



# Liver pathophysiology at the single cell level: characterization of cellular heterogeneity and identification of novel therapeutic targets for chronic liver diseases and hepatocellular carcinoma

Antonio Saviano

## ► To cite this version:

Antonio Saviano. Liver pathophysiology at the single cell level: characterization of cellular heterogeneity and identification of novel therapeutic targets for chronic liver diseases and hepatocellular carcinoma. Quantitative Methods [q-bio.QM]. Université de Strasbourg, 2019. English. NNT : 2019STRAJ093 . tel-03270802

**HAL Id: tel-03270802**

**<https://theses.hal.science/tel-03270802>**

Submitted on 25 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ**  
**UMR\_S1110 Institut de Recherche sur Les Maladies Virales et Hépatiques**

# **THÈSE** présentée par :

**Antonio SAVIANO**

soutenue le : **28 Novembre 2019**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Sciences de la Vie et de la Santé, Biologie des systèmes

## **Physiopathologie du foie à l'échelle de la cellule unique : caractérisation de l'hétérogénéité cellulaire et identification de nouvelles cibles thérapeutiques dans les maladies hépatiques chroniques et le cancer hépatocellulaire**

### **THÈSE dirigée par :**

**Prof. Thomas F. BAUMERT**  
**Prof. Patrick PESSAUX**

PU-PH, Université de Strasbourg  
PU-PH, Université de Strasbourg

### **RAPPORTEURS :**

**Prof. Tarik ASSELAH**  
**Dr. Magdalena FILIPOWICZ**

PU-PH, Université Paris Diderot  
Université de Bâle, Suisse

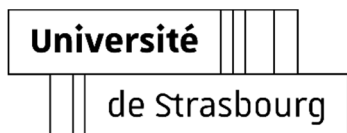
### **AUTRES MEMBRES DU JURY :**

**Prof. Gabriel MALOUF**  
**Dr. Nicolas GOOSSENS**

PU-PH, Université de Strasbourg  
Université de Genève, Suisse

---





**UNIVERSITE DE STRASBOURG**



***ECOLE DOCTORALE 414***

UMR\_S1110 Institut de Recherche sur Les Maladies Virales et Hépatiques

Thèse présentée par :

**Antonio SAVIANO**

Soutenue le : **28 Novembre 2019**

Pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/S spécialité : Sciences de la Vie et de la Santé, Biologie des systèmes

**Liver pathophysiology at the single cell level:  
characterization of cellular heterogeneity and identification  
of novel therapeutic targets for chronic liver diseases and  
hepatocellular carcinoma**

**THESE dirigée par :**

**Prof. Thomas F. BAUMERT** Université de Strasbourg

**Prof. Patrick PESSAUX** Université de Strasbourg

**RAPPORTEURS :**

**Prof. Tarik ASSELAH** Université Paris Diderot

**Dr. Magdalena FILIPOWICZ** Université de Bâle, Suisse

**AUTRES MEMBRES DU JURY :**

**Prof. Gabriel MALOUF** Université de Strasbourg

**Dr. Nicolas GOOSSENS** Université de Genève, Suisse



# TABLE OF CONTENTS

|   |           |
|---|-----------|
| <b>ABBREVIATIONS.....</b>   | <b>5</b>  |
| <b>1. INTRODUCTION.....</b>   | <b>9</b>  |
| 1.1    GLOBAL BURDEN OF LIVER DISEASES.....   | 10        |
| 1.2    MORTALITY CHANGES IN THE ANTIVIRAL ERA.....  | 12        |
| 1.3    FROM CHRONIC LIVER DISEASE TO LIVER CIRRHOSIS AND<br>HEPATOCELLULAR CARCINOMA..... | 15        |
| 1.4    HEPATOCARCINOGENESIS IN NAFLD.....   | 19        |
| 1.5    HBV-INDUCED HEPATOCARCINOGENESIS.....  | 23        |
| 1.6    HETEROGENEITY, MOLECULAR AND HISTOLOGICAL HCC<br>CLASSIFICATION.....               | 28        |
| 1.7    TREATMENT OF LIVER FIBROSIS.....   | 36        |
| 1.8    HCC CHEMOPREVENTION.....   | 39        |
| 1.9    HCC TREATMENT.....   | 43        |
| 1.10   SINGLE-CELL RNA SEQUENCING PRINCIPLES AND TECHNOLOGIES.....                        | 47        |
| 1.11   SINGLE-CELL RNA-SEQUENCING TO STUDY LIVER PHYSIOLOGY AND<br>DISEASES.....          | 50        |
| 1.11.1 LIVER HETEROGENEITY AND ZONATION.....  | 50        |
| 1.11.2 HEPATOBILIARY PROGENITORS.....   | 56        |
| 1.11.3 NASH AND LIVER FIBROSIS.....   | 60        |
| 1.11.4 PRIMARY LIVER CANCER.....  | 68        |
| <b>2. RESEARCH OBJECTIVES.....</b>  | <b>72</b> |

|   |     |
|---|-----|
| <b>3. RESULTS</b>   | 73  |
| 3.1 A HUMAN LIVER CELL ATLAS REVEALS HETEROGENEITY AND EPITHELIAL PROGENITORS   | 73  |
| 3.2 VIRAL COMPARTMENTALIZATION, CANCER HETEROGENEITY IN HBV-INDUCED HEPATOCELLULAR CARCINOMA                            | 102 |
| <b>4. CONCLUSIONS AND PERSPECTIVES</b>  | 136 |
| <b>5. RESUME DE LA RECHERCHE</b>  | 141 |
| 5.1 INTRODUCTION  | 141 |
| 5.2 RESULTATS   | 144 |
| 5.2.1 UN ATLAS DES CELLULES DU FOIE HUMAIN REVELE SON HETEROGENEITE CELLULAIRE ET IDENTIFIE LES PROGENITEURS HEPATIQUES | 144 |
| 5.2.3 COMPARTIMENTATION VIRALE ET HETEROGENEITE CANCEREUSE DANS LE CHC INDUITS PAR LE VIRUS DE L'HEPATITE B             | 148 |
| 5.3 DISCUSSION  | 150 |
| <b>6. REFERENCES</b>  | 151 |
| <b>7. SUPPLEMENTARY MATERIAL</b>  | 166 |
| 7.1 Hamdane N, Jühling F et al. Gastroenterology, 2019 Jun  | 166 |
| 7.2 Felli E, Saviano A et al. Hepatology 2019, Oct 9  | 191 |
| 7.3 Wakabayashi T et al. Hepatol Int 2019, Aug 31   | 198 |
| 7.4 Turon-Lagot V et al. Virologie 2019, Jun 1  | 213 |
| 7.5 Saviano A et al. Hepatobiliary Surg Nutr 2019, Jun 8  | 225 |
| 7.6 Saviano A, Roehlen N et al. Hepatocellular carcinoma, Springer 2019   | 230 |
| <b>8. CURRICULUM VITAE</b>  | 246 |

## ABBREVIATIONS

|          |   |
|----------|---|
| AASLD    | American association for the study of liver diseases        |
| AFP      | Alpha-fetoprotein   |
| ALD      | Alcoholic liver disease                                     |
| AMLN     | Amylin diet   |
| ASK1     | Apoptosis signal-regulating kinase 1                        |
| BCLC     | Barcelona Clinic Liver Cancer                               |
| BECs     | Biliary epithelial cells                                    |
| CAFs     | Cancer-associated fibroblasts                               |
| CCL2     | C-C motif chemokine ligand 2                                |
| CCl4     | Carbon tetrachloride  |
| cDCs     | Conventional Dendritic cells                                |
| cDNA     | Complementary DNA   |
| chCC-ICC | Combined hepatocellular and intrahepatic cholangiocarcinoma |
| COX2     | Cyclooxygenase-2  |
| CTCs     | Circulating tumor cells                                     |
| CTLA-4   | T-lymphocyte associated protein 4                           |
| DAA      | Direct antiviral agents                                     |
| DAMPs    | Damage-associated molecular patterns                        |
| DCs      | Dendritic cells   |
| DDC      | 3,5-diethoxycarbonyl-1,4-dihydrocollidine                   |
| DDR      | DNA damage response   |
| EASL     | European association for the study of the liver             |

|               |   |
|---------------|---|
| EGFR          | Epidermal growth factor receptor            |
| EMT           | Epithelial-mesenchymal transition           |
| ERK           | Extracellular signal-regulated kinases      |
| ETV           | Entecavir                                   |
| FACS          | Flow-activated cell sorting                 |
| FFAs          | Fat- tissue-derived free fatty acids        |
| FLC           | Fibrolamellar carcinoma                     |
| G-CSF         | Granulocyte-colony stimulating factor       |
| KCs           | Kupffer cells                               |
| FXR           | Farnesoid X receptor                        |
| HAEnd         | Arterial endothelial cells                  |
| HBV           | Hepatitis B                                 |
| HCC           | Hepatocellular carcinoma                    |
| HCV           | Hepatitis C                                 |
| HHyP          | Hepatobiliary hybrid progenitor ()          |
| HIVID         | High-throughput viral integration detection |
| HSC           | Hepatic stellate cell                       |
| ICC           | Intrahepatic cholangiocarcinoma             |
| IFN- $\gamma$ | Interferon gamma                            |
| IL            | Interleukin                                 |
| IVT           | In vitro transcription                      |
| LECs          | Liver endothelial cells                     |
| LEL-C         | Lymphoepithelioma-like carcinoma            |
| LINE          | Long interspersed nuclear element           |

|                |  |
|----------------|--|
| LOXL2          | Lysyl oxidase homolog 2                          |
| LPA            | Lysophosphatidic acid                            |
| LPAR1          | LPA receptor 1                                   |
| LPS            | Lipopolysaccharide                               |
| LSECs          | Liver sinusoidal endothelial cells               |
| MFB            | Myofibroblast                                    |
| mTOR           | Mammalian target of rapamycin                    |
| MPTP           | Mitochondrial permeability transition pore       |
| MTM-HCC        | Macrotrabecular-massive hepatocellular carcinoma |
| NAFL           | Non-alcoholic fatty liver                        |
| NAFLD          | Non-alcoholic fatty liver disease                |
| NAS            | NAFLD activity score                             |
| NSAIDs         | Non-steroidal anti-inflammatory drugs            |
| NASH           | Non-alcoholic steatohepatitis                    |
| NF- $\kappa$ B | Nuclear factor kappa B                           |
| NUCs           | Nucleos(t)ide analogues                          |
| OCA            | Obethicolic acid                                 |
| OS             | Overall survival                                 |
| PAMPs          | Pathogen-associated molecular patterns           |
| PPARs          | Peroxisome proliferator-activated receptors      |
| PD-1           | Programmed death 1                               |
| PDGF-B         | Platelet derived growth factor subunit B         |
| PG-HCC         | Progenitor hepatocellular carcinoma              |
| PSC            | Primary sclerosing cholangitis                   |

|                   |  |
|-------------------|--|
| RCT               | Randomized clinical trial                          |
| ROS               | Reactive oxygen species                            |
| SAEndo            | Scar-associated endothelial cells                  |
| SAM $\Phi$        | Scar-associated macrophages                        |
| SAMes             | Scar-associated mesenchymal cells                  |
| ScRNA-seq         | Single-cell RNA-sequencing                         |
| SmRNA FISH        | Single molecule RNA FISH                           |
| STAT3             | Signal transducer and activator of transcription 3 |
| TDF               | Tenofovir disoproxil fumarate                      |
| TGF- $\beta$      | Transforming growth factor beta                    |
| T <sub>H</sub> 17 | T helper 17  |
| TKI               | Tyrosine kinase inhibitor                          |
| TMo               | Tissue monocyte-derived macrophages                |
| TNF- $\alpha$     | Tumor necrosis factor alpha                        |
| Treg              | T regulatory cells                                 |
| UMIs              | Unique molecular identifiers                       |



# 1. INTRODUCTION

Liver cirrhosis, viral hepatitis and hepatocellular carcinoma (HCC) account for at least 2 million deaths per year worldwide<sup>1</sup>. Liver cirrhosis is an irreversible condition, few drugs limited by low efficacy and tolerability are available for liver cancer. Remarkable advancements were made in the last years in the treatment of viral hepatitis. High genetic barrier nucleos(t)ide analogues (NUCs) are now available for the treatment of chronic hepatitis B (HBV) and highly effective direct antiviral agents (DAAs) that lead to a definitive viral cure are available for hepatitis C (HCV). However, NUCs are not effective in eradicate HBV infection and the HCC risk is not completely eliminated upon viral treatment in patients with severe liver fibrosis<sup>2</sup>. The absence of effective therapies for liver cirrhosis and HCC is mainly attributable to the complexity of liver physiopathology. This complexity is poorly understood and translates in difficulties in identifying effective and safe therapeutic targets. Recently, single-cell analysis has open new insights and research possibilities in biology and medicine. The investigation at single-cell resolution of the molecular mechanisms underlying liver disease progression and HCC development will have major impact in the next future helping in uncovering novel targets for liver fibrosis and HCC prevention and treatment.

## 1.1 GLOBAL BURDEN OF LIVER DISEASES

Liver cirrhosis is the first cause of mortality for digestive disease worldwide<sup>3</sup>. Cirrhosis and liver cancer represent, respectively, the 13<sup>th</sup> and the 19<sup>th</sup> most common cause of death. Combined, they account for 3.8% of all deaths in the world<sup>3</sup>. This marks an increase from 2000 when liver-related mortality accounted for 3% of all deaths. The burden is even higher if deaths related to acute hepatitis (126,400) and alcohol-use disorders are taken into account (184,900)<sup>3</sup>.

In 2018 there were approximately 840,000 new cases of liver cancer, the 6<sup>th</sup> leading cause of cancer worldwide<sup>4</sup>. Primary liver cancers are more common in men representing the 5<sup>th</sup> leading cancer for incidence (9<sup>th</sup> among women) and the 2<sup>nd</sup> leading cause of cancer death (6<sup>th</sup> among women)<sup>4</sup>. According to the last Global Disease Burden report, in 2017 about 40% of HCC were due to HBV, 29% to HCV, 16% to alcohol, 8% to non-alcoholic steatohepatitis (NASH) and 7% due to other causes<sup>3</sup>. Importantly, mortality of NASH-related liver cancer increased of 42.3% in the period 2007-2017 showing a dramatic changing in the epidemiology of liver diseases<sup>3</sup>.

Data from The U.S. Surveillance, Epidemiology, and End Results Program estimated the 5-year survival for liver and intrahepatic ducts cancer to be only 18%<sup>5</sup>. Hepatobiliary cancer has, therefore, the second worst survival rate among cancers in US, second only to pancreatic cancer (5-year survival 8%) and 5 times worse than colorectal cancer (5-year survival 65%)<sup>5</sup>.

The quality of life of patients with chronic liver diseases and HCC is low and the economic impact is significant<sup>5,6</sup>. In US, patients with chronic liver disease – compared to those without chronic

liver disease - are more likely to be unemployed (55.3% vs. 30.7%), to have higher rates of disability/illness related unemployment (30.5% vs. 6.6%), job loss (10.2% vs. 3.4%), health care use and health care expenditures (\$19,390 vs. \$5,567 per year)<sup>6</sup>. Furthermore, liver diseases are associated with a variety of extrahepatic morbidities (e.g. NASH and cardiovascular diseases), which significantly contribute to mortality and reduced quality of life. It is not surprising that liver cirrhosis is within the top 20 causes of disability-adjusted life years and years of life lost <sup>5</sup>.

Accurate statistics on liver diseases burden are difficult to obtain since cause-specific mortality data for many countries, especially regions where liver diseases are highly prevalent as in Africa, are missing. Moreover, the incidence and prevalence of chronic liver diseases are not well reported even in areas where population-based studies are available<sup>5</sup>. The quality of data is additionally limited by referral bias (e.g. primary vs. tertiary centers), population composition (e.g. outpatients or inpatients), lack of standardization in definition and assessment of liver disease etiology (e.g. NASH) and stage (e.g. liver biopsy or non-invasive methods)<sup>5</sup>.

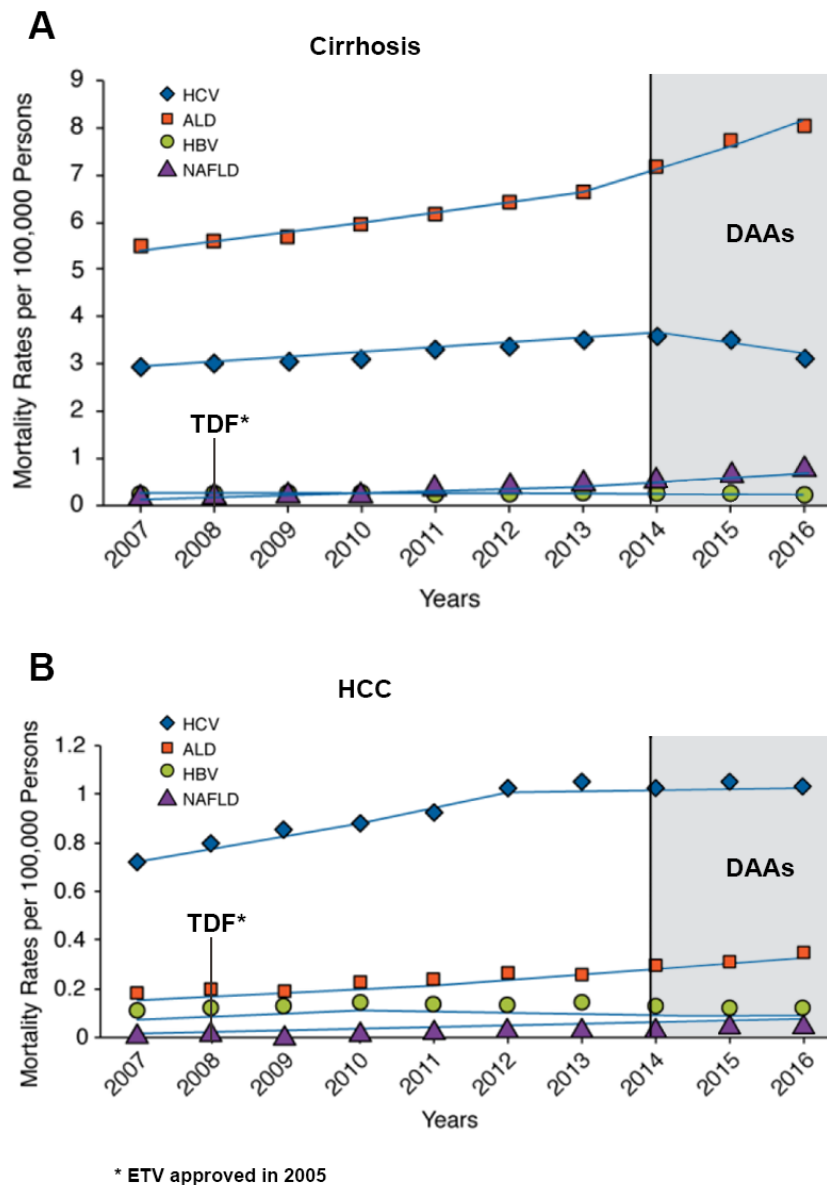
## 1.2 MORTALITY CHANGES IN THE ANTIVIRAL ERA

From 1980 to 2010 liver cirrhosis and primary liver cancer deaths increased worldwide of 52.2% and 62.4% respectively<sup>1,7</sup>. Chronic viral liver diseases were among the most important etiologies of cirrhosis and HCC. The development of direct antiviral agents (DAAs) has revolutionized HCV patients care. DAAs are highly effective antiviral drugs (success rates higher than 90%) and very well tolerated by difficult-to-treat patients with advanced liver disease. Antiviral agents for HBV with high genetic barrier were also introduced in the clinical setting in the last 15 years: entecavir was approved by FDA in 2005 and tenofovir disoproxil fumarate in 2008.

Antiviral treatments demonstrated to reverse and prevent liver decompensation and to decrease the prevalence of HBV and HCV-related end-stage liver disease<sup>8,9</sup>. HCV viral cure was associated with a decreased HCC risk in several large clinical cohorts<sup>10-13</sup>. However, HCC risk persists after HCV treatment in particular in patients with advanced liver fibrosis<sup>12,14</sup>. Moreover, in HCV patients with a history of HCC, unexpected high early recurrence rates after DAAs raised concerns on the safety profile of these drugs in terms of HCC prevention<sup>15</sup>.

Recent data from U.S. have shown that, since DAAs introduction in late 2013, HCV-related cirrhosis mortality significantly decreased (**Fig. 1a**). Despite the alert on early HCC recurrence in DAA-treated patients<sup>16,17</sup>, HCV-related HCC mortality did not increase and HBV-related HCC overall mortality decreased after 2010 (**Fig. 1b**). In the natural history of chronic liver diseases, HCC occurrence is a late event. In clinical cohorts, 1–8% per year of patients with liver cirrhosis

will develop HCC<sup>12,18</sup> suggesting that long-term studies (> 10 years) are needed to evaluate the real benefit of antiviral treatments on HCC mortality.



**Figure 1. Age-standardized mortality rates for liver cirrhosis (A) and HCC (B) in U.S. between 2007 and 2016.** HCV-related cirrhosis mortality decreased after DAAs while HCV-related HCC mortality did not increase. With the use of entecavir (ETV) and tenofovir (TDF), mortality for HBV-related cirrhosis significantly decreased and HBV-related HCC mortality showed a decline trend. Alcoholic liver disease (ALD) and NAFLD mortality for cirrhosis constantly raised in the last years. Adapted from Kim et al. and Saviano et al.<sup>19,20</sup>.

As viral liver disease mortality started to decrease, mortality rates for alcoholic liver disease and non-alcoholic fatty liver disease (NAFLD) dramatically increased in western countries in the last 10-years<sup>19</sup>.

In US, NAFLD is currently the third cause of cirrhosis mortality and the fourth cause of HCC-related mortality with dramatic annual percent changes of respectively +15.4% and +19.1%<sup>19</sup>. In spite of these impressive trends for both ALD and NAFLD, HCV still accounted for most of HCC deaths during the study period confirming the different disease prevalence and cancer risk between viral and metabolic diseases (**Fig. 1**)<sup>18</sup>.

Epidemiology changes in the antiviral era represent new challenges for the medical community in terms of a better understanding of metabolic liver diseases physiopathology and HCC development and establishing new treatments for liver cancers patients that will be more likely affected by severe extrahepatic metabolic disorders.

### **1.3 FROM CHRONIC LIVER DISEASE TO LIVER CIRRHOSIS AND HEPATOCELLULAR CARCINOMA**

The liver is a complex multifunctional organ with key roles in metabolism, detoxification, immune response regulation and tolerance. In normal conditions, the liver is continuously exposed to pathogens and toxins derived from the gut and removes large amount of bacteria, microbes, pathogen-associated and damage-associated molecular patterns (PAMPs and DAMPs) to maintain a tolerant and immunosuppressive environment<sup>21</sup>. In response to a chronic hepatocyte damage, immune and stromal liver cells modify the immune tolerant environment and promote and sustain chronic inflammation that ultimately drives to liver fibrosis and HCC<sup>22</sup>.

This deregulation in the liver immune balance is common to all chronic liver diseases and triggers cellular stress and death, apoptosis, liver fibrosis, hepatocyte regeneration and proliferation<sup>22</sup>.

Chronic hepatocyte injury activates liver immune system including T, B, NK, NKT, dendritic cells (DCs), monocyte-derived and liver-resident macrophages (the latter also known as Kupffer cells, KCs). Upon stimulation, these cells produce pro-inflammatory cytokines and chemokines and activate quiescent hepatic stellate cells (HSCs) into myofibroblasts (MFBs) that are finally responsible for collagen and extra-cellular matrix accumulation a hallmark of liver cirrhosis<sup>23</sup>.

During chronic liver diseases, the presence of PAMPs or the release of DAMPs resulting from hepatocytes apoptosis and death activate immune cells and especially KCs. KCs present viral antigens to T cells and secrete chemokines and cytokines to recruit circulating immune cells<sup>24</sup>.

Proinflammatory cytokines such as IL-17 secreted by helper T cells, transforming growth factor

beta (TGF- $\beta$ ) and platelet derived growth factor subunit B (PDGF-B) secreted by monocyte-derived macrophages and KCs, activate and differentiate HSCs into collagen-producing MFBs in the attempt of promoting compensatory tissue repair mechanisms<sup>25,26</sup>. Importantly, HSCs can also directly be activated to MFBs by DAMPs<sup>27,28</sup> and play a key role in both progression to advanced liver fibrosis and HCC development<sup>29</sup>.

TGF- $\beta$  also induces epithelial-mesenchymal transition (EMT) in hepatocytes increasing the population of collagen-producing cells. Indeed, EMT is a process in which epithelial cells develop higher motility, invasive properties and mesenchymal features.

Other important proinflammatory molecules that alter liver immune balance are interferon gamma (IFN- $\gamma$ ), tumor necrosis factor alpha (TNF- $\alpha$ ), C-C motif chemokine ligand 2 (CCL2) and interleukins (IL) such as IL-6, IL-1 $\beta$ , IL-2, IL-7 and IL-15.

Etiology-independent mechanisms that from chronic inflammation and fibrosis lead to carcinogenesis are less described. Certainly, the inflammation generates hepatocellular stress and induce DNA damage, chromosomal aberrations, epigenetic modifications and mitochondrial alterations<sup>27</sup>. On the signaling level, proinflammatory molecules upregulate signal transducer and activator of transcription 3 (STAT3) and nuclear factor kappa B (NF- $\kappa$ B) both promoting cell proliferation, escape from apoptosis, survival and angiogenesis<sup>30,31</sup>.

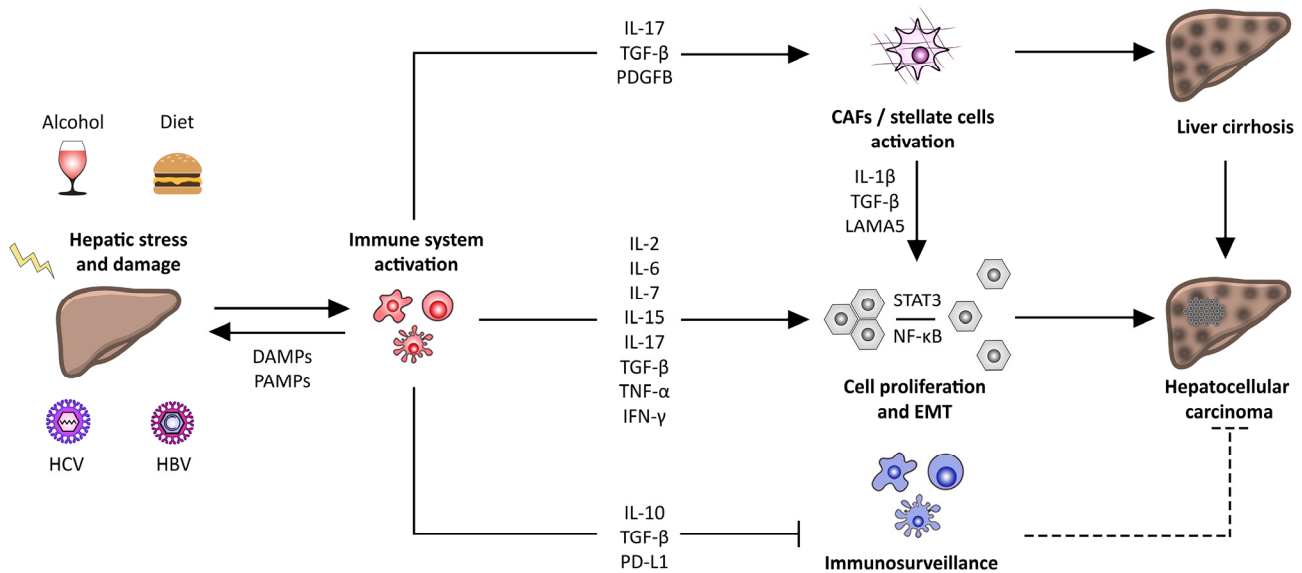
Chronic inflammation and compensatory immune changes induce T cell exhaustion attenuating immune surveillance and contributing to HCC risk<sup>32-34</sup>. Immunosuppressive functions are carried out by regulatory T cells (Treg) secreting IL-10 and TGF- $\beta$  and presenting checkpoint molecules



as cytotoxic programmed death 1 (PD-1) and T-lymphocyte associated protein 4 (CTLA-4)<sup>35,36</sup>. TGF- $\beta$  reduces IL2-induced T cell proliferation, secretion of proinflammatory cytokines and activation of effector cells<sup>37-39</sup>. IL-10 is also secreted by KCs and suppress T-cell proliferation, macrophages activation and IFN- $\gamma$  production furtherly reducing immune surveillance<sup>40,41</sup>. The relevance of these cytokines is also indirectly suggested by the evidences reporting that elevated levels of IL-10 and TGF- $\beta$  in patients with chronic liver diseases correlate with disease progression and survival<sup>40,42</sup>.

Liver and tumor microenvironment have also a relevant effect on tumor progression and treatment response<sup>43</sup>. Liver tumor microenvironment is a dynamic structure of tumor cells within the extracellular matrix populated by stromal cells and the proteins that they secrete. The interactions of tumor cells with sinusoidal and extrasinusoidal cells affect tumor cell fate and participate in cancer development and progression. Stromal and immune cells contribute to these processes providing signals to induce angiogenesis, reprogram cell metabolism, support cell proliferation, escape apoptosis and immune surveillance and promote cellular immortality, invasion and metastasis<sup>43-45</sup>.

Liver fibrosis and HCC development in chronic liver disease are multifactorial event in which the stromal and the immune non-parenchymal cells play an important role (**Fig. 2**).



**Figure 2. Chronic inflammation as a driver of hepatocarcinogenesis and liver fibrosis.** Hepatocarcinogenesis can be induced by multiple etiological and environmental conditions that trigger the activation of the immune system via release of Damage-Associated Molecular Patterns (DAMPs) and/or Pathogen Associated Molecular Patterns (PAMPs). The persistent dysregulation of the immunological network of the liver, promoted by the secretion of pro-inflammatory cytokines/chemokines, leads to cells death, compensatory hepatocellular proliferation, activation of cancer-associated fibroblasts (CAFs) and hepatic stellate cells (HSCs) as well as epithelial-mesenchymal transition (EMT). The sustained necro-inflammatory status attenuates immune-surveillance and anti-tumor immune response via anti-inflammatory molecules (e.g. IL-10, TGF-β, PD-L1). HSCs activation contributes to cirrhosis and HCC development. From Saviano et al.<sup>46</sup>

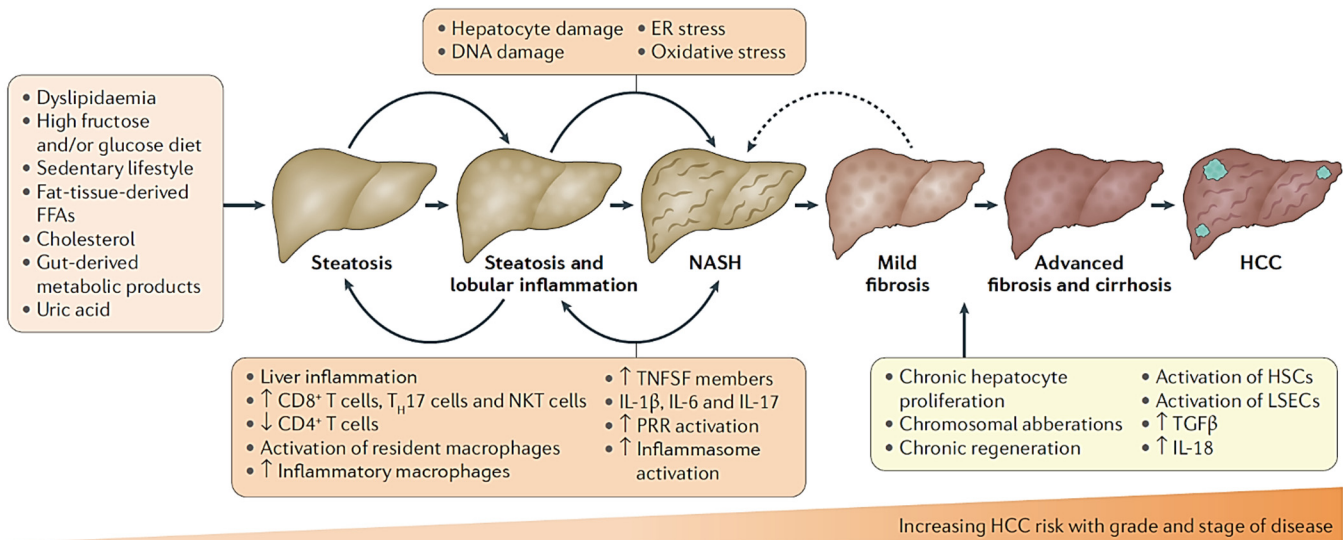
In the immunotherapy era, it is clear that understanding the biology and targeting liver and tumor microenvironment constitutes a rationale for novel therapeutic strategies. Even though there are robust evidences linking perturbed microenvironment, liver fibrosis and HCC, little is known about the single-cell transcriptomic profiles of liver microenvironment.

## 1.4 HEPATOCARCINOGENESIS IN NAFLD

NAFLD has emerged as a leading cause of end-stage liver disease as well as HCC. NAFLD encompasses a clinico-pathologic spectrum of diseases ranging from hepatic steatosis (non-alcoholic fatty liver, NAFL) to non-alcoholic steatohepatitis (NASH) that can progress to advanced fibrosis, cirrhosis and HCC. Studies have demonstrated that the incidence of HCC in patients with NAFLD ranges from 2.4% over 7 years to 12.8% over 3 years<sup>47</sup>. In US, NAFLD-related mortality dramatically increased in the last 10 years<sup>19</sup>.

NAFLD-related HCC has unique features in terms of carcinogenesis and clinical presentation. HCC in NAFLD is diagnosed more often in patients without cirrhosis and is associated with late diagnosis and higher tumor burden<sup>48</sup>. Moreover, patients with NAFLD-related cirrhosis receive sub-optimal HCC surveillance in comparison to patients with HCV cirrhosis and ultrasound surveillance can be inadequate due to the general body habitus of the patients<sup>49</sup>.

The carcinogenetic process in NAFLD is not completely understood. Steatosis alone is not a driver of HCC, chronic inflammation is needed to induce cancer<sup>27</sup>. Fat- tissue-derived free fatty acids (FFAs) and gut-derived products lead to hepatocyte steatosis and lobular inflammation prevalently sustained by intrahepatic lymphocytes such as CD8+ T cells, T helper 17 (T<sub>H</sub>17) cells, NKT cells and infiltrating inflammatory macrophages. HCC develops in a chronic regenerative environment shaped by chronic hepatocyte cell death and chromosomal aberrations, compensatory proliferation, increased levels of TNF superfamily members, TGF $\beta$ , IL-18 and activation of HSCs and liver sinusoidal endothelial cells<sup>50</sup> (**Fig. 3**).



**Figure 3. Overview of NASH pathogenesis and hepatocarcinogenesis.** Free fatty acids (FFAs) and other metabolic and gut- derived products induce liver steatosis and lobular inflammation. Intrahepatic lymphocytes, such as CD8<sup>+</sup> T cells, T<sub>H</sub>17 cells, NKT cells and macrophages together with pro-inflammatory cytokines lead to chronic necroinflammation. Early stages of the disease are, to some extent, reversible. Chronic hepatocyte cell death and subsequent compensatory proliferation along with increased levels of TNF superfamily (TNFSF) members, TGF $\beta$  and IL-18, participate to HCC risk. These factors associated to hepatic stellate cells (HSCs) and liver sinusoidal endothelial cells (LSECs) activation as well as chronic hepatic proliferation accompanied by chromosomal aberrations, contribute to HCC development. Adapted from Anstee et al.<sup>50</sup>.

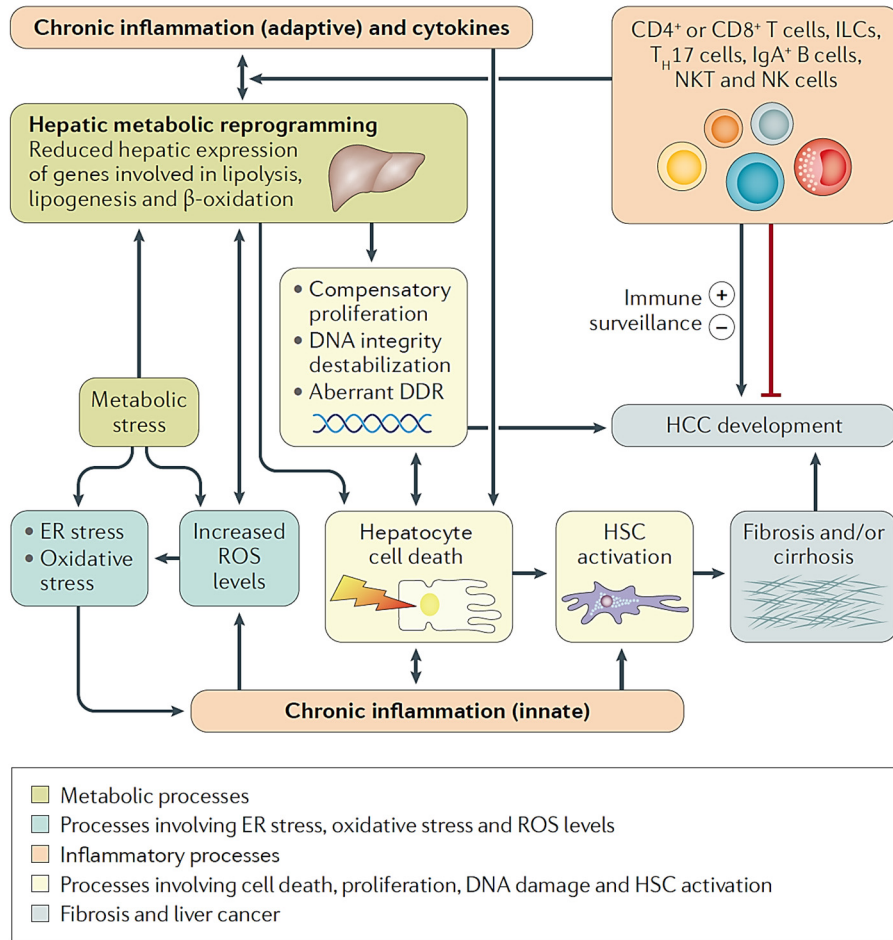
The increased hepatocyte metabolism and oxidation of fatty acid induce an overproduction of reactive oxygen species (ROS)<sup>51</sup>. The excess of triglycerides and FFAs impaired the initiation of autophagy through activation of mammalian target of rapamycin (mTOR). When the antioxidant capacity of the hepatocytes is exceeded, DNA damage and oxidation occurs, eventually resulting in chromosomal aberrations and cancer development<sup>52,53</sup>.

Increased gut permeability has also been suggested as responsible of the inflammatory and cancer development in NASH. Even though it is unclear if a leaky gut is the consequence or the cause of NASH, it is evident that the translocation of lipopolysaccharide (LPS) from Gram-negative bacteria is an important aggravating factor of liver inflammation and fibrosis<sup>50,54</sup>.

Metabolic alterations and chronic inflammation establish a vicious circle that sustain inflammation and enhance hepatocyte death and compensatory proliferation increasing HCC risk. Indeed, metabolic alterations trigger liver inflammation and chronic inflammation induces metabolic reprogramming and stress via a reduction of activity and expression of proteins involved in lipogenesis, lipolysis and  $\beta$ -oxidation<sup>55</sup> (**Fig. 4**).

Finally, persistent inflammation induces immune cell exhaustion and reduce immune surveillance for HCC development.

Hepatocarcinogenesis in NAFLD is a multifactorial process in which liver microenvironment play an important role in initiating inflammation, inducing hepatocyte death, proliferation and genetic aberrations finally leading to HCC.



**Figure 4. Metabolic reprogramming and chronic inflammation vicious cycle in NAFLD.** Chronic inflammation affects metabolism of hepatocytes through cytokine expression and downregulates metabolic genes implicated in the lipolysis and  $\beta$ -oxidation. The metabolic reprogramming and inflammation lead to chromosomal aberration and DNA damage response (DDR), hepatocyte death, HSC activation, liver cirrhosis and HCC development. Chronic inflammation and immune cell exhaustion reduces immune surveillance and increase the HCC risk. Adapted from Anstee et al.<sup>50</sup>.

## 1.5 HBV-INDUCED HEPATOCARCINOGENESIS

HBV patients have a high risk of HCC that is not dependent on fibrosis stage only. A direct role of the virus in HCC development is broadly reported. The genetic landscape of HBV-related HCC has some peculiar features. HBV-related HCC has higher frequency of TP53 mutations while  $\beta$ -catenin mutations are more frequent in non HBV-related HCC<sup>56,57</sup>. HBV-related HCC is also associated with an upregulation of miRNAs different from the ones in HCV-related HCC<sup>58</sup>.

HBV viral genome integrates into the genome of regenerating infected hepatocytes<sup>59,60</sup>. The integration is a main feature of HBV pathogenesis even though HBV does not contain any enzyme that facilitates this process and it is not an essential step in HBV replication<sup>61</sup>. The integration does not support viral replication since the circular structure of the genome is lost but participates in viral pathogenesis by activating or disrupting the expression or activity of cellular proteins (e.g. proto-oncogene or oncosuppressor)<sup>61</sup>.

The absence of repetitive, constant HBV integration sites in HCC tissues has made difficult to clearly understand the magnitude of HBV integration in HCC development. Integration is usually an early event during HBV infection and HCC cells have multiple integration sites making difficult to determine whether integration had a role in tumor development and progression<sup>62</sup>. HBV integrated genome has been reported in the proximity of genes controlling cell survival, proliferation and immortalization such as RAR $\beta$ , cyclin A or TERT<sup>60,63-67</sup>. Breakpoints in the HBV genome are often located close to the viral enhancer, HBx gene and core ORF<sup>68</sup>. Most of HBV breakpoints in human genome are near to coding genes usually upregulated in tumors such as

TERT, MML4, CCNE1 and ROCK1<sup>68</sup>. Interestingly, HBx gene can integrate in long interspersed nuclear element (LINE) to generate a HBx-LINE1 chimeric transcript that activates WNT signaling, sequesters miR-122 promoting EMT-like changes and liver injury and it is associated to poor patients' outcomes<sup>69,70</sup>. So far, HBx-LINE1 fusion transcript has been detected only in Asian patients suggesting that it is not an universal mechanism of carcinogenesis but might be related to HBV genotype variation<sup>71</sup>.

Large scale analysis of HBV-DNA integration sites using high-throughput viral integration detection (HIVID) of 426 paired biopsies of HBV-related HCC showed that HBV integration events occur more frequently in fragile sites, CpG islands, in proximity of telomeres, into chromosome 2 and 17 and in male patients suggesting a non-random integration followed by a positive selection during carcinogenesis<sup>72</sup>. HBV gene integrations in HCC were frequently associated with altered expression of the corresponding proteins and transcripts<sup>72</sup>.

Integrated HBV genomes also indirectly participate in oncogenesis by the production of altered or truncated viral proteins of HBx gene and HBV envelope proteins<sup>73,74</sup>. HBV genome encodes three envelope proteins (i.e. small, middle, large). Overproduction of complete or truncated forms of the envelope proteins from HBV integrated genome can induce endoplasmic reticulum stress and activated transduction pathways inducing upregulation of cyclin A, c-Raf-1, extracellular signal-regulated kinases (ERK) signaling finally stimulating cell proliferation<sup>61</sup>. Indeed, transgenic mice overexpressing HBV large envelope proteins present hepatocyte stress, liver inflammation and HCC<sup>75</sup>. Moreover, some of the mutations in the HBV envelope proteins (preS1 and preS2



regions) can involve regions containing immune cell epitopes producing immune escape variants of HBV<sup>76</sup>. Clinical data confirm that preS mutations are associated with an increased HCC risk of about 3.77 fold<sup>77</sup>.

HBx is the most relevant protein directly involved in HBV-related hepatocarcinogenesis. HBx has multiple roles: increases HBV replication, regulates host transcription factors, miRNAs levels and signal transduction pathways, apoptosis and cell proliferation. Although HBx is not essentially required for HBV infection, *in vitro* and *in vivo* data clearly demonstrated that it enhances and stimulates viral replication<sup>78-81</sup>. How HBx influences viral replication is still not fully understood. The most accredited hypothesis suggests that the modulation of proteasome activity and the interaction with the ultraviolet-damaged DNA-binding complex plays a key role<sup>82,83</sup>.

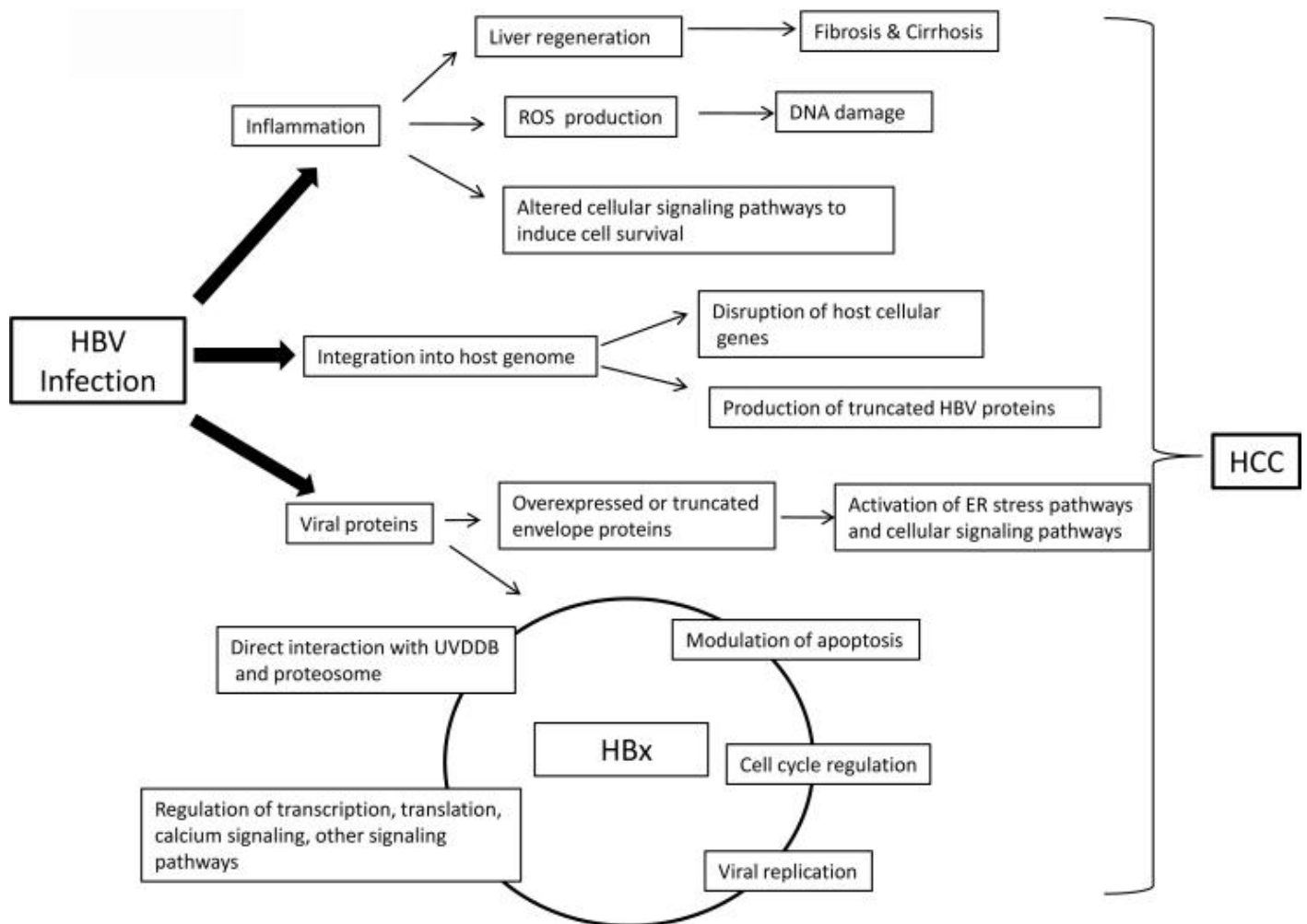
HBx modulates host transcription machinery by both direct interactions with transcription factors and activation of intracellular signaling. Indeed, HBx was found to interact with several transcription factors such as RPB5, TFIIB, TBP, TFIIH and CREB<sup>80</sup>. With an indirect, transduction-mediated effect, HBx activates NF- $\kappa$ B, AP-1 and NFAT<sup>80</sup>. HBx activates also Pyk2 and Src kinases and the Ras/Raf/MAPK pathways. Importantly, HBx regulates calcium signaling pathways through the interaction with mitochondrial permeability transition pore (MPTP) and modulates in this way several signaling pathways such as STAT3<sup>84-86</sup>. Interacting with mitochondria, HBx also elevates the levels of ROS that through a positive feedback increase intracellular calcium<sup>61</sup>. In addition, HBx was found to upregulate SMAD-dependent and non SMAD-dependent pathways of TGF $\beta$ 1 that mediate EMT changes in hepatocytes<sup>87,88</sup>.

The effects of HBx on apoptosis are variable since it has been reported that it can both prevent and induce apoptosis. The apoptotic effect is mainly dependent on the NF- $\kappa$ B and calcium levels. HBx prevents apoptosis when NF- $\kappa$ B is activated and induces apoptosis when NF- $\kappa$ B is blocked through regulation of intracellular calcium levels<sup>89</sup>. NF- $\kappa$ B activation and both inhibition and induction of apoptosis followed by hepatocyte regeneration can increase the risk of HCC<sup>90</sup>.

HBx can facilitate angiogenesis stabilizing and upregulating HIF1 $\alpha$  and ANG2 and alter cell cycle progression either inducing G1, S and G2/M progression in differentiated cells or promoting cell entry in G1 and S phase in less differentiated cells<sup>91-93</sup>. Finally, HBx induces also the expression of hepatic stem cells markers as EPCAM, NANOG, OCT4, MYC and the  $\beta$ -catenin<sup>94</sup>.

The natural history of HCC development during chronic HBV infection does not robustly support a direct strong link between viral proteins, including HBx, and carcinogenesis. Indeed, HCC is a late event in chronic HBV infection and the incidence is higher in patients with advanced liver disease and more aggressive hepatitis. It is therefore, more likely that HBV proteins produce an intracