes of
437 *Fah*$^{-/-}$ mice with different concentration of AICAR. qRT-PCR and western blot
438 showed the expression of *Cyp26a1* was also suppressed by AICAR (Fig. 4A-B). We

speculated that the AMPK may be involved in the decreased expression of *Cyp26a1* induced by metformin and AICAR. Compound C (dorsomorphin) is an effective, reversible and selective AMPK inhibitor, and is widely used to imply AMPK-dependence experimentally (49). Compound C (80 μM) attenuates the ability of metformin in activating AMPK in mouse hepatocytes (Fig. 4D). qRT-PCR and western blot showed Compound C attenuates the ability of metformin in suppressing the *Cyp26a1* expression which means the AMPK is a major effector of metformin in suppressing *Cyp26a1* gene expression (Fig. 4C-D).

To determine whether metformin mediated *Cyp26a1* expression through the repression of its promoter, we constructed plasmid pGL3-cyp26a1-promoter-2000bp-luciferase, which was generated by using the promoter of mouse *Cyp26a1* gene (-2000 to +60 relation to transcription start site) to replace SV40 promoter of pGL3-promoter-luciferase. The mouse *Cyp26a1* promoter activity was significantly suppressed after metformin treatment (Figure 4E) while pGL3-promoter-luciferase is not changed (Fig. 3J), which means metformin-suppressed *Cyp26a1* expression is mediated through the repression of its promoter.

To define the roles of the *cis*-regulatory elements of the *Cyp26a1* promoter in response to metformin regulation, a series of truncated mutants of the *Cyp26a1* promoter were generated (-750 to +60, -500 to +60, -250 to +60 and -50 to +60). pGL3-cyp26a1-promoter-750bp-luciferase also showed significantly decreased luciferase activity by metformin. The shorter truncated mutants (-500 to +60, -250 to

+60 and -50 to +60) of the *Cyp26a1* promoter blocked metformin-decreased *Cyp26a1* promoter activity, indicating that the sequence between nucleotide -750 and -500 was critical for the suppression of the *Cyp26a1* by metformin treatment (Fig. 4F). We predicted the transcription factors of this sequence (nucleotide -750 to -500 relation to transcription start site) using PROMO database (http://alggen.lsi.upc.es/cgibin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3). There are forty predicted transcription factors (Fig. S3), and among them a predicted c-Jun binding site is located in this region. Previous study showed the JNK/c-Jun signal pathway could be regulated by metformin-AMPK pathway (51). We found P-JNK and P-c-Jun was inhibited by metformin in hepatocytes (Fig. 4B). When activity of AMPK was pharmacologically inhibited by Compound C, such reduction by metformin treatment in *Cyp26a1*, P-JNK and P-c-Jun expression were not observed (Fig. 4D). Furthermore, we used JNK inhibitor SP600125 to determine whether metformin-decreased *Cyp26a1* expression is mediated through the repressed of JNK. In the presence of SP600125, JNK activity is suppressed and metformin lost the effect of inhibiting JNK because JNK activity has been inhibited in advance by SP600125. In parallel, *Cyp26a1* could not be decreased by metformin even AMPK is activated when JNK activity was not suppressed by metformin (Figs. 4G-H). Taken together, the results demonstrate that the *Cyp26a1* expression is inhibited by metformin through the AMPK/JNK/c-Jun signal pathway.

atRA supplementation decreases HCC formation in chronic liver injury model

Metformin could promote the level of atRA through downregulation of *Cyp26a1* gene expression (Fig. 3). The atRA is usually considered as an anti-tumor small molecule (41). We wanted to study the antitumor effect of atRA in our chronic liver injury model. IP injection of atRA was administered to *Fah*$^{-/-}$ mice at the beginning of chronic liver injury (Fig. 5A). Tumor incidence, tumor number and tumor diameter in atRA-treated mice were significantly reduced compared to vehicle-treated mice, respectively (Fig. 5B-D). We also found that the levels of serum aspartate transaminase (AST), serum alanine transaminase (ALT), and AST/ALT ratio were accordingly not significantly changed in atRA treatment compared to untreated control (Fig. 5E). These data demonstrate a robust therapeutic benefit conferred by atRA supplementation in our chronic liver injury model.

Metformin and atRA both reduce CD8$^+$ T cell number and ratio in CLI mouse model

Recent publications have shown that metformin is able to modulate the interaction between tumor cells and their microenvironment and thus presenting an immune-mediated antitumor effect (52-55). In addition, the antitumor properties of atRA are also closely related to immune cells (45). To explore the mechanism behind tumor suppression, we studied immune cell subsets in the CLI *Fah*$^{-/-}$ mice on metformin or atRA treatment. A prominent expansion of the proportion of hepatic CD8$^+$ cells among the hepatic CD45$^+$ cells was found in C12 mice when compared with C0 mice, and the proportion of CD8$^+$ T cells were significantly reduced in

metformin-treated mice (Met12) compared with C12 mice, whereas no changes were found in other immune subsets (Kupffer cells, $CD4^+$ T, NK and $CD3^+$ cells) (Fig. 6A-B). Although the proportion of $CD4^+$ in all $CD45^+$ cells were significantly increased in metformin-treated mice (Met12) compared with C12 mice, there is no change between C12 and C0 mice, so we focused on the role of $CD8^+$ T cells in the metformin-induced tumor formation hindering. A similarity result was found by IHC staining, chronic injured liver (C12) contained much more $CD45^+$ cells and $CD8^+$ T cells than the liver of mice on 100% NTBC (C12 VS C0), the number of $CD8^+$ T cells were reduced by metformin treatment while the number of $CD45^+$ cells remained unchanged (Met12 VS C12) (Figs. 6C-D). Accordingly, the number of hepatocytes per field significantly dropped in the histological sections of the CLI *Fah*$^{-/-}$ mice compared with normal liver tissue sections, which hinted that the size of hepatocytes enlarged significantly in the CLI *Fah*$^{-/-}$ mice compared with those in normal liver (C12 VS C0) (Fig. 6D). Hepatic $CD8^+$ T cells are integral to antitumor immunity via direct antigen-specific cytotoxic targeting of tumors. However, recent researches showed that the $CD8^+$ T cells contribute to HCC tumor formation in chronic liver injury model (19, 25, 56, 57). These papers showed the $CD8^+$ T cells secrete pro-tumor cytokines (IL-1β, IL-6, Tnfa and Ltβ) in the process of chronic liver injury. Especially, $CD8^+$ T cells as pro-tumor immune cell has been demonstrated in CLI *Fah*$^{-/-}$ mice, although the adopted mouse model is a little different to our CLI *Fah*$^{-/-}$ mice (25). We aimed to delineate which cytokine mediates the impact of $CD8^+$ T cells

523 on hepatocarcinogenesis in our disease model. Pro-tumor cytokines mRNA levels
524 were measured by qRT-PCR, and the induction of Tnfa and Ltβ strikingly correlated
525 with tumor development in the CLI *Fah*$^{-/-}$ mice. Moreover, chemokines like the Tnfa
526 and Ltβ were significantly reduced in metformin treatment (Fig. 6E). Together, our
527 results indicate that metformin could reduce the number and ratio of CD8^{+} T cells to
528 suppress HCC formation in the CLI *Fah*$^{-/-}$ mice.

529 Next, we sought to clarify the mechanism that was responsible for the anti-tumor
530 effect of atRA. Since CD8^{+} T cells play a crucial role in pro-tumor immunity in our
531 model, we examined the CD8^{+} T cell and other immune cells in the atRA-treated CLI
532 *Fah*$^{-/-}$ mice. A significant decrease in the percentage of CD8^{+} T cells was found after
533 atRA treatment compared with vehicle treatment (atRA12 vs C12) (Fig. 6F). Besides,
534 a significant decrease in the percentage of NK cells was found after atRA treatment
535 while was not found after metformin treatment (Fig. 6B). There were no changes in
536 other immune subsets after atRA treatment compared with vehicle treatment (Kupffer
537 cells, CD4^{+} T and CD3^{+} cells). Different effects between metformin and atRA
538 treatment indicate that the signal pathways regulated by metformin and atRA are not
539 completely consistent in our mouse model. IHC staining also showed that with atRA
540 treatment, the number of CD8^{+} T cells decreased significantly while the number of
541 CD45^{+} cells was not changed (Fig. 6G-H). We also tested the expression of pro-tumor
542 cytokines that secreted by CD8^{+} T cells. qRT-PCR showed Tnfa and Ltβ was
543 decreased significantly with atRA treatment compared to untreated, similar to

544 metformin treatment (Fig. 6I). These data indicated the inhibition of HCC by
545 metformin is at least partly due to the promotion of elevated atRA level.

546 **scRNA-seq of immune cells of *Fah*$^{-/-}$ mouse model reveals hepatic resident-like**
547 **CD8$^+$ T cells are gradually increased following the development of chronic liver**
548 **injury**

549 A previous study has reported that the CLI environment of *Fah*$^{-/-}$ mouse model
550 enables to promote the carcinogenesis of implanted HBsAg$^+$ hepatocytes and found
551 that HBsAg-specific CD8$^+$ T cells participate in this process (35). To figure out
552 whether CD8$^+$ T cells could express pro-tumor cytokines and atRA receptor, we
553 obtained related scRNA-seq data. In order to confirm the consistency of transcriptome
554 level with the conclusion we found using flow cytometry assays, CD8$^+$ T cells were
555 divided into Cd8a-high expression group and Cd8a-low expression group to better
556 observe the dynamic changes of CD8$^+$ T cells during tumor progression at the mRNA
557 level. The distribution of CD8$^+$ T cells with high Cd8a expression was more likely to
558 be in the same clusters while CD8$^+$ T cells with low Cd8a expression were distributed
559 separately and mixed with other immune cells (Fig. 7A). Interestingly, results showed
560 CD8$^+$ T cells will gradually increase after suffering CLI, especially CD8$^+$ T cells with
561 high Cd8a expression (Figs. 7A-C and S4A-C). Then, we could find that the cell
562 proportion of CD8$^+$ T cells expressing pro-tumor cytokines/molecules was higher
563 than other immune cells, especially Ltb, Pdcd1, Cxcr6 and Il1b (Fig. 7D). And
564 expression levels of most pro-tumor cytokines were higher in CD8$^+$ T cells with high

565 Cd8a expression (Fig. S4D). In addition, the expression levels of atRA receptor (Rara
566 and Rarb) were gradually increased following the CLI time (Figs. 7E, S4E). However,
567 $CD8^+$ T cells tended to express Rara while little could express Rarb (Figs. 7F, S4F).
568 And the expression level of both Rara and Rarb were correlated with $CD8^+$ T cells
569 (Figs. 7G, S4G). Then, we further explored the expression characters of pro-tumor
570 molecules in $CD8^+$ T cells. Intriguingly, Tnf, Pdcd1, Cxcr6, Il1b and Il6, excluding
571 Ltb, were all gradually up-regulated along the tumorigenesis of *Fah*$^{-/-}$ mouse model
572 (Figs. S5-10A). Then, Tnf, Ltb, Pdcd1, Il1b and Cxcr6, excluding Il6, tended to
573 highly express and distributed in the same $CD8^+$ T cell clusters (Figs. S5-10B). And
574 all these pro-tumor genes were correlated with the expression of Cd8a (Figs. S5-10C)
575 and Rara (Figs. S5-10D). In short, $CD8^+$ T cells were responsible for the liver
576 tumorigenesis of *Fah*$^{-/-}$ mouse via expressing pro-tumor molecules.

577 **Hepatic resident-like $CD8^+$ T cells are increased in patients with chronic liver**
578 **injury**

579 To assess the relevance of our findings between precancerous mouse livers and
580 human HCCs, adjacent liver tissue and HCC samples from The Cancer Genome Atlas
581 (TCGA) and normal liver tissue from GTEx database were downloaded and analyzed
582 for the expression levels of CD8A and pro-tumor cytokines. We found that $CD8^+$ T
583 cells increased in the adjacent tumor tissues compared with both HCC tissues and
584 normal liver tissues (Fig. 8A). Also, we compared healthy controls (HC), simple
585 steatosis (SS) and nonalcoholic steatohepatitis (NASH) from GSE89632 database,

586 and found the number of $CD8^+$ T cells increased in the human chronic injured liver
587 compared with normal liver (Fig. 8B). Notably, previous paper also found that the
588 fraction of $CD8^+$ T cells are higher in HCC and HCC adjacent tissue than in healthy
589 liver tissue, while HCC adjacent tissue contained even more T cells than HCC (58), in
590 concordant with our analysis (Fig. 8A). We also found that all pro-tumor cytokines
591 were up-regulated in adjacent tumor tissues while they were not consistent in tumor
592 tissues when both compared to normal tissues (Figs. 8C and 8D, S11). Moreover, we
593 found a significant higher positive regression coefficient between CD8A expression
594 and TNF or LTβ in adjacent tumor tissues other than HCC tissues and normal tissues
595 (Figs. 8D, S12-14A). In parallel, other four pro-tumor molecules had consistent
596 pattern like TNF or LTβ (Figs. S12-14B). And pan-tissues and pan-cancers analyses
597 show significant correlation between $CD8^+$ T cells and these pro-tumor cytokines in
598 most tissues and cancer types, especially liver tissues and HCC, hinting that $CD8^+$ T
599 cells contributed to the expression of pro-tumor cytokines (Figs. 8D, S12-14B). Taken
600 together, these results may indicate that the $CD8^+$ T cells could express pro-tumor
601 cytokines and might have different expression levels and functions in adjacent tumor
602 tissue and HCC tissues.

603 Next, we examined the clinical significance of CD8A in liver cancer. When
604 patients were divided into "high" and "low" CD8 expression based on top 20% value
605 of CD8A, we observed that elevated CD8A expression was associated with better
606 patient survival in TCGA liver cancer database (Fig. 8E). This association means

607 $CD8^+$ T cells may function as pro-tumor cells in adjacent tumor tissues while as anti-
608 tumor cells in tumor tissues, dependent on the expression levels of pro-tumor or anti-
609 tumor cytokines secreted by $CD8^+$ T cells. Altogether, these results showed the
610 number and function of hepatic resident-like $CD8^+$ T cells were changed in livers with
611 varying degrees of injury (normal liver, chronic injured liver and HCC).

612 Taken together, these results show metformin suppressed the expression of
613 *Cyp26a1* through the AMPK/JNK/c-Jun pathway and leading to the elevation of atRA
614 level. The atRA treatment inhibited $CD8^+$ T cells proliferation by binding to atRA
615 receptor (RAR) on the cell membrane of $CD8^+$ T cells. The decrease of pro-tumor
616 $CD8^+$ T cells can reduce the release of pro-tumor cytokines (especially Tnfa and Ltβ)
617 and inhibit hepatocarcinogenesis (Fig. 8F)

Discussion

Currently, systemic options for the prevention and treatment of HCC are limited. New therapeutic targets, preventive strategies as well as biomarkers for patient stratification need to be discovered. Therefore, it is a wise choice to test new therapeutic targets and preventive strategies using animal model. HCC development is a complex and multistep process caused by diverse risk factors, determining that no adaptive animal model can fully mimic the occurrence and development of human HCC. Human *Fah*$^{-/-}$ induced HCC occurs in a very low incidence. The murine model of *Fah*-deficiency is a suitable animal model, which represents all phenotypic and biochemical characterizations of patients with FAH deficiency (24). Furthermore, the HCC model of CLI *Fah*$^{-/-}$ mice is highly similar to human alcohol-induced HCC or c-myc-altered HCC (25). So far, there are many studies using this animal model to discover and verify relevant biological phenomena and mechanisms related to HCC. These indicates that the study of *Fah*-deficient chronic injury mice could be a good indicator of human HCC.

In this study, we revealed altered transcriptomic profiles and signaling pathways in the early stages of chronic liver injury during hepatocarcinogenesis, and identified hepatic responses to metformin treatment. Our results uncovered that the inhibition of AMPK activity and the symptoms of hypoglycemia are co-existent in the model of chronic liver injury (Fig. 1D, F). Besides, AMPK activity is repressed in kinds of human liver disease (Fig, 1G-I) (37, 38). AMPK activity is reduced by inflammation,

obesity, and diabetes. The activation of AMPK pathway has been viewed as a viable therapeutic strategy to improve HCC (59). These results suggest that metformin may improve various liver diseases by up-regulating the repressed AMPK activity under disease conditions.

Metformin is an AMPK activator and blood glucose regulator, which prompted us to investigate the role and mechanism of metformin in our mouse HCC model. Epidemiological studies have shown a reduction in incidence rate and mortality of liver cancer in type 2 diabetes patients treated with metformin (60-62). In our mouse model, metformin treatment significantly reduced the HCC incidence of *Fah*$^{-/-}$ mice (Fig. 2C-E). Chronic inflammation contributes to hepatocarcinogenesis through multiple mechanisms, one of which is the secretion of pro-tumor cytokines by immune cells. The number and proportion of CD8^{+} T cells increased significantly during chronic liver injury, and metformin could reverse this effect in *Fah*$^{-/-}$ mice (Fig. 6B-D). CD8^{+} T cells could be classified based on their molecular and functional characteristics. Generally, tumor infiltrating CD8^{+} T cells are anti-tumor immune cells and associated with a favorable prognosis (63, 64). On the other hand, CD8^{+} T cells with high expression of inhibitory receptors (PD-1, TIGIT, TIM-3, LAG-3) is defined as exhausted CD8^{+} T cells, which progressively loses effector function and possesses a poor outcome (65, 66).

Recently, several studies showed CD8^{+} T cells contributes to HCC development in many types of mouse chronic liver injury (19, 25, 56, 57). Endig *et al.* showed that

660 $CD8^+$ T cells and LTβ signaling contributes to HCC development in *Fah*$^{-/-}$ mice with
661 chronic liver injury, which is exactly the same as the mouse model we used, but the
662 way of inducing chronic liver injury is slightly different (25). Pfister *et al.* reported in
663 a preclinical model of NASH-induced HCC, $CD8^+$ T cells contribute to the induction
664 of NASH-HCC rather than invigorating or executing immune surveillance (57).
665 Concordantly, another group has revealed that auto-aggression of $CD8^+$ T cells in the
666 liver may be involved in the development of HCC in patients with NASH (56).
667 Williams *et al.* also observed the enrichment of exhausted $CD8^+$ T cells in the
668 precancerous liver of *Ncoa5*$^{+/-}$ mouse and these T cells function as a pro-tumor
669 microenvironment. Meanwhile, metformin treatment is able to reduce hepatic
670 infiltration of $CD8^+$ T cells and reduce HCC formation in the chronic liver injury of
671 *Ncoa5*$^{+/-}$ mice(19). All these findings suggest that the classification (effector or
672 exhaustion) and location (tumor or adjacent tumor) of $CD8^+$ T cells determine
673 whether they are pro-tumor or anti-tumor T cells. Our data proved the number and
674 ratio of $CD8^+$ T cells were increased in the process of chronic liver injury, and
675 metformin could reduce the number and ratio of $CD8^+$ T cells in liver and then
676 suppressed the HCC formation in *Fah*$^{-/-}$ mice (Figs. 6B-D). The single cell RNA-seq
677 analysis of $CD8^+$ T cells in *Fah*$^{-/-}$ mice reveals that $CD8^+$ T cells could express pro-
678 tumor cytokines and atRA receptors (Fig. 7). Interestingly, this study also found
679 that the fraction of $CD8^+$ T cells was higher in HCC and HCC adjacent tissues than in
680 healthy liver tissue, while HCC adjacent tissue contained even more $CD8^+$ T cells

than HCC (Fig. 8) (58). Thus, in addition to gaining insights into the mechanism of actions of metformin, our data could be valuable in supporting that metformin treatment may be beneficial in reversing $CD8^+$ T cell-exhausted tumor microenvironment for HCC patients.

Retinoic acid (RA) is an active metabolite of vitamin A. Retinoic acid includes all-*trans* retinoic acid (atRA), 9-*cis*-retinoic acid (9-*cis* RA), and 13-*cis*-retinoic acid (13-*cis* RA)(41). Vitamin A is irreversibly converted to retinoic acid or atRA by the aldehyde dehydrogenase (ALDH) enzyme family members. The atRA was degraded by atRA-degrading cytochrome P450 reductases, such as Cyp26a1, which converts atRA to inactive metabolites(41). The up-regulation of ALDH or down-regulation of Cyp26a1 will lead to the increase of atRA level. In our study, we found the expression of Cyp26a1 was decreased by metformin treatment *in vivo* and *in vitro* (Fig. 3). The decrease of *Cyp26a1* gene expression will result in the increase of atRA level. Besides, metformin directly increased the luciferase activity of atRA response element plasmid, further demonstrating that metformin promotes the level of atRA in hepatocytes (Fig. 3J). Preclinical and clinical studies of atRA against tumor are reported increasingly. A classic example of clinical use is the treatment of acute promyelocytic leukemia (APL), which is the most efficacious use of atRA in cancer therapy (42, 43). Bryan *et al.* investigated that atRA combined with paclitaxel had better overall clinical efficacy than paclitaxel alone in the treatment of recurrent or metastatic breast cancer (44). Han *et al.* showed that the storage of retinaldehyde was

significantly decreased in patients with HCC and found that retinol metabolism has great prospects for clinical application in diagnosis, prognosis and chemotherapy of HCC (67). There are several mechanisms for the anti-tumor activity of atRA, one of which is the effect on the immune systems (45, 68). In our study, we also proved that atRA agents possess anti-tumor activity through their ability to reduce the number and ratio of $CD8^+$ T cells in *Fah*$^{-/-}$ mice (Fig. 6). $CD8^+$ T cells expressed the atRA receptor RAR (Fig. 7 and S4E) and responded to atRA (69). The atRA treatment inhibited T cell proliferation in a dose-dependent manner *in vivo* and *in vitro* (68, 70). These may illustrate metformin can reduce the number and ratio of $CD8^+$ T cells through elevate the level of atRA in our mouse model. Metformin/Cyp26a1/atRA signaling pathway may partly explain how metformin affected $CD8^+$ T cell proliferation and differentiation.

Metformin and atRA are both identified as anti-neoplastic agents. Our data show both metformin and atRA reduce the number and proportion of $CD8^+$ T cells which was considered as pro-tumor immune cells in our mouse model (Fig. 6). In summary, we firstly revealed that metformin increases the level of atRA by inhibiting the expression of *Cyp26a1* gene, and then inhibits HCC caused by chronic liver injury (Fig. 8F).

Acknowledgments

The authors thank Yanxiang Song, Hua Qiu, Cheng Wang for tissue preparation,

sectioning, and staining; Xiuhua Li, Changcheng Liu for Flow Cytometric analysis; Yanli Lu, Xiaolan Mu, Xinyue Du, Shengwei Shen, Mingyang Xu, Shihan Sun for their kindly assistance and meaningful discussion during manuscript preparation. This work was funded by Major Program of National Key Research and Development Project (2020YFA0112600, 2019YFA0801502), National Natural Science Foundation of China (82002945, 82173019, 82103095), Jiangxi Provincial Natural Science Foundation (20212ACB206033), the Project of Shanghai science and technology commission (19140902900), Program of Shanghai Academic/Technology Research Leader (20XD1434000), and Peak Disciplines (Type IV) of Institutions of Higher Learning in Shanghai.

Author Contributions

WH, MC, CL designed and performed experiments, analyzed data; WC, YY, YC, ZY performed animal experiments; XW, GW, LP performed bioinformatics analysis; WH, YC performed immunohistochemistry analysis; ZH, JW, QT supervised and planned research; ZH conceptualized study, supervised and planned research. WH, ZH wrote the paper.

Competing Interests

The authors declare no competing interests.

Data Availability Statement

All data generated or analysed during this study are included in the manuscript and supporting file; Source Data files have been provided for Figures 1 and 3.

Reference

1. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet. 2018;391(10127):1301-14.
2. Villanueva A. Hepatocellular Carcinoma. N Engl J Med. 2019;380(15):1450-62.
3. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. Nat Rev Dis Primers. 2021;7(1):6.
4. Seo W, Gao Y, He Y, Sun J, Xu H, Feng D, et al. ALDH2 deficiency promotes alcohol-associated liver cancer by activating oncogenic pathways via oxidized DNA-enriched extracellular vesicles. J Hepatol. 2019;71(5):1000-11.
5. Gao S, Li A, Liu F, Chen F, Williams M, Zhang C, et al. NCOA5 haploinsufficiency results in glucose intolerance and subsequent hepatocellular carcinoma. Cancer Cell. 2013;24(6):725-37.
6. van Spronsen FJ, Bijleveld CM, van Maldegem BT, Wijburg FA. Hepatocellular carcinoma in hereditary tyrosinemia type I despite 2-(2 nitro-4-3 trifluoro-methylbenzoyl)-1, 3-cyclohexanedione treatment. J Pediatr Gastroenterol Nutr. 2005;40(1):90-3.
7. Bueno MJ, Ruiz-Sepulveda JL, Quintela-Fandino M. Mitochondrial Inhibition: a Treatment Strategy in Cancer? Curr Oncol Rep. 2021;23(4):49.
8. Andrzejewski S, Gravel SP, Pollak M, St-Pierre J. Metformin directly acts on mitochondria to alter cellular bioenergetics. Cancer Metab. 2014;2:12.
9. Owen MR, Doran E, Halestrap AP. Evidence that metformin exerts its anti-diabetic effects through inhibition of complex 1 of the mitochondrial respiratory chain. Biochem J. 2000;348 Pt 3:607-14.
10. El-Mir MY, Nogueira V, Fontaine E, Averet N, Rigoulet M, Leverve X. Dimethylbiguanide inhibits cell respiration via an indirect effect targeted on the respiratory chain complex I. J Biol Chem. 2000;275(1):223-8.
11. Cui Q, Wen S, Huang P. Targeting cancer cell mitochondria as a therapeutic approach: recent updates. Future Med Chem. 2017;9(9):929-49.
12. Ma T, Tian X, Zhang B, Li M, Wang Y, Yang C, et al. Low-dose metformin targets the lysosomal AMPK pathway through PEN2. Nature. 2022;603(7899):159-65.
13. Viale A, Pettazzoni P, Lyssiotis CA, Ying H, Sanchez N, Marchesini M, et al. Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function. Nature. 2014;514(7524):628-32.
14. Villani LA, Smith BK, Marcinko K, Ford RJ, Broadfield LA, Green AE, et al. The diabetes medication Canagliflozin reduces cancer cell proliferation by inhibiting mitochondrial complex-I supported respiration. Mol Metab. 2016;5(10):1048-56.
15. Roesch A, Vultur A, Bogeski I, Wang H, Zimmermann KM, Speicher D, et al. Overcoming intrinsic multidrug resistance in melanoma by blocking the mitochondrial respiratory chain of slow-cycling JARID1B(high) cells. Cancer Cell. 2013;23(6):811-25.

16. Shackelford DB, Abt E, Gerken L, Vasquez DS, Seki A, Leblanc M, et al. LKB1 inactivation dictates therapeutic response of non-small cell lung cancer to the metabolism drug phenformin. Cancer Cell. 2013;23(2):143-58.
17. Chen HP, Shieh JJ, Chang CC, Chen TT, Lin JT, Wu MS, et al. Metformin decreases hepatocellular carcinoma risk in a dose-dependent manner: population-based and in vitro studies. Gut. 2013;62(4):606-15.
18. Tseng CH. Metformin and risk of hepatocellular carcinoma in patients with type 2 diabetes. Liver Int. 2018;38(11):2018-27.
19. Williams M, Liu X, Zhang Y, Reske J, Bahal D, Gohl TG, et al. NCOA5 deficiency promotes a unique liver protumorigenic microenvironment through p21(WAF1/CIP1) overexpression, which is reversed by metformin. Oncogene. 2020;39(19):3821-36.
20. Shankaraiah RC, Callegari E, Guerriero P, Rimessi A, Pinton P, Gramantieri L, et al. Metformin prevents liver tumourigenesis by attenuating fibrosis in a transgenic mouse model of hepatocellular carcinoma. Oncogene. 2019;38(45):7035-45.
21. Jorquera R, Tanguay RM. Fumarylacetoacetate, the metabolite accumulating in hereditary tyrosinemia, activates the ERK pathway and induces mitotic abnormalities and genomic instability. Hum Mol Genet. 2001;10(17):1741-52.
22. Daou KN, Barhoumi A, Bassyouni A, Karam PE. Diagnostic and Therapeutic Challenges of Hereditary Tyrosinemia Type 1 in Lebanon: A 12-Year Retrospective Review. Front Pediatr. 2021;9:698577.
23. McKiernan PJ. Nitisinone in the treatment of hereditary tyrosinaemia type 1. Drugs. 2006;66(6):743-50.
24. Grompe M, Overturf K, al-Dhalimy M, Finegold M. Therapeutic trials in the murine model of hereditary tyrosinaemia type I: a progress report. J Inherit Metab Dis. 1998;21(5):518-31.
25. Endig J, Buitrago-Molina LE, Marhenke S, Reisinger F, Saborowski A, Schutt J, et al. Dual Role of the Adaptive Immune System in Liver Injury and Hepatocellular Carcinoma Development. Cancer Cell. 2016;30(2):308-23.
26. Yoon SH, Choi SW, Nam SW, Lee KB, Nam JW. Preoperative immune landscape predisposes adverse outcomes in hepatocellular carcinoma patients with liver transplantation. NPJ Precis Oncol. 2021;5(1):27.
27. Arendt BM, Comelli EM, Ma DW, Lou W, Teterina A, Kim T, et al. Altered hepatic gene expression in nonalcoholic fatty liver disease is associated with lower hepatic n-3 and n-6 polyunsaturated fatty acids. Hepatology. 2015;61(5):1565-78.
28. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14:7.
29. Wang C, Chen WJ, Wu YF, You P, Zheng SY, Liu CC, et al. The extent of liver injury determines hepatocyte fate toward senescence or cancer. Cell Death Dis. 2018;9(5):575.
30. He Z, Zhang H, Zhang X, Xie D, Chen Y, Wangensteen KJ, et al. Liver xeno-repopulation with human hepatocytes in Fah-/-Rag2-/- mice after pharmacological immunosuppression. Am J Pathol. 2010;177(3):1311-9.

826 31. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
827 sequence data. Bioinformatics. 2014;30(15):2114-20.
828 32. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory
829 requirements. Nat Methods. 2015;12(4):357-60.
830 33. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-
831 throughput sequencing data. Bioinformatics. 2015;31(2):166-9.
832 34. Wang MJ, Chen F, Liu QG, Liu CC, Yao H, Yu B, et al. Insulin-like growth factor
833 2 is a key mitogen driving liver repopulation in mice. Cell Death Dis. 2018;9(2):26.
834 35. Hao X, Chen Y, Bai L, Wei H, Sun R, Tian Z. HBsAg-specific CD8(+) T cells as
835 an indispensable trigger to induce murine hepatocellular carcinoma. Cell Mol Immunol.
836 2021;18(1):128-37.
837 36. Marhenke S, Lamle J, Buitrago-Molina LE, Canon JM, Geffers R, Finegold M, et
838 al. Activation of nuclear factor E2-related factor 2 in hereditary tyrosinemia type 1 and
839 its role in survival and tumor development. Hepatology. 2008;48(2):487-96.
840 37. Zheng L, Yang W, Wu F, Wang C, Yu L, Tang L, et al. Prognostic significance of
841 AMPK activation and therapeutic effects of metformin in hepatocellular carcinoma.
842 Clin Cancer Res. 2013;19(19):5372-80.
843 38. Zheng LY, Wu L, Lu J, Zou DJ, Huang Q. Expression of Phosphorylated AMP-
844 Activated Protein Kinase Predicts Response to Transarterial Chemoembolization in
845 Postoperative Cases of Hepatocellular Carcinoma. Medicine (Baltimore).
846 2016;95(11):e2908.
847 39. Yang X, Liu Y, Li M, Wu H, Wang Y, You Y, et al. Predictive and preventive
848 significance of AMPK activation on hepatocarcinogenesis in patients with liver
849 cirrhosis. Cell Death Dis. 2018;9(3):264.
850 40. Zhao P, Sun X, Chaggan C, Liao Z, In Wong K, He F, et al. An AMPK-caspase-6
851 axis controls liver damage in nonalcoholic steatohepatitis. Science.
852 2020;367(6478):652-60.
853 41. Hunsu VO, Facey COB, Fields JZ, Boman BM. Retinoids as Chemo-Preventive
854 and Molecular-Targeted Anti-Cancer Therapies. Int J Mol Sci. 2021;22(14).
855 42. Breitman TR, Selonick SE, Collins SJ. Induction of differentiation of the human
856 promyelocytic leukemia cell line (HL-60) by retinoic acid. Proc Natl Acad Sci U S A.
857 1980;77(5):2936-40.
858 43. Sanz MA, Fenaux P, Tallman MS, Estey EH, Lowenberg B, Naoe T, et al.
859 Management of acute promyelocytic leukemia: updated recommendations from an
860 expert panel of the European LeukemiaNet. Blood. 2019;133(15):1630-43.
861 44. Bryan M, Pulte ED, Toomey KC, Pliner L, Pavlick AC, Saunders T, et al. A pilot
862 phase II trial of all-trans retinoic acid (Vesanoid) and paclitaxel (Taxol) in patients with
863 recurrent or metastatic breast cancer. Invest New Drugs. 2011;29(6):1482-7.
864 45. Bhattacharya N, Yuan R, Prestwood TR, Penny HL, DiMaio MA, Reticker-Flynn
865 NE, et al. Normalizing Microbiota-Induced Retinoic Acid Deficiency Stimulates
866 Protective CD8(+) T Cell-Mediated Immunity in Colorectal Cancer. Immunity.
867 2016;45(3):641-55.

868 46. Topletz AR, Tripathy S, Foti RS, Shimshoni JA, Nelson WL, Isoherranen N.
869 Induction of CYP26A1 by metabolites of retinoic acid: evidence that CYP26A1 is an
870 important enzyme in the elimination of active retinoids. Mol Pharmacol.
871 2015;87(3):430-41.
872 47. Xi J, Yang Z. Expression of RALDHs (ALDH1As) and CYP26s in human tissues
873 and during the neural differentiation of P19 embryonal carcinoma stem cell. Gene Expr
874 Patterns. 2008;8(6):438-42.
875 48. Hoffman LM, Garcha K, Karamboulas K, Cowan MF, Drysdale LM, Horton WA,
876 et al. BMP action in skeletogenesis involves attenuation of retinoid signaling. J Cell
877 Biol. 2006;174(1):101-13.
878 49. Zhou G, Myers R, Li Y, Chen Y, Shen X, Fenyk-Melody J, et al. Role of AMP-
879 activated protein kinase in mechanism of metformin action. J Clin Invest.
880 2001;108(8):1167-74.
881 50. Corton JM, Gillespie JG, Hawley SA, Hardie DG. 5-aminoimidazole-4-
882 carboxamide ribonucleoside. A specific method for activating AMP-activated protein
883 kinase in intact cells? Eur J Biochem. 1995;229(2):558-65.
884 51. Woo SL, Xu H, Li H, Zhao Y, Hu X, Zhao J, et al. Metformin ameliorates hepatic
885 steatosis and inflammation without altering adipose phenotype in diet-induced obesity.
886 PLoS One. 2014;9(3):e91111.
887 52. Pearce EL, Walsh MC, Cejas PJ, Harms GM, Shen H, Wang LS, et al. Enhancing
888 CD8 T-cell memory by modulating fatty acid metabolism. Nature.
889 2009;460(7251):103-7.
890 53. Cha JH, Yang WH, Xia W, Wei Y, Chan LC, Lim SO, et al. Metformin Promotes
891 Antitumor Immunity via Endoplasmic-Reticulum-Associated Degradation of PD-L1.
892 Mol Cell. 2018;71(4):606-20 e7.
893 54. Eikawa S, Nishida M, Mizukami S, Yamazaki C, Nakayama E, Udono H. Immune-
894 mediated antitumor effect by type 2 diabetes drug, metformin. Proc Natl Acad Sci U S
895 A. 2015;112(6):1809-14.
896 55. de Oliveira S, Houseright RA, Graves AL, Golenberg N, Korte BG, Miskolci V, et
897 al. Metformin modulates innate immune-mediated inflammation and early progression
898 of NAFLD-associated hepatocellular carcinoma in zebrafish. J Hepatol.
899 2019;70(4):710-21.
900 56. Dudek M, Pfister D, Donakonda S, Filpe P, Schneider A, Laschinger M, et al.
901 Auto-aggressive CXCR6(+) CD8 T cells cause liver immune pathology in NASH.
902 Nature. 2021;592(7854):444-9.
903 57. Pfister D, Nunez NG, Pinyol R, Govaere O, Pinter M, Szydlowska M, et al. NASH
904 limits anti-tumour surveillance in immunotherapy-treated HCC. Nature.
905 2021;592(7854):450-6.
906 58. Rohr-Udilova N, Klinglmuller F, Schulte-Hermann R, Stift J, Herac M, Salzmann
907 M, et al. Deviations of the immune cell landscape between healthy liver and
908 hepatocellular carcinoma. Sci Rep. 2018;8(1):6220.

59. Zhang Y, Wang H, Xiao H. Metformin Actions on the Liver: Protection Mechanisms Emerging in Hepatocytes and Immune Cells against NASH-Related HCC. Int J Mol Sci. 2021;22(9).
60. Donadon V, Balbi M, Mas MD, Casarin P, Zanette G. Metformin and reduced risk of hepatocellular carcinoma in diabetic patients with chronic liver disease. Liver Int. 2010;30(5):750-8.
61. Schulte L, Scheiner B, Voigtlander T, Koch S, Schweitzer N, Marhenke S, et al. Treatment with metformin is associated with a prolonged survival in patients with hepatocellular carcinoma. Liver Int. 2019;39(4):714-26.
62. Giovannucci E, Harlan DM, Archer MC, Bergenstal RM, Gapstur SM, Habel LA, et al. Diabetes and cancer: a consensus report. Diabetes Care. 2010;33(7):1674-85.
63. Itoh S, Yoshizumi T, Yugawa K, Imai D, Yoshiya S, Takeishi K, et al. Impact of Immune Response on Outcomes in Hepatocellular Carcinoma: Association With Vascular Formation. Hepatology. 2020;72(6):1987-99.
64. Sia D, Jiao Y, Martinez-Quetglas I, Kuchuk O, Villacorta-Martin C, Castro de Moura M, et al. Identification of an Immune-specific Class of Hepatocellular Carcinoma, Based on Molecular Features. Gastroenterology. 2017;153(3):812-26.
65. McLane LM, Abdel-Hakeem MS, Wherry EJ. CD8 T Cell Exhaustion During Chronic Viral Infection and Cancer. Annu Rev Immunol. 2019;37:457-95.
66. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. Nat Rev Immunol. 2015;15(8):486-99.
67. Han J, Han ML, Xing H, Li ZL, Yuan DY, Wu H, et al. Tissue and serum metabolomic phenotyping for diagnosis and prognosis of hepatocellular carcinoma. Int J Cancer. 2020;146(6):1741-53.
68. Devalaraja S, To TKJ, Folkert IW, Natesan R, Alam MZ, Li M, et al. Tumor-Derived Retinoic Acid Regulates Intratumoral Monocyte Differentiation to Promote Immune Suppression. Cell. 2020;180(6):1098-114 e16.
69. Ma J, Liu Y, Li Y, Gu J, Liu J, Tang J, et al. Differential role of all-trans retinoic acid in promoting the development of CD4+ and CD8+ regulatory T cells. J Leukoc Biol. 2014;95(2):275-83.
70. Yang H, Gu J, Zhu Q, Lu H, Wang K, Ni X, et al. Protection of acute GVHD by all-trans retinoic acid through suppression of T cell expansion and induction of regulatory T cells through IL-2 signaling. Int Immunopharmacol. 2015;28(2):911-6.

943 **Figure Legend**

944 **Figure 1. Characteristics of *Fah*$^{-/-}$ mice with chronic liver injury**

945 (A) RNA sequencing showing volcano plot of differential expression analysis results
946 from 4 weeks of chronic liver injury (C4) vs normal *Fah*$^{-/-}$ mouse liver (C0). (B) Heat
947 map illustrates differential gene expression of liver tissue from 4 weeks chronic liver
948 injury (C4, labeled as FT4 here) vs normal *Fah*$^{-/-}$ mouse (C0, labeled as FT0 here).
949 (C) Gene Ontology (GO) classification, and Kyoto Encyclopedia of Genes and
950 Genomes (KEGG) pathway enrichment analyses of the differentially expressed genes
951 from 4 weeks chronic liver injury (C4) vs normal *Fah*$^{-/-}$ mouse liver (C0). (D) Body
952 weight (BW) in *Fah*$^{-/-}$ mice with chronic liver injury (2W, 4W, 8W and 12W) and
953 normal *Fah*$^{-/-}$ mouse (0W). (E) Blood glucose levels in *Fah*$^{-/-}$ mice with chronic liver
954 injury (2W, 4W, 8W and 12W) and normal *Fah*$^{-/-}$ mouse (0W). (F) Western blot
955 analysis of AMPK activity from 4 weeks (C4), 12 weeks (C12) chronic liver injury
956 and normal *Fah*$^{-/-}$ mouse liver (C0). (G) Western blot analysis of P-AMPK level in
957 healthy human liver tissues and adjacent noncancerous liver tissues of HCC patients.
958 (H) Enrichment score of AMPK pathway in normal and premalignant tissues. The
959 enrichment score of AMPK pathway was calculated by ssGSEA algorithm. (I)
960 Enrichment score of AMPK pathway in normal tissues and NAFLD tissues. Normal:
961 non-tumor Normal control, FL: low Fibrosis, FH: high Fibrosis, CS: Cirrhosis, DL:
962 Dysplastic nodule Low, DH: Dysplastic Nodule high, NASH: non-alcoholic
963 steatohepatitis. All data represented as “mean ± SEM”. Statistical significance was

964 determined by unpaired two-tailed t-test. *P < 0.05; **P < 0.01; ***P < 0.001.

965 **Figure 2. Effects of metformin on the characteristics of precancerous livers and**

966 **HCC incidence in *Fah*$^{-/-}$ mice**

967 (A) Schematic diagram showing the experimental set. Representative photographs of

968 livers about chronic liver injury for 12 weeks without metformin (C12) and with

969 metformin (Met12) are showing. (B) Western blot analysis of P-AMPK and AMPK

970 protein level in liver tissues of C12 and Met12. (C) Graphs representing tumor

971 incidence of *Fah*$^{-/-}$ mice with and without metformin (Met12 and C12). (D-E) Scatter

972 plots displaying the tumor numbers (D) and size of tumors (E) in Fah-deficient livers

973 at C12 and with Met12. (F) Serological indexes of liver injury (ALT, AST and

974 AST/ALT) are almost not changed in the serum of *Fah*$^{-/-}$ mice at C12 and Met12. *P

975 < 0.05; **P < 0.01; ***P < 0.001; NS: not significant.

976 **Figure 3. Effects of metformin on *Cyp26a1* Gene in mouse hepatocyte**

977 **(**A) Differentially up- or down- regulated genes between C12 and Met12. (Log2 Fold

978 Change >1, FDR<0.05). (B) The intersection of our sequencing results and

979 GSE110524. (C) RNA-seq data of *Cyp26a1* gene. (D) Quantitative analysis of

980 *Cyp26a1* gene in liver tissues from C12 and Met12 was performed by qRT-PCR. (E)

981 Western blot analysis of relative protein in liver tissues from C12 and Met12. (F)

982 qRT-PCR analysis of *Cyp26a1* gene expression between mouse hepatocyte and non-

983 parenchymal cells. perfused hepatocyte and non-parenchymal cells were separated by

984 low-speed centrifugation. (G) Perfused hepatocyte treated with metformin in vitro.

985 qRT-PCR analysis results of *Cyp26a1* gene expression with different concentrations
986 of metformin. (H) Western blot analysis of *Cyp26a1* gene expression in perfused
987 adherent hepatocyte with different concentrations of metformin *in vitro*. (I) RARE
988 luciferase assay on *Fah*$^{-/-}$ hepatocyte cultured with increasing amounts of atRA. Data
989 representative of 3 independent experiments. Two-way ANOVA was used. (J) RARE
990 luciferase assay on *Fah*$^{-/-}$ hepatocyte cultured with increasing amounts of metformin.
991 Data representative of 3 independent experiments. Two-way ANOVA was used. Data
992 are represented as "mean ± SEM". *P < 0.05; **P < 0.01; ***P < 0.001; NS: not
993 significant.

994 **Figure 4. Metformin suppressed *Cyp26a1* gene by inhibiting JNK/c-Jun pathway**
995 (A) qRT-PCR analysis results of *Cyp26a1* gene expression of primary hepatocytes
996 with different concentrations of AICAR. (B) Western blot analysis of relative gene
997 expression in perfused adherent hepatocyte with different concentrations of
998 metformin and AICAR *in vitro*. (C) qRT-PCR analysis results of *Cyp26a1* gene
999 expression. Primary hepatocytes treated with AMPK inhibitor Compound C (80 μM)
1000 in the absence or presence of metformin for 6 hours. Treatment with the inhibitor
1001 started 1 h before metformin treatment. (D) Western blot analysis of relative gene
1002 expression. Primary hepatocytes treated with AMPK inhibitor in the absence or
1003 presence of metformin for 6 hours. Treatment with the inhibitor started 1 h before
1004 metformin treatment. (E) We transfected plasmid (pGL3-cyp26a1-promoter-2000bp-
1005 luciferase) into primary hepatocyte cultured with increasing amounts of metformin.

Data representative of 3 independent experiments. Two-way ANOVA was used. (F) The truncated mouse *Cyp26a1* promoter mutation plasmids were transfected into primary hepatocyte cultured with increasing amounts of metformin, separately. Data representative of 3 independent experiments. Two-way ANOVA was used. (G) qRT-PCR analysis results of *Cyp26a1* gene expression. Primary hepatocytes treated with JNK inhibitor SP 600125 (100 μM) with increasing amounts of metformin for 6 hours. Treatment with the inhibitor started 1 h before metformin treatment. (H) Western blot analysis of relative gene expression. Primary hepatocytes treated with JNK inhibitor SP600125 (100 μM) with increasing amounts of metformin for 6 hours. Treatment with the inhibitors started 1 h before metformin treatment. *P < 0.05; **P < 0.01; ***P < 0.001; NS: not significant.

Figure 5. Effects of atRA on HCC formation in *Fah*$^{-/-}$ mice with chronic liver injury

(A) Schematic diagram showing the experimental set. Representative photographs of livers about chronic liver injury for 12 weeks without atRA (C12) and with atRA treatment (atRA12) are showing. (B) Graphs representing tumor incidence of *Fah*$^{-/-}$ mice with and without atRA treatment. (C) Scatter plots displaying the tumor numbers of *Fah*$^{-/-}$ mice with and without atRA treatment. (D) Scatter plots displaying the tumor size in *Fah*$^{-/-}$ mice between C12 group and atRA12 group. (E) Indicators of hepatocyte injury (ALT, AST and AST/ALT) are almost not changed in the serum of *Fah*$^{-/-}$ mice at C12 and atRA12 groups. *P < 0.05; **P < 0.01; ***P < 0.001; NS: not

1027 significant.

1028 **Figure 6. Effects of Metformin and *at*RA on the characteristics of immune cells**

1029 **in *Fah*$^{-/-}$mice**

1030 (A) Single immune cells were stained with anti-CD3, CD8, CD45, CD4, CD49B,

1031 F4/80, CD11B and Live/Dead antibodies. The phenotype of liver immune cells was

1032 analyzed by FACS. (B) Liver infiltrating immune cells were measured treated with or

1033 without metformin. Data represent "mean ± SEM" of four pooled experiments.

1034 Statistical significance was determined by unpaired two-tail t-test. (C) Representative

1035 pictures of the indicated H&E staining and immunohistochemistry staining. Scale bar,

1036 50μm. (D) Quantification of hepatocyte cell number, $CD45^+$ cell number and $CD8^+$

1037 cell number per square millimeter. (E) Quantitative analysis of pro-tumor cytokines of

1038 liver tissues. (F) Liver infiltrating immune cells were measured treated with or

1039 without atRA. Data represent "mean ± SEM" of four pooled experiments. Statistical

1040 significance was determined by unpaired two-tail t-test. (G) Representative pictures of

1041 the indicated H&E staining and immunohistochemistry staining. Scale bar, 50μm. (H)

1042 Quantification of $CD45^+$ cell number and $CD8^+$ cell number per square millimeter in

1043 the liver. (I) qRT-PCR analysis of pro-tumor cytokines of liver tissues. *P < 0.05; **P

1044 < 0.01; ***P < 0.001; NS: not significant.

1045 **Figure 7. Single cell RNA-sequencing analysis reveals the expression**

1046 **characteristics of $CD8^+$ T cells in *Fah*$^{-/-}$ mouse with progressive chronic liver**

1047 **injury**

(A) T-distributed stochastic neighbor embedding (t-SNE) plot of immune cells of *Fah*$^{-/-}$ mouse at 0, 12, and 18 weeks under CLI. (B) Histogram indicating the proportion of the CD8^{+} T cell at 0, 12, and 18 weeks. (C) Violin plot presenting the expression levels of Cd8a at 0, 12, and 18 weeks. (D) Expression proportion of Tnf, Ltb, Pdcd1, Cxcr6, Il1b, Il6, Rara and Rarb in CD8^{+} T cell and other immune types group at 0, 12, and 18 weeks. (E) Violin plot presenting the expression levels of Rara at 0, 12, and 18 weeks. (F) Uniform manifold approximation and projection (UMAP) plot showing the expression levels of Rara. (G) Correlation analysis of Cd8a and Rara in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green).

Figure 8. CD8^{+} T cells have different role in HCC and HCC adjacent tumor based on their molecular and functional characteristics

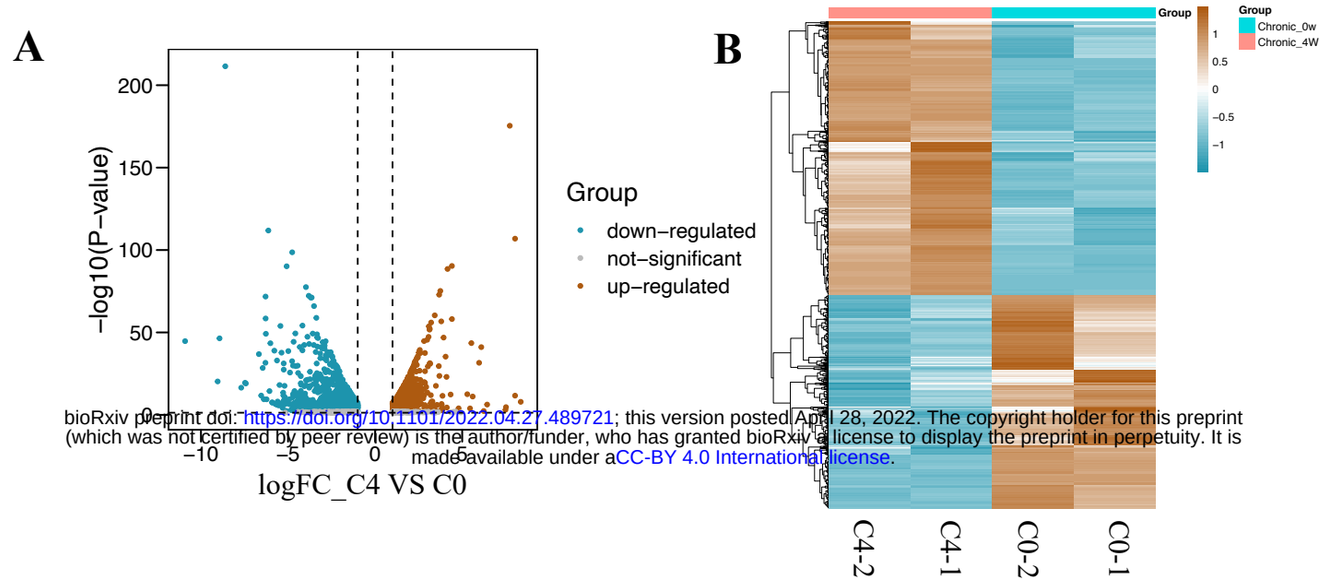
(A) Violin plot presenting the expression levels of CD8A in GTEx normal, TCGA para-tumor and TCGA tumors tissues. (B) Boxplot presenting the expression levels of CD8A in GSE89632 HC, SS and NASH tissues. (C) Violin plot presenting the expression levels of TNF and LTB in GTEx normal, TCGA para-tumor and TCGA tumors tissues. (D) Correlation analysis of CD8A and LTB or TNF in 30 GTEx normal, 17 TCGA para-tumor and 34 TCGA tumors tissue types. Red color indicates the Liver organ in GTEx or HCC para-tumor tissues, and HCC tumor in TCGA tumor tissues. (E) Kaplan-Meier survival analysis of overall survival (OS) (left) or disease-free survival (DFS) for CD8A transcript level of HCC patients from TCGA database. (F) The schematic diagram illustrating the mechanism of appropriately metformin-

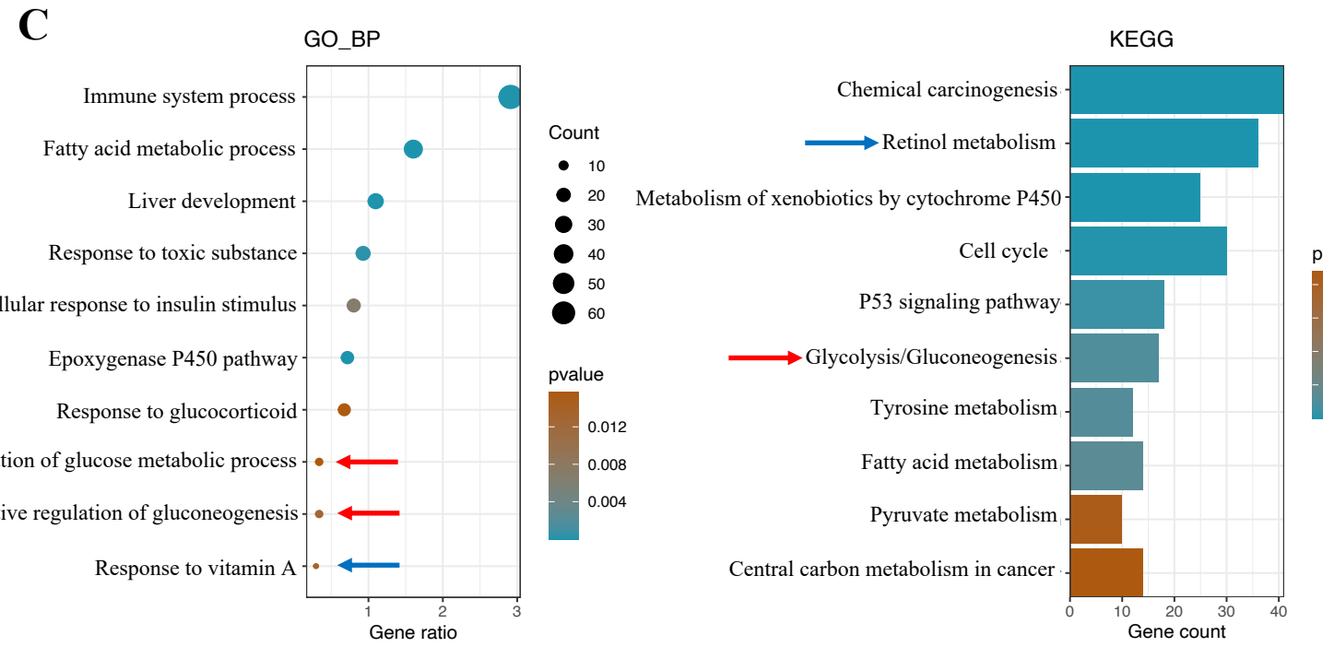
1069 downregulated tumor promotion in hepatocarcinogenesis. The icons of cell types were

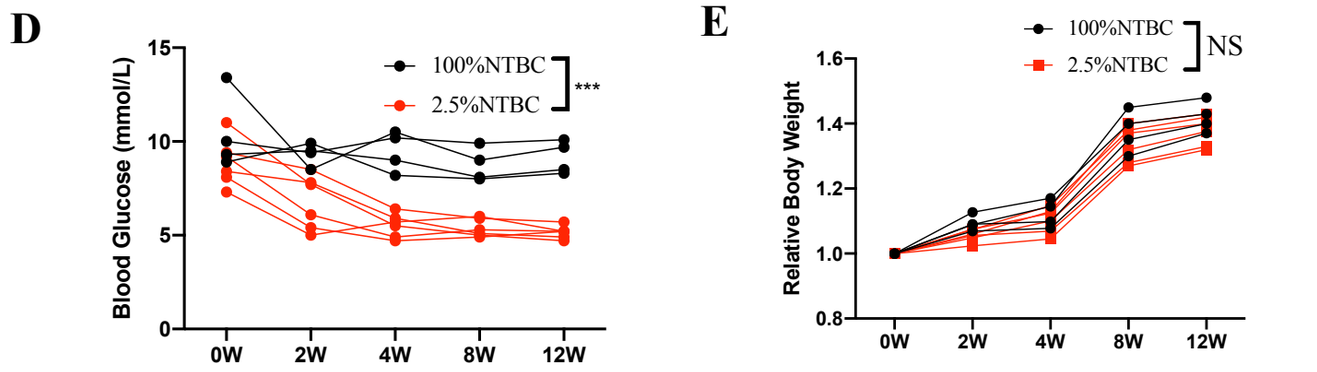
1070 obtained from SERVIER MEDICAL ART (https://smart.servier.com/).

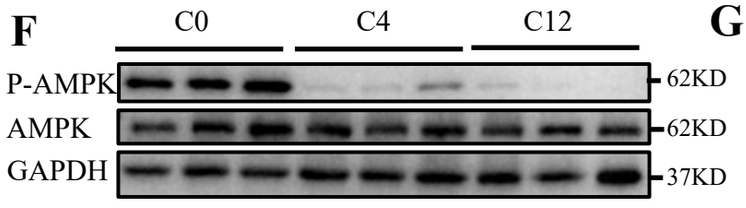
1071

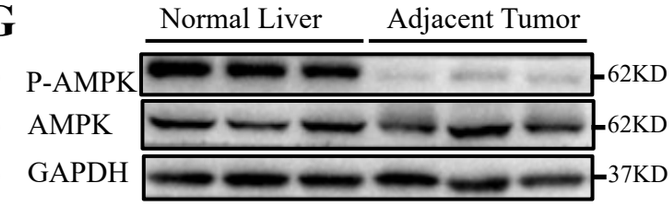
Figure 1

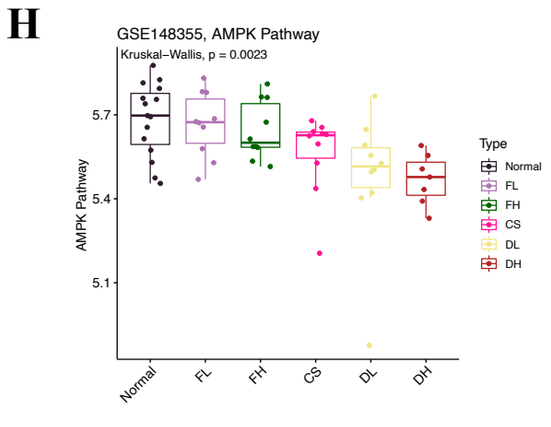












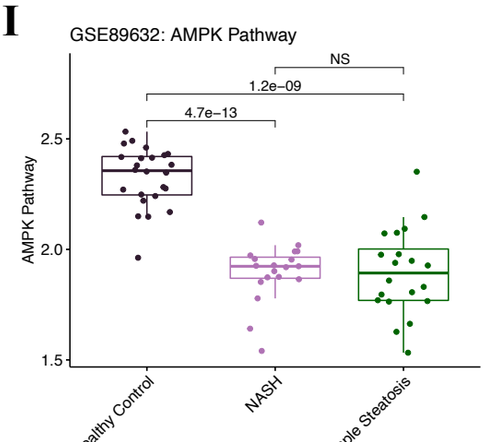
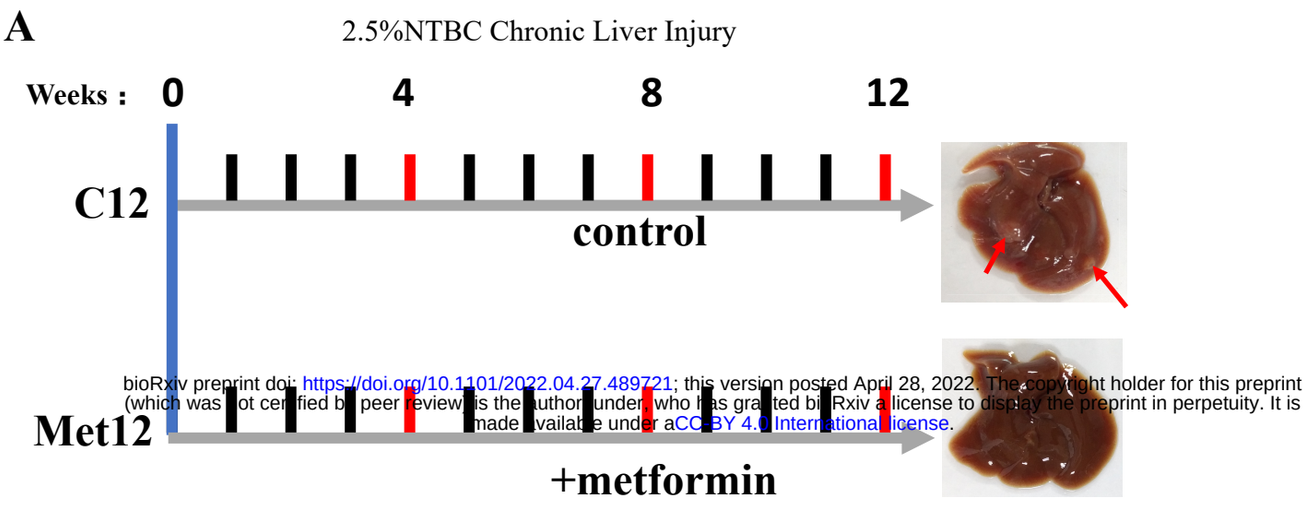
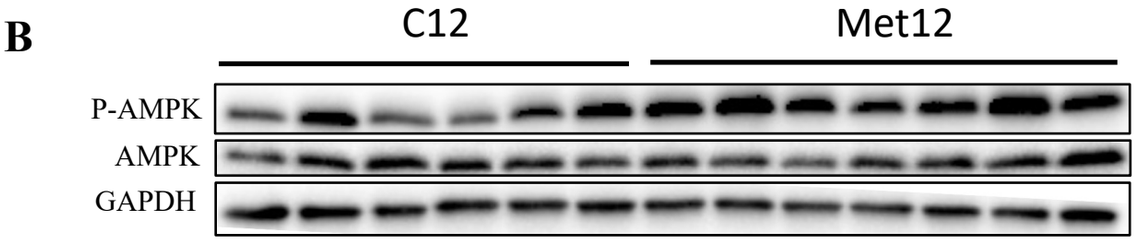
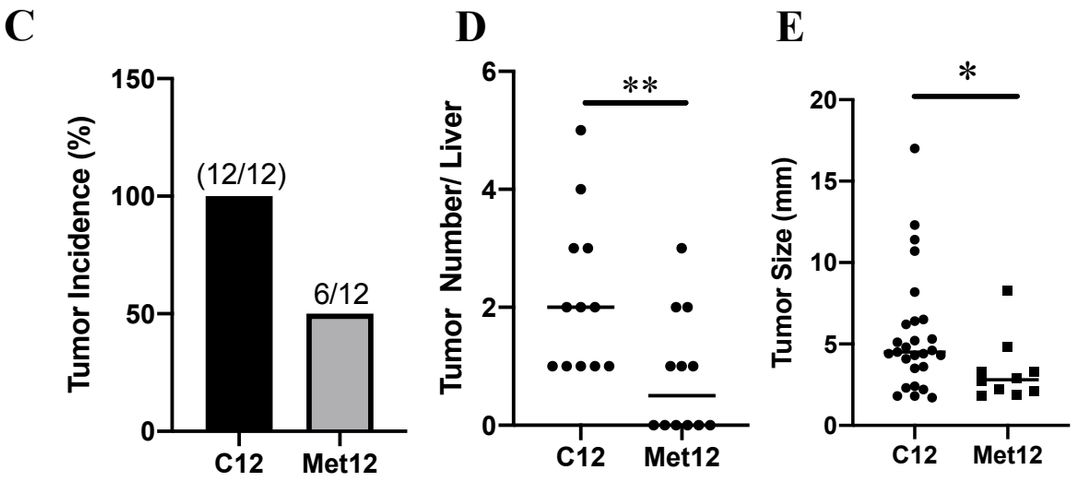


Figure 2







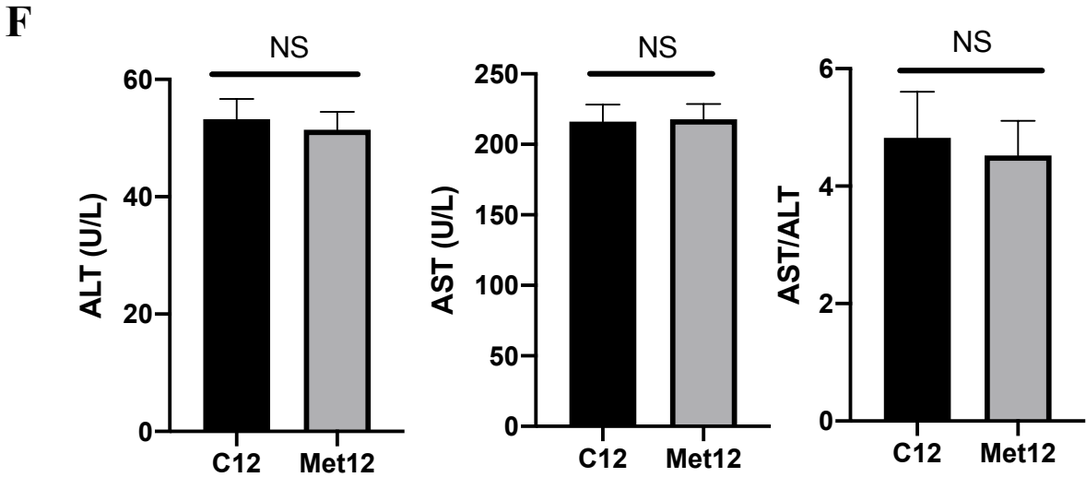
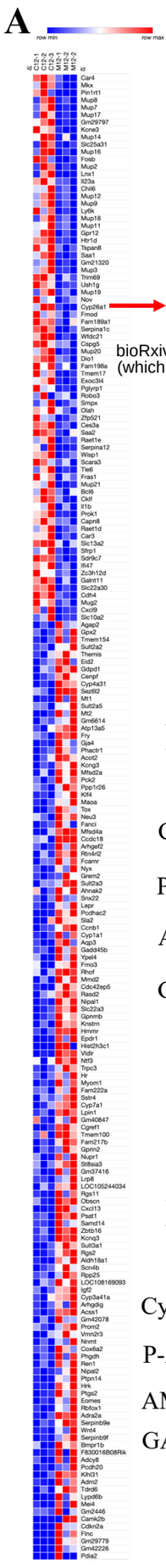
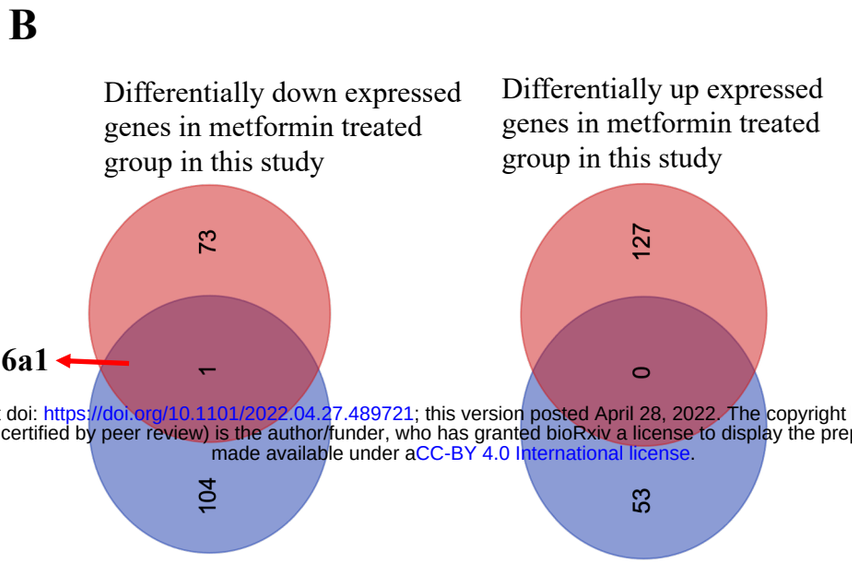
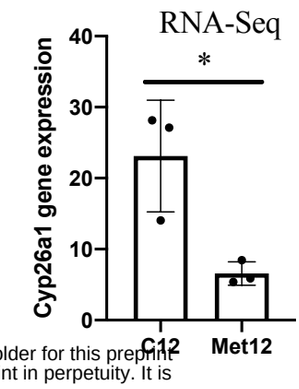
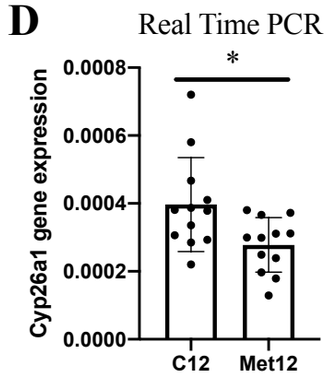


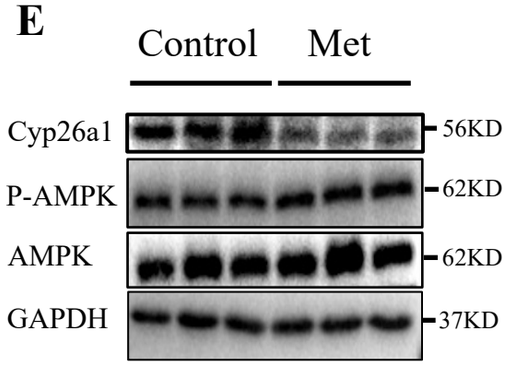
Figure 3

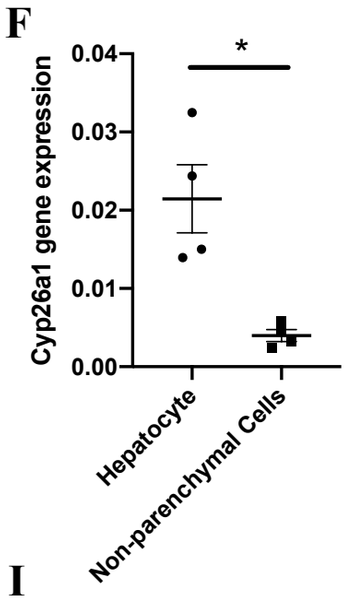


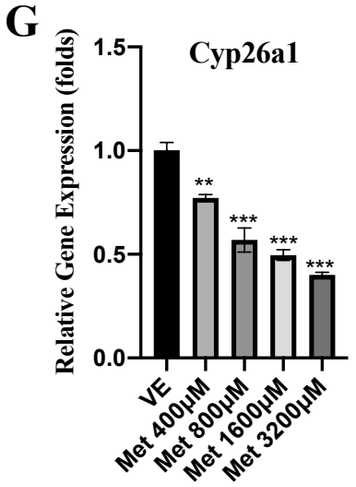


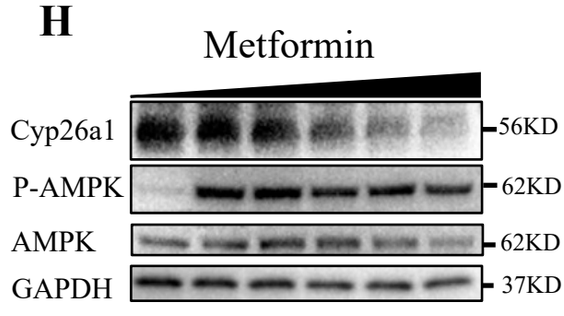


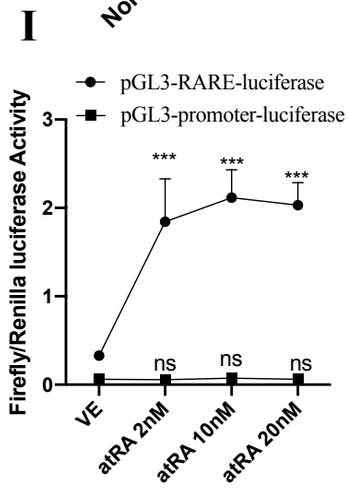












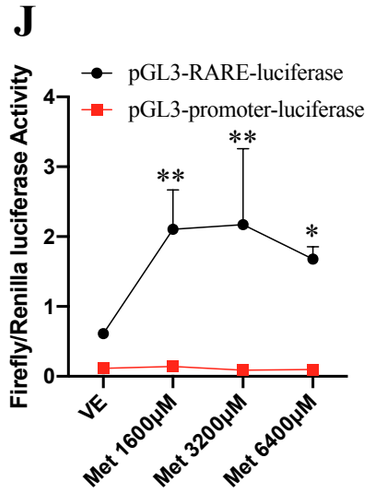


Figure 4

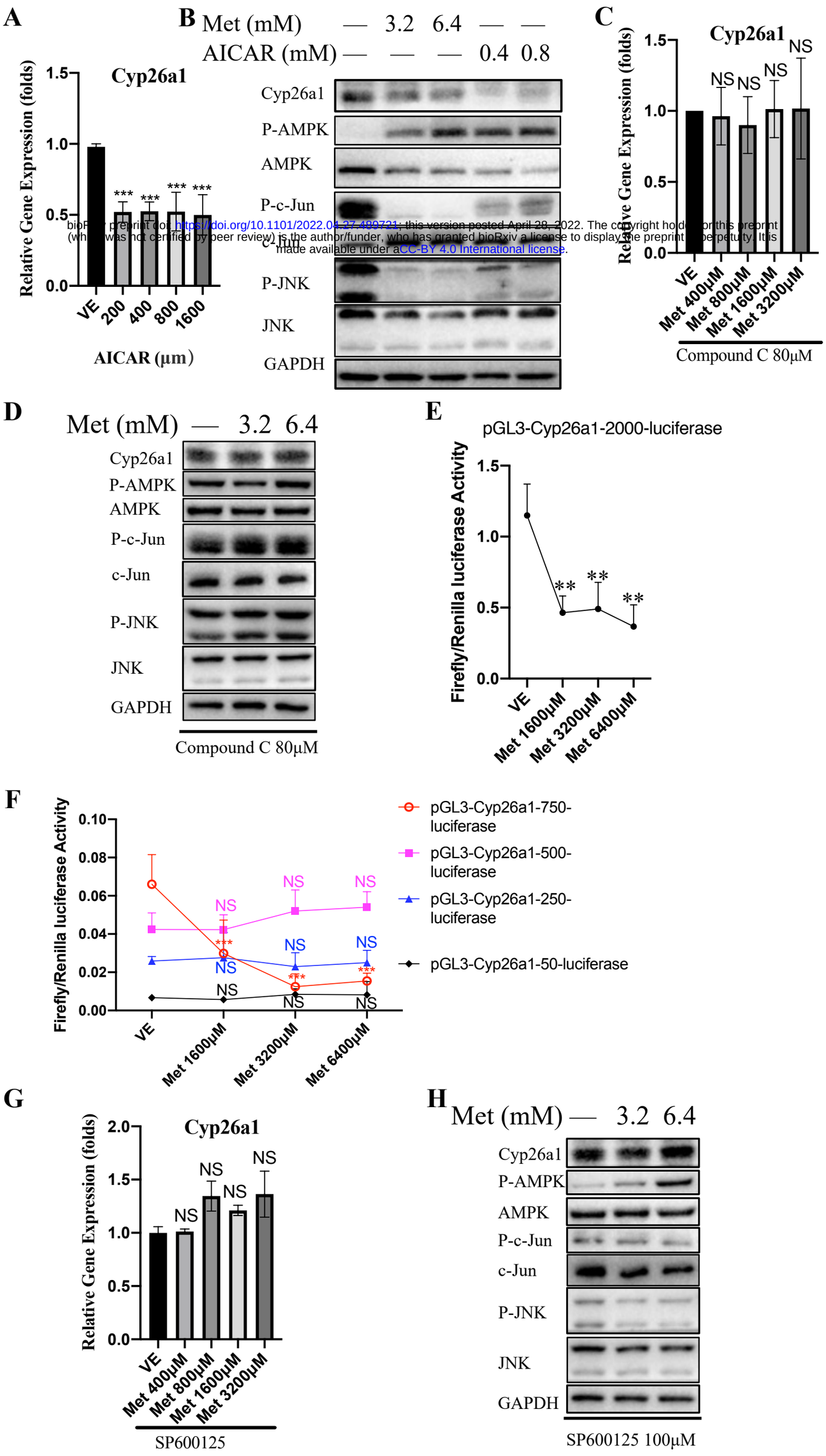
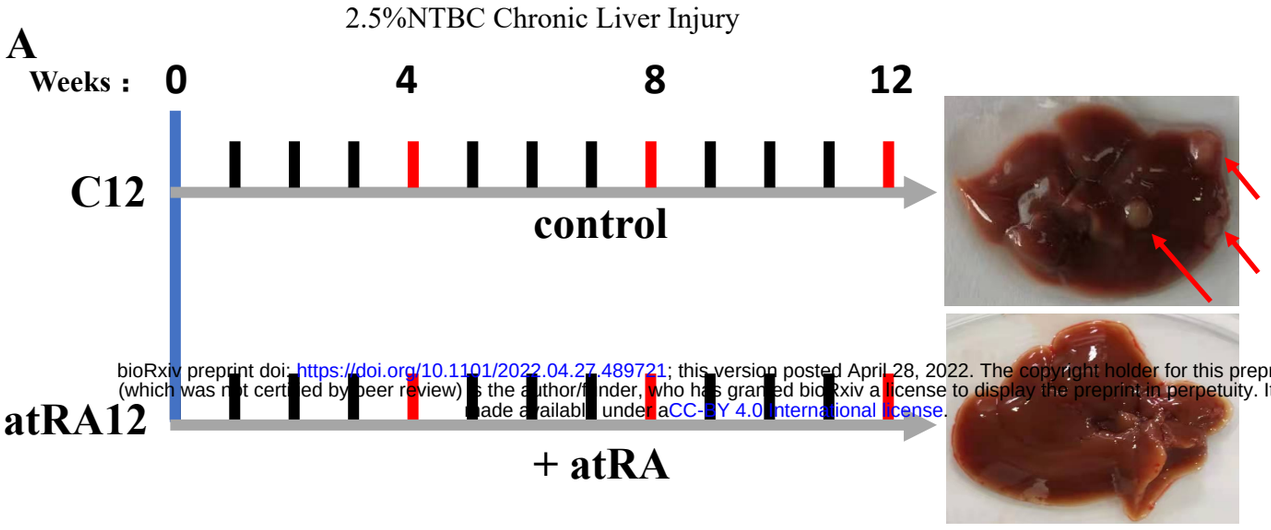


Figure 5



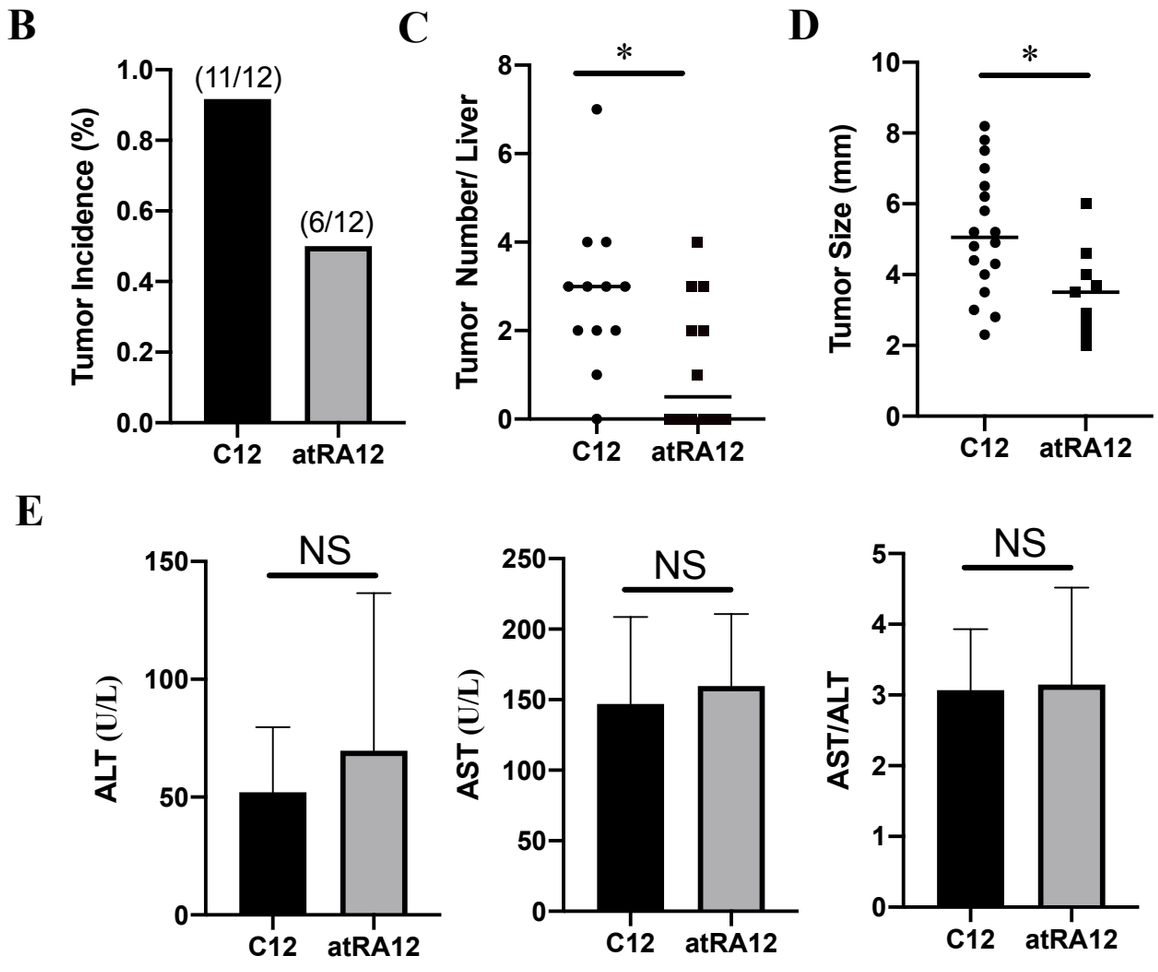
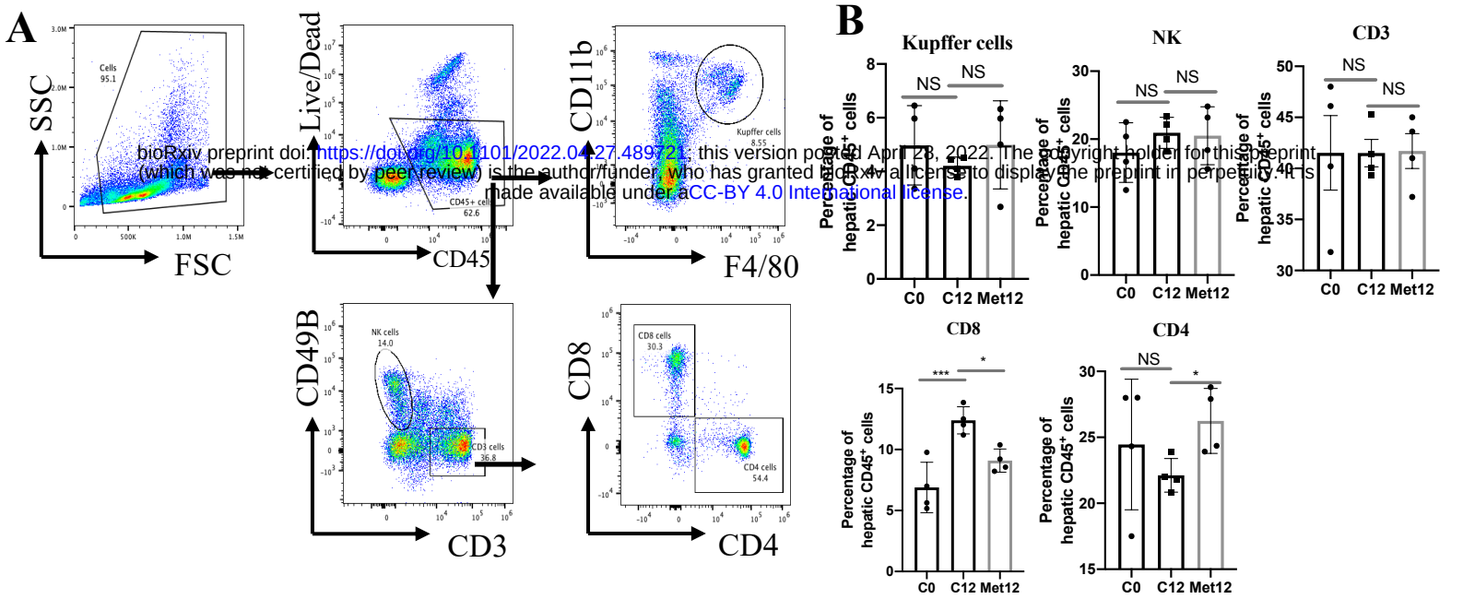
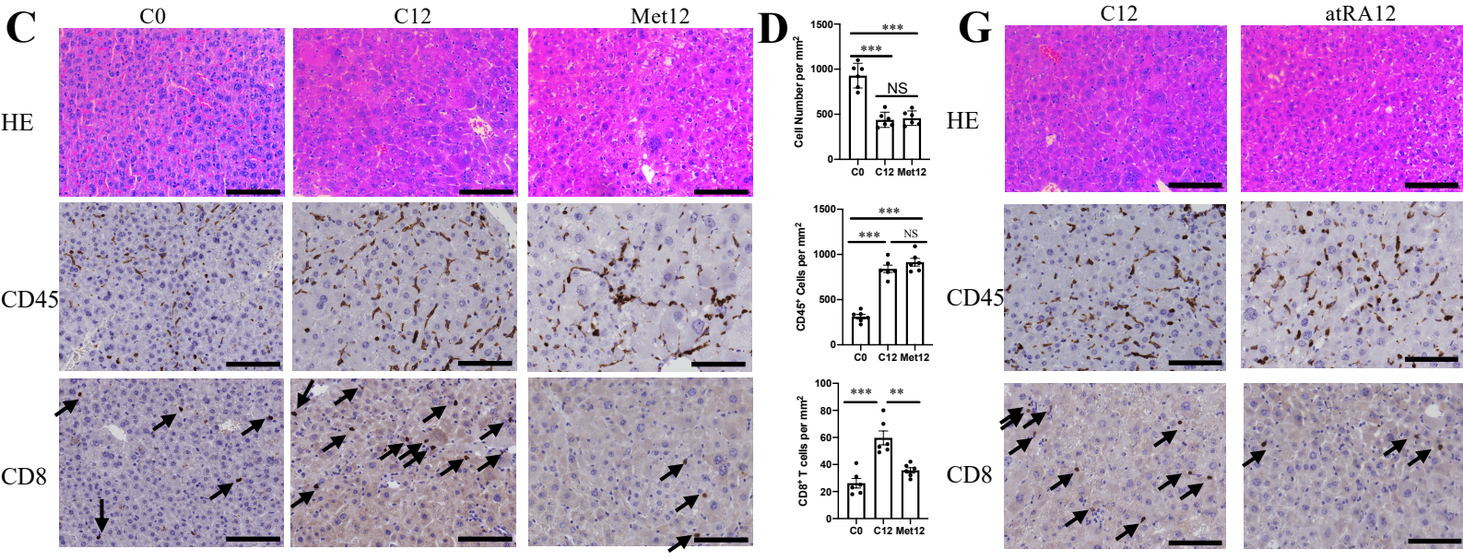


Figure 6





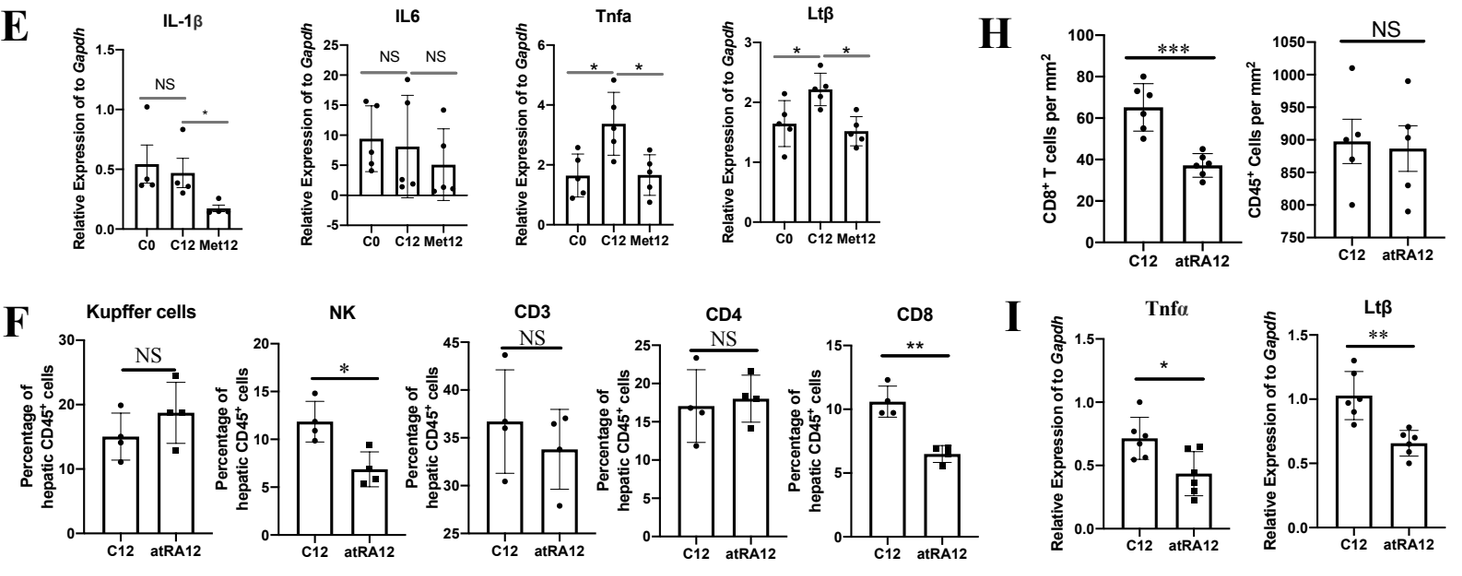


Figure 7

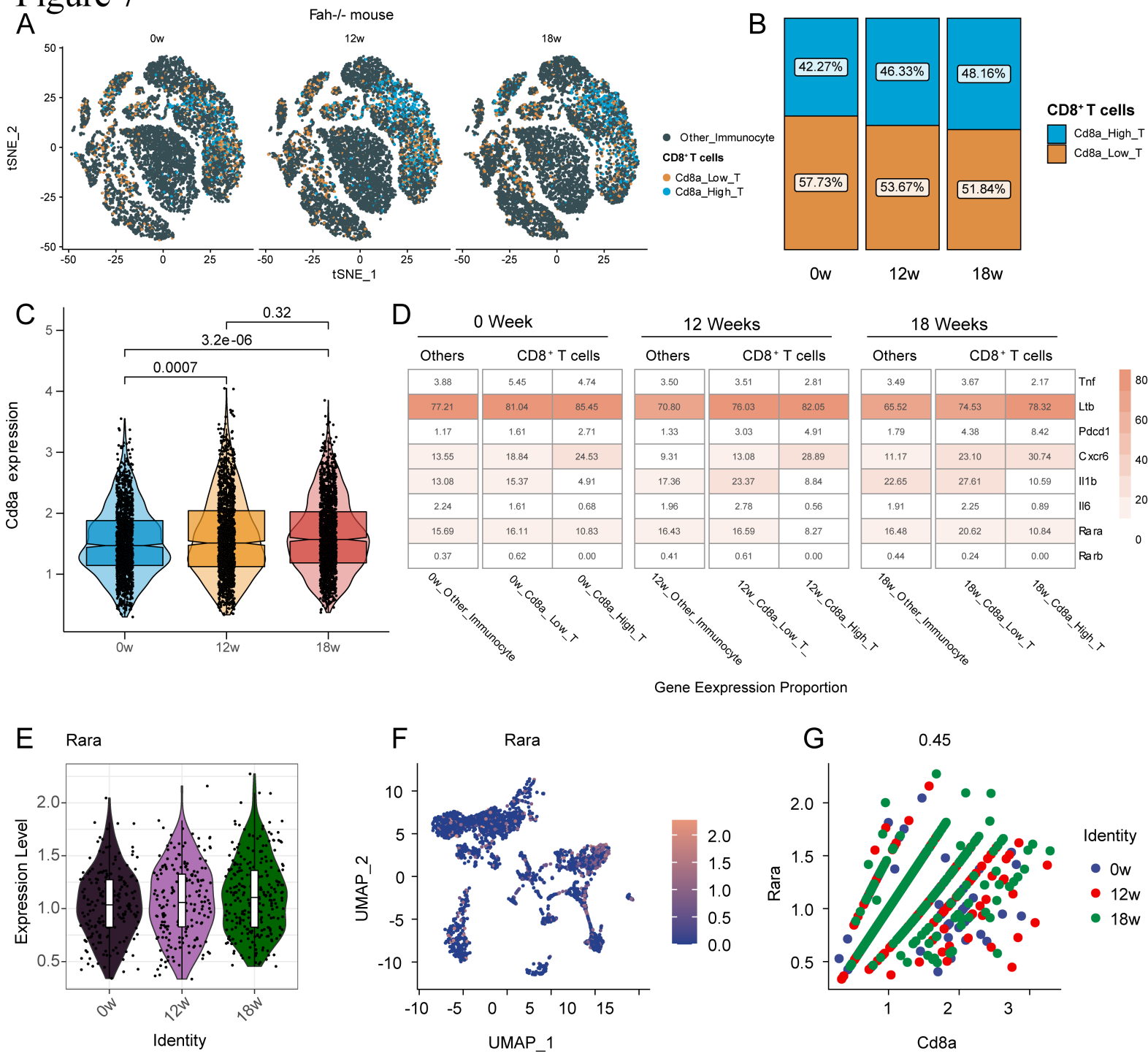
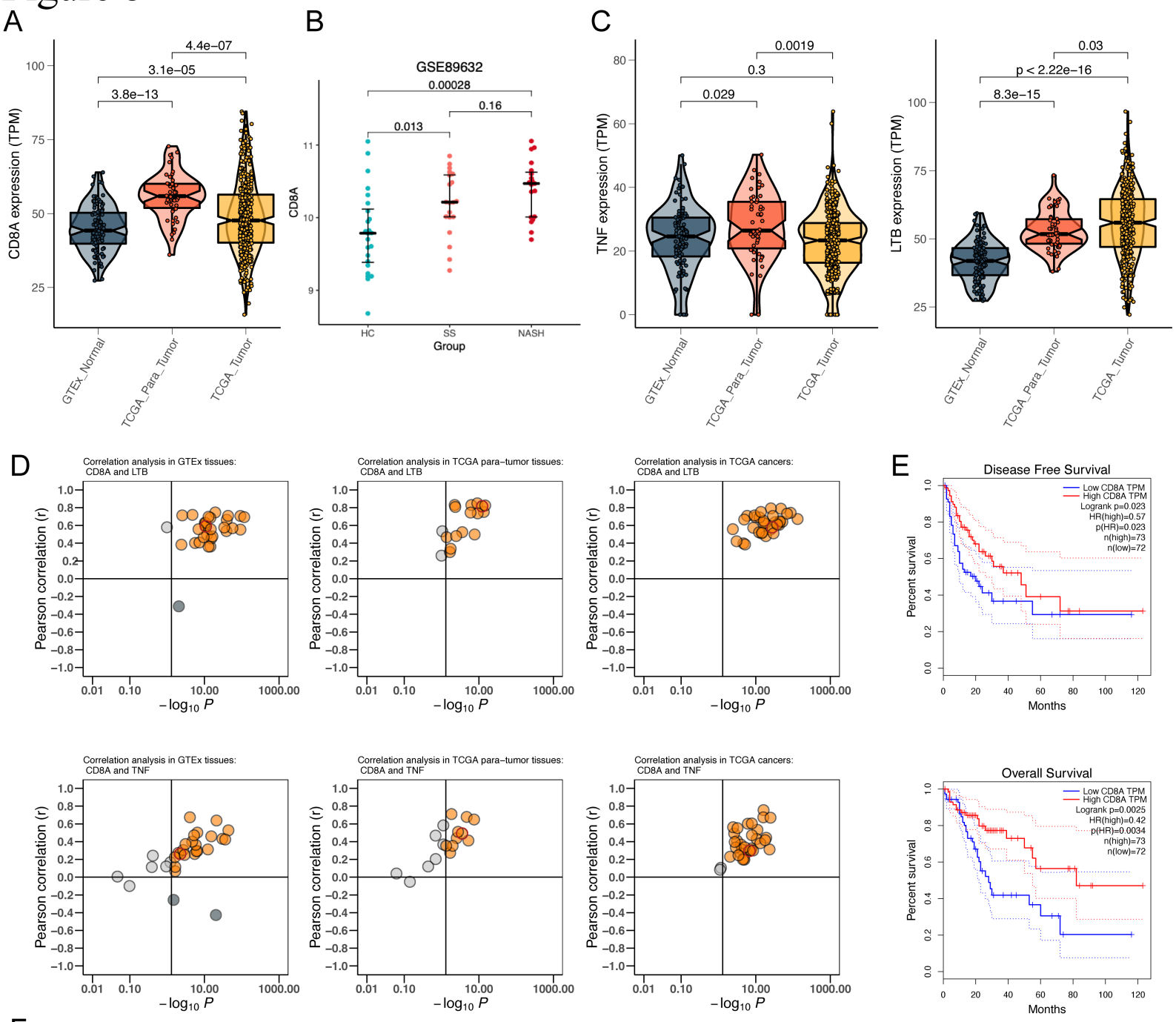


Figure 8



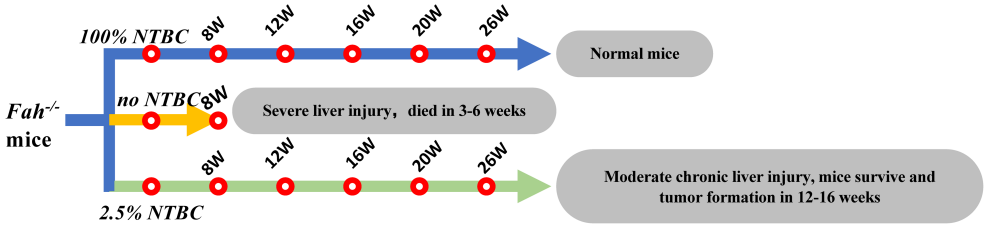


Figure S1. characteristics of *Fah*$^{-/-}$ mice. under 100% NTBC, the liver of Fah$^{-/-}$ mice was preserved normally; Fah$^{-/-}$ mice developed acute liver injury and died at 3-6 weeks without any NTBC; Fah$^{-/-}$ mouse can survive under 2.5% NTBC, but experience chronic liver injury and form liver cancer after 12 weeks.

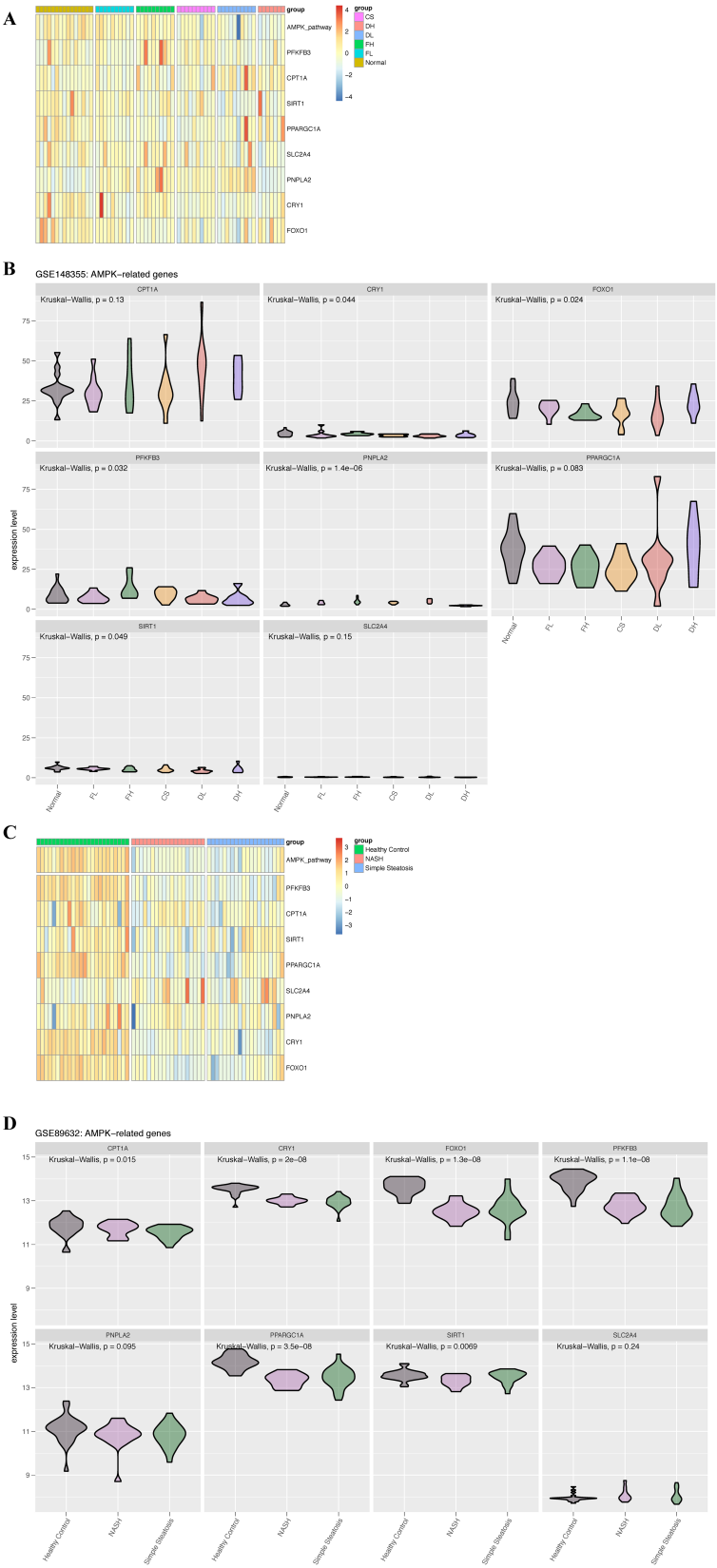


Figure S2. AMPK pathway was suppressed in chronic liver injury tissues compare with normal tissues.

(A)Heatmap of AMPK pathway and related genes in normal and premalignant tissues. (B)Expression level of AMPK pathway related genes in normal and premalignant tissues (GSE148355 dataset). (C) Heatmap of AMPK pathway and related genes in normal tissues and NAFLD tissues. (D) Expression level of AMPK pathway related genes in normal tissues and NAFLD tissues (GSE89632 dataset). Normal: non-tumor Normal control, FL: low Fibrosis, FH: high Fibrosis, CS: Cirrhosis (CS), DL: Dysplastic nodule Low, DH: Dysplastic Nodule high, NASH: nonalcoholic steatohepatitis

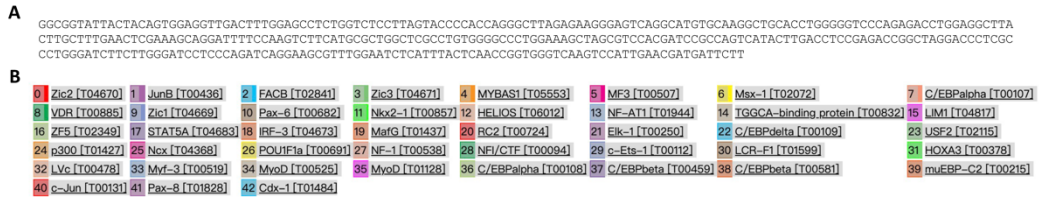


Figure S3. Predict transcription factors that bind in the promoter region of nucleotide -750 to -500 relation to transcription start site of mouse cyp26a1 gene.

(A)The detail sequence of nucleotide -750 to -500 relation to transcription start site of mouse cyp26a1 gene.

(B) All the predicted transcription factors binding to nucleotide -750 to -500 relation to transcription start site of mouse cyp26a1 gene using PROMO database (http://alggen.lsi.upc.es/cgibin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3)

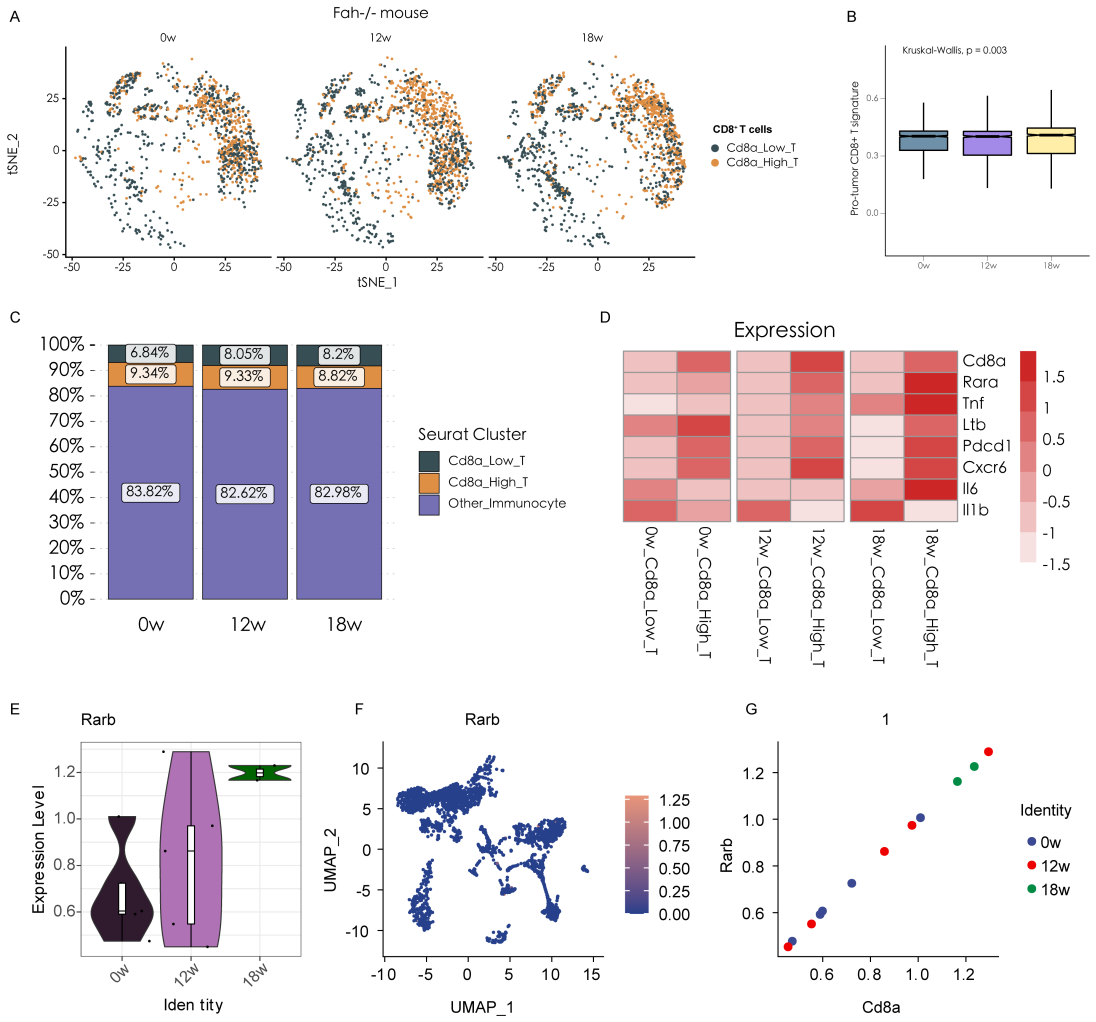


Figure S4. The expression levels of CD8^{+} T cells in *Fah*$^{-/-}$ mouse at 0, 12, and 18 weeks

(A) t-SNE visualization of Cd8^{High} T cell (grey) and Cd8^{Low} T cell group (orange) in *Fah*$^{-/-}$ mouse at 0, 12, and 18 weeks. (B) Expression levels of tumor promoting CD8^{+} T cell signature in *Fah*$^{-/-}$ mouse at 0, 12, and 18 weeks. (C) Histogram indicating the proportion of cells in the Cd8^{High} T cell (grey), Cd8^{Low} T cell (orange) and other immune types group (pruple) at 0, 12, and 18 weeks. (D) Heatmap presenting the expression levels of Cd8a, Rara, Tnf, Ltb, Pdcd1, Cxcr6, Il6 and Il1b in the Cd8^{High} T cell and Cd8^{Low} T cell group at 0, 12, and 18 weeks. (E) Violin plot presenting the expression levels of Rarb at 0, 12, and 18 weeks. (F) UMAP plot showing the expression levels of Rarb. (G) Correlation analysis of Cd8a and Rarb in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green).

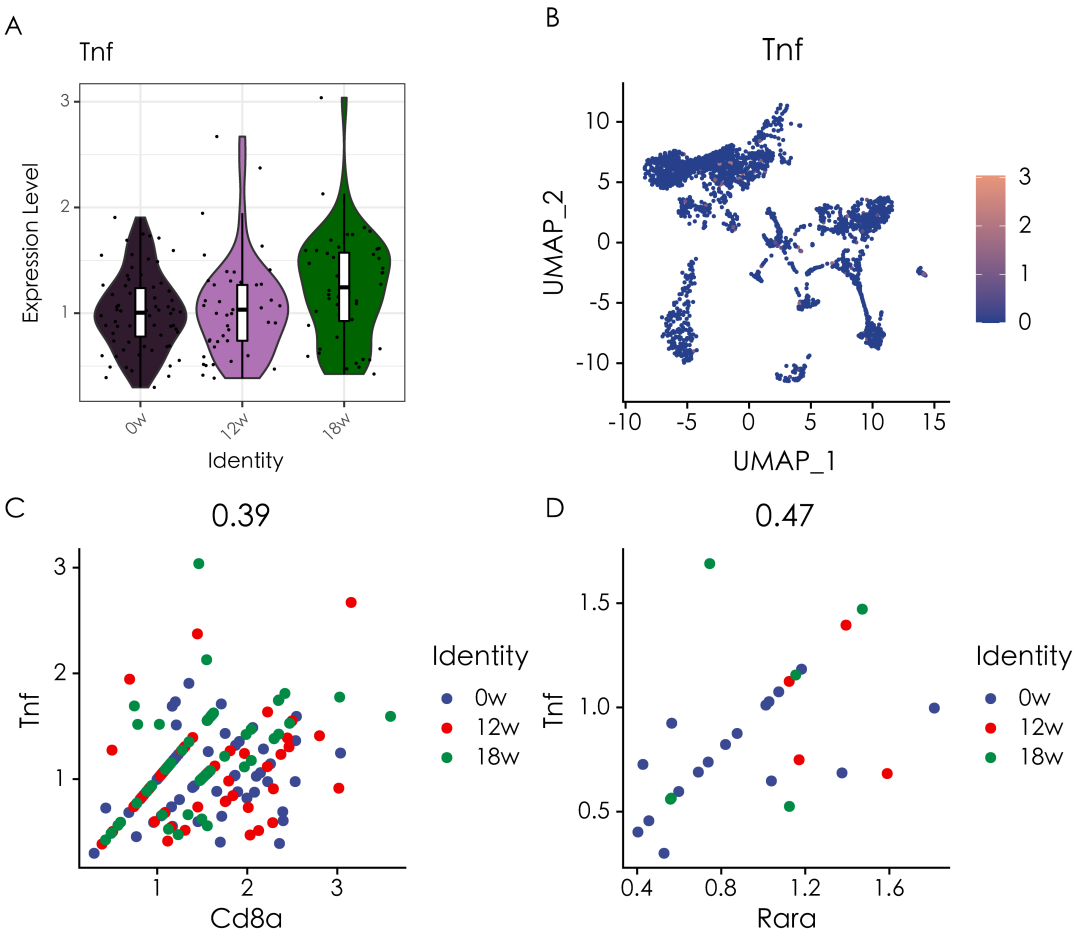


Figure S5. The expression levels of Tnf and its correlation with Cd8a and Rara

(A) Violin plot presenting the expression levels of Tnf at 0, 12, and 18 weeks. (B) UMAP plot showing the expression levels of Tnf. (C) Correlation analysis of Cd8a and Tnf in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green). (D) Correlation analysis of Rara and Tnf in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green).

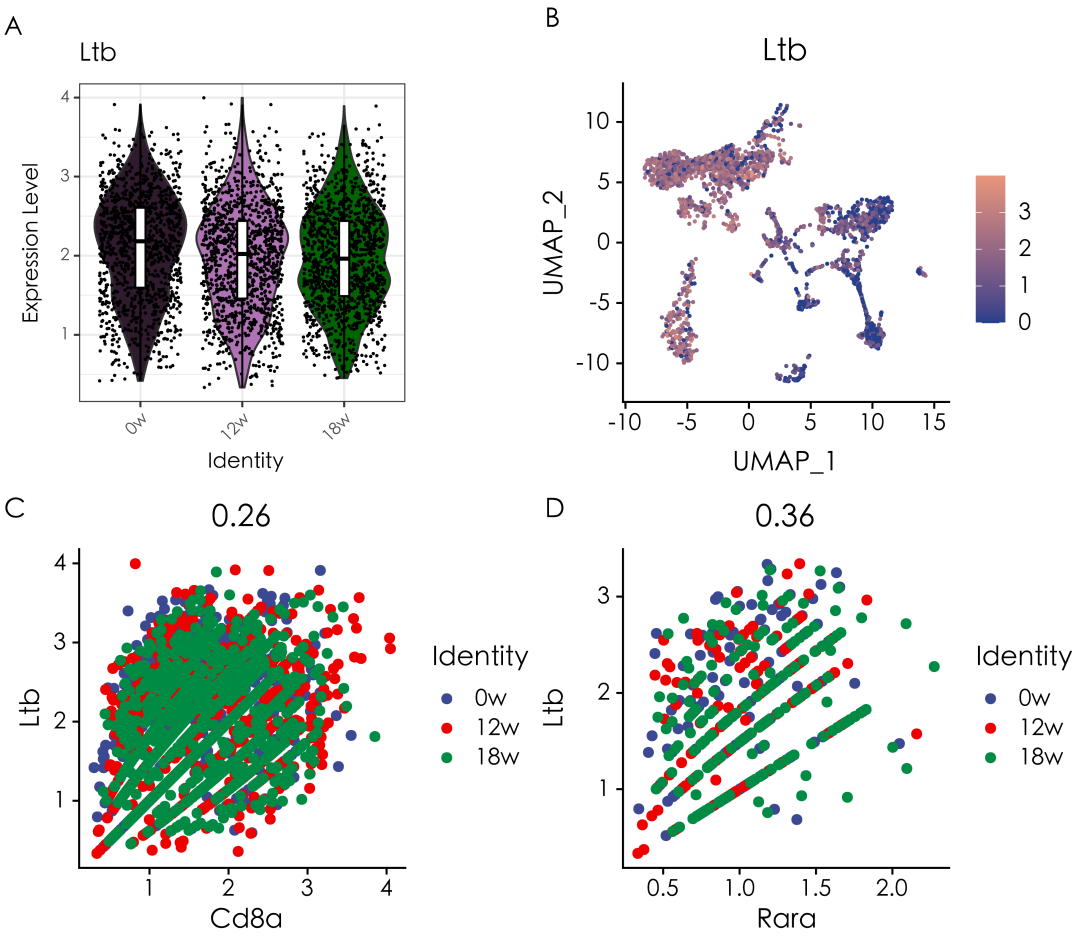


Figure S6. The expression levels of Ltb and its correlation with Cd8a and Rara

(A) Violin plot presenting the expression levels of Ltb at 0, 12, and 18 weeks. (B) UMAP plot showing the expression levels of Ltb. (C) Correlation analysis of Cd8a and Ltb in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green). (D) Correlation analysis of Rara and Ltb in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green).

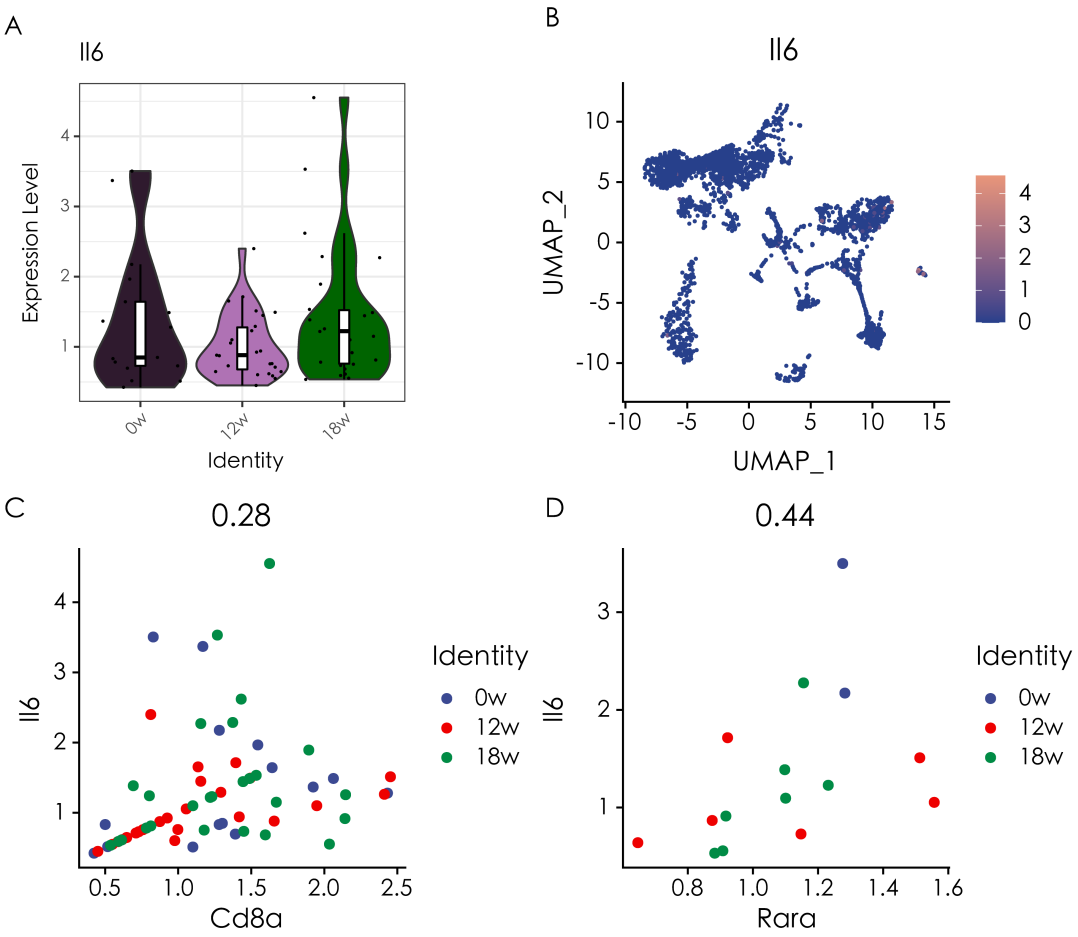


Figure S7. The expression levels of Il6 and its correlation with Cd8a and Rara

(A) Violin plot presenting the expression levels of Il6 at 0, 12, and 18 weeks. (B) UMAP plot showing the expression levels of Il6. (C) Correlation analysis of Cd8a and Il6 in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green). (D) Correlation analysis of Rara and Il6 in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green).

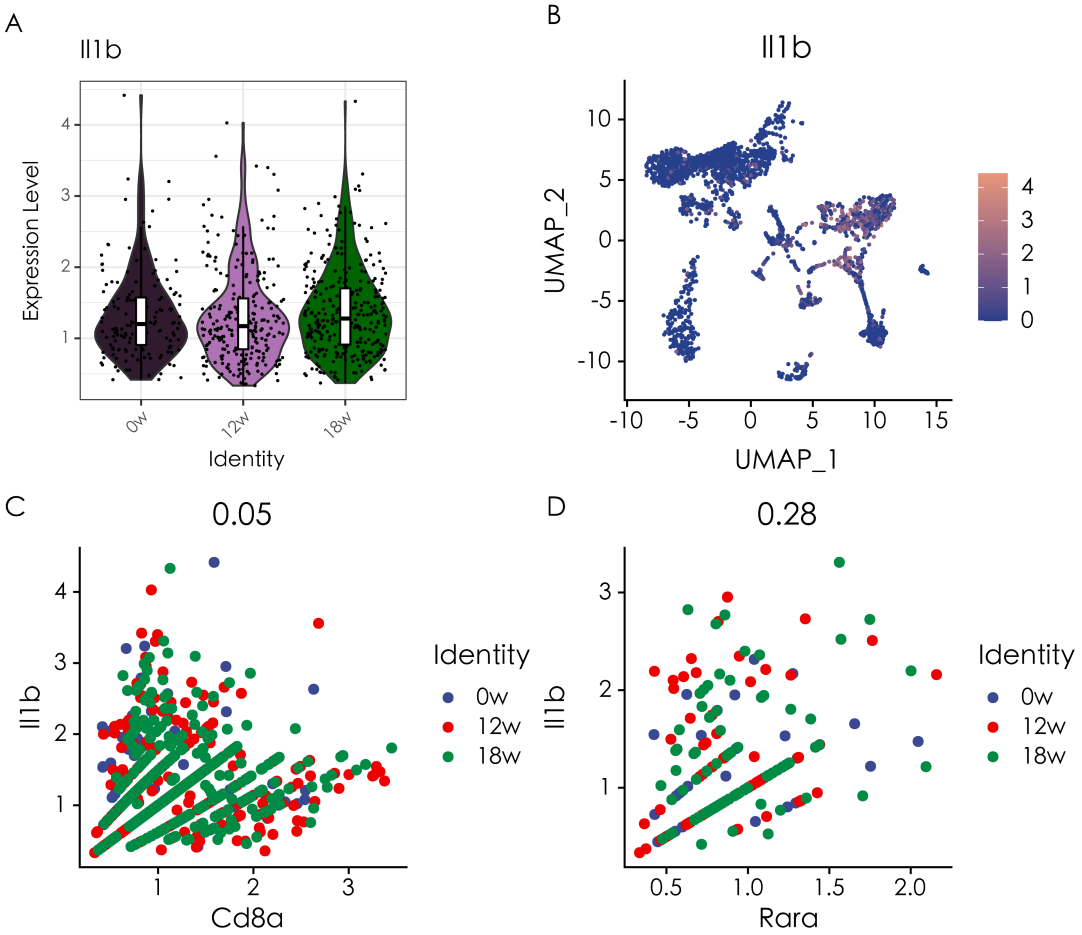


Figure S8. The expression levels of Il1b and its correlation with Cd8a and Rara

(A) Violin plot presenting the expression levels of Il1b at 0, 12, and 18 weeks. (B) UMAP plot showing the expression levels of Il1b. (C) Correlation analysis of Cd8a and Il1b in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green). (D) Correlation analysis of Rara and Il1b in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green).

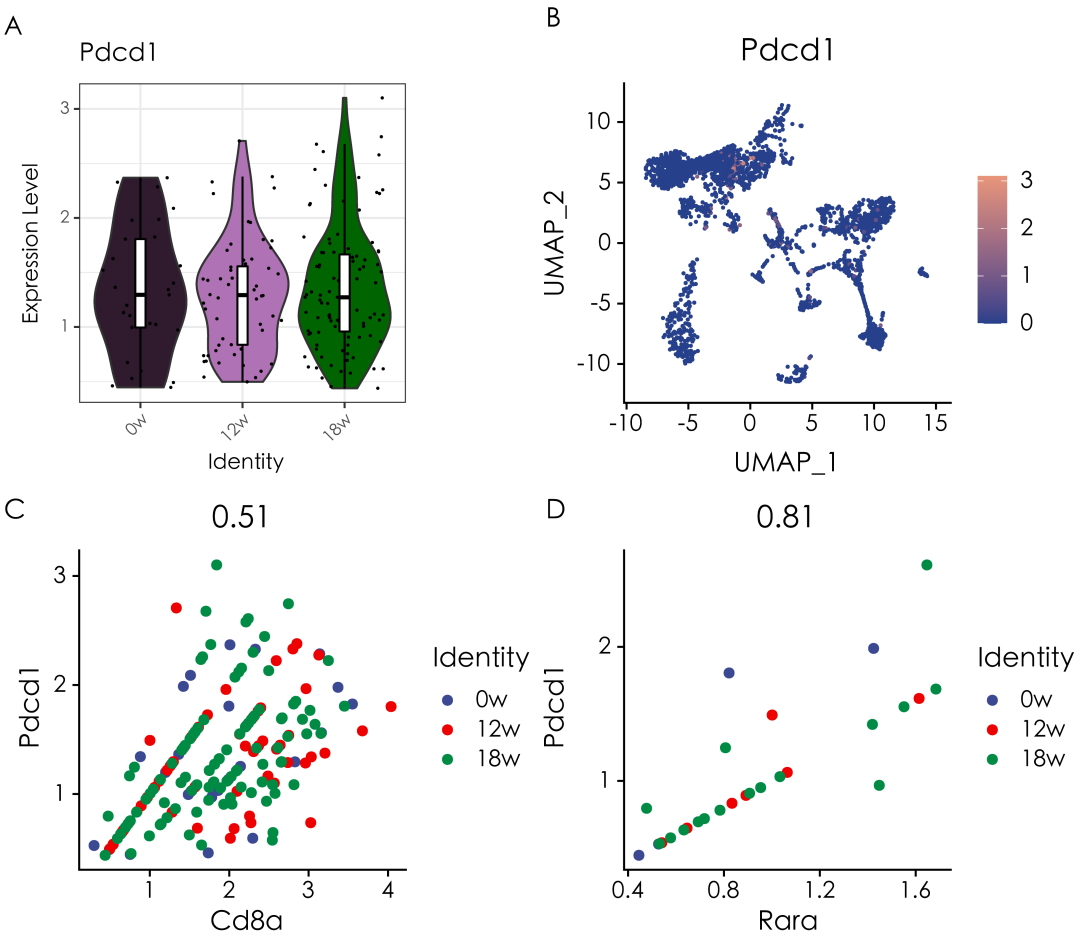


Figure S9. The expression levels of Pdcd1 and its correlation with Cd8a and Rara

(A) Violin plot presenting the expression levels of Pdcd1 at 0, 12, and 18 weeks. (B) UMAP plot showing the expression levels of Pdcd1. (C) Correlation analysis of Cd8a and Pdcd1 in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green). (D) Correlation analysis of Rara and Pdcd1 in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green).

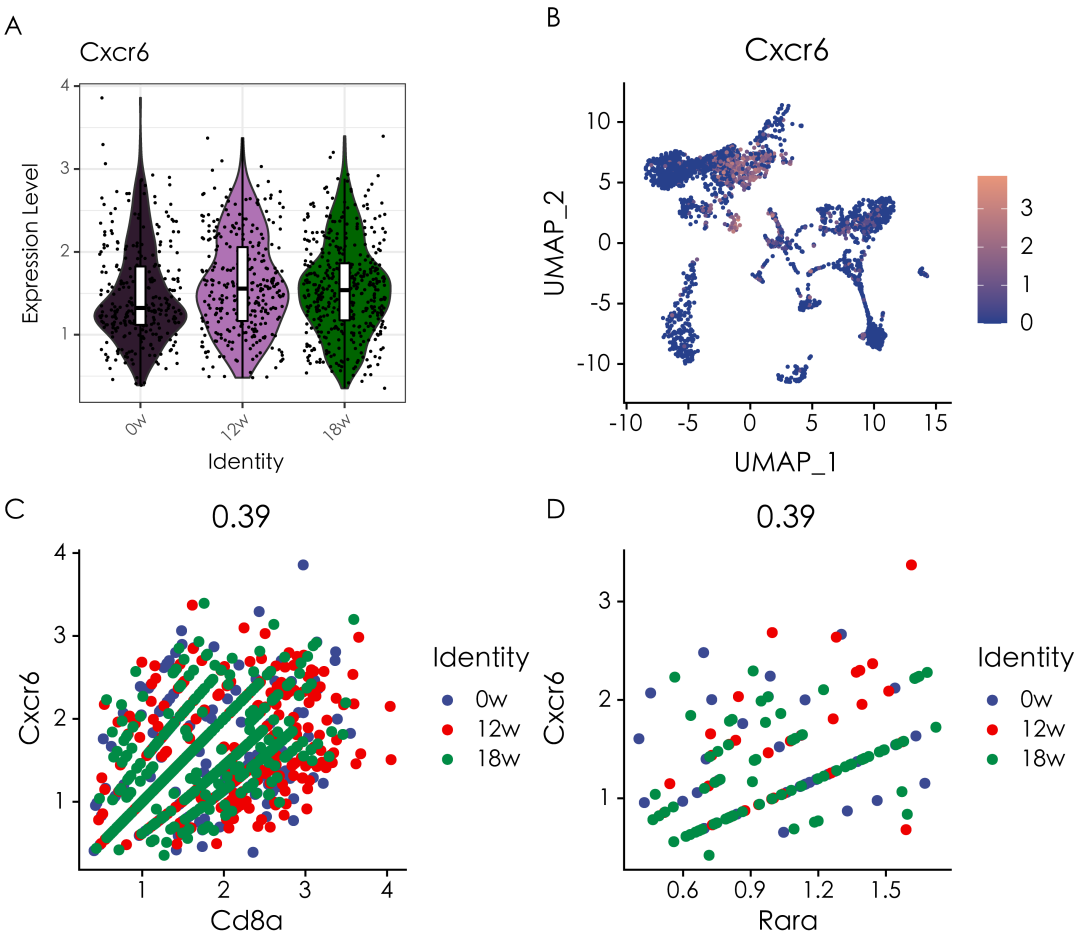


Figure S10. The expression levels of Cxcr6 and its correlation with Cd8a and Rara

(A) Violin plot presenting the expression levels of Cxcr6 at 0, 12, and 18 weeks. (B) UMAP plot showing the expression levels of Cxcr6. (C) Correlation analysis of Cd8a and Cxcr6 in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green). (D) Correlation analysis of Rara and Cxcr6 in *Fah*$^{-/-}$ mouse at 0 (blue), 12 (red), and 18 weeks (green).

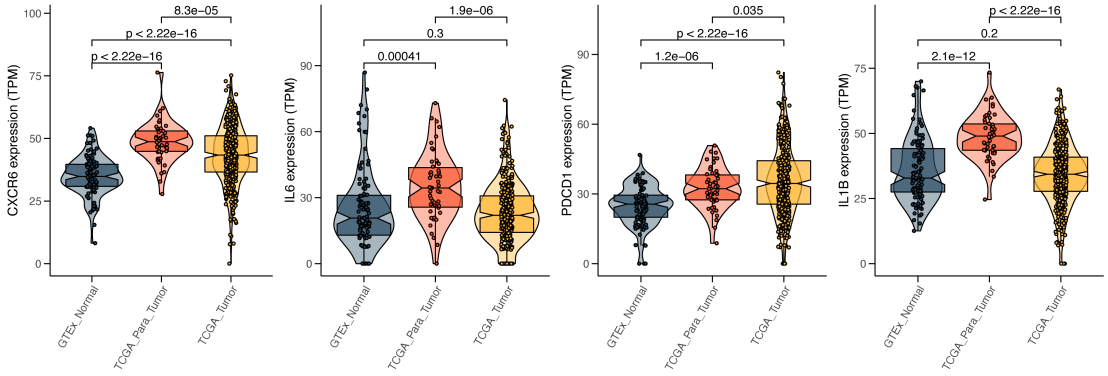


Figure S11. The expression characteristics of CD8+ T cells in GTEx normal, TCGA para-tumor and TCGA tumors tissues

Violin plot presenting the expression levels of CXCR6, IL6, PDCD1 and IL1B in GTEx normal, TCGA para-tumor and TCGA tumors tissues.

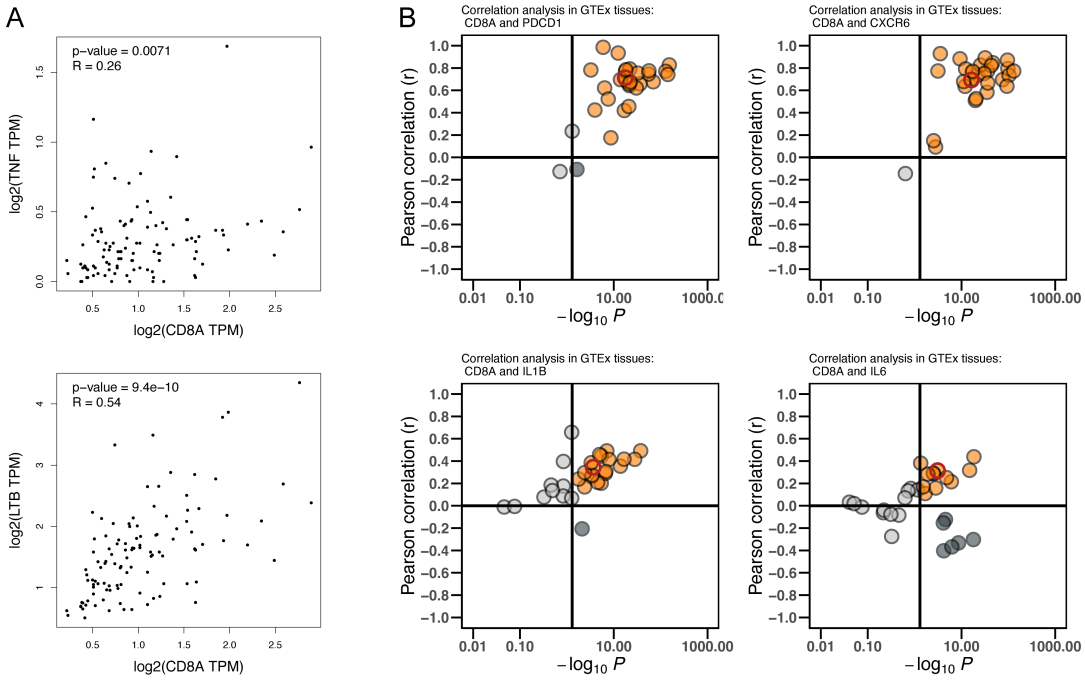


Figure S12. CD8A is positively correlated with pro-tumor cytokines in GTEx normal tissues

(A) Correlation analysis of CD8A and TNF and LTB in 30 GTEx normal tissues.

(B) Correlation analysis of CD8A and PDCD1, CXCR6, IL6, IL1B in 30 GTEx normal tissues.

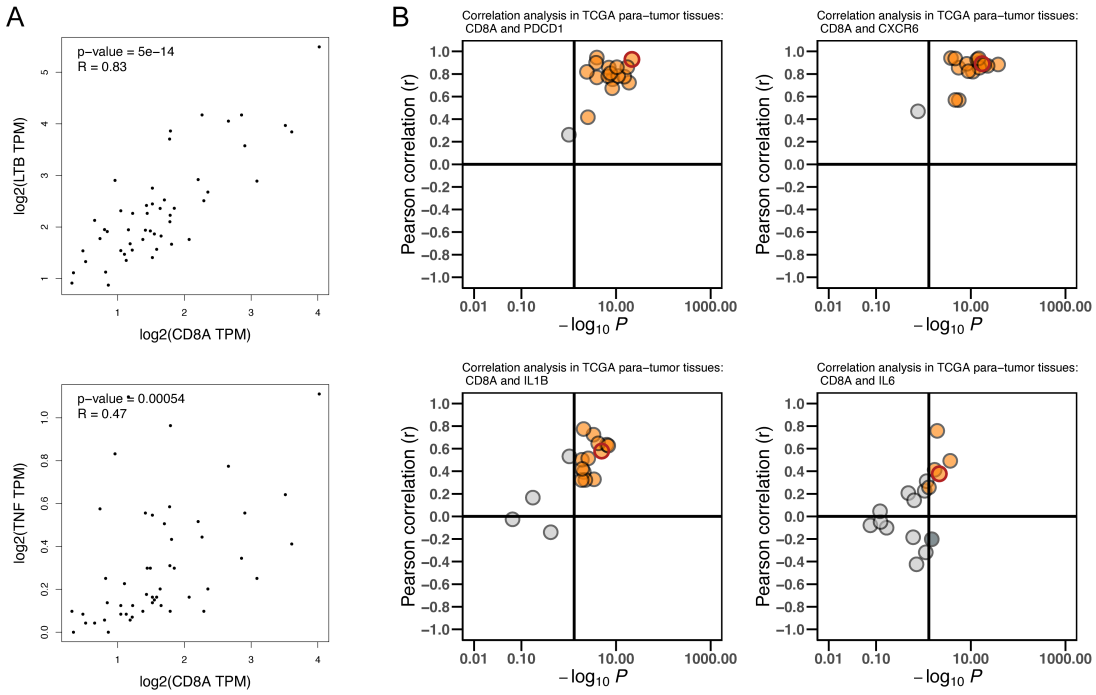


Figure S13. CD8A is positively correlated with pro-tumor cytokines in TCGA para-tumor tissues

(A) Correlation analysis of CD8A and TNF and LTB in 17 TCGA para-tumor tissues.

(B) Correlation analysis of CD8A and PDCD1, CXCR6, IL6, IL1B in 17 TCGA para-tumor tissues.

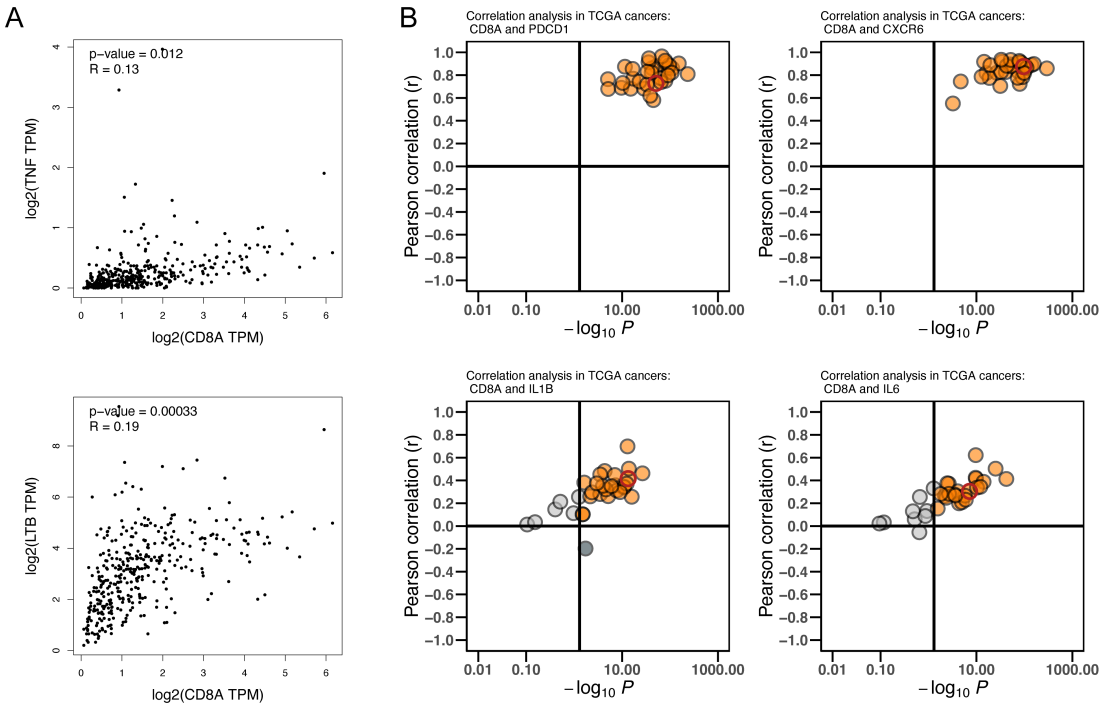


Figure S14. CD8A is positively correlated with pro-tumor cytokines in TCGA tumors tissues

(A) Correlation analysis of CD8A and TNF and LTB in 34 TCGA tumors tissues.

(B) Correlation analysis of CD8A and PDCD1, CXCR6, IL6, IL1B in 34 TCGA tumors tissues.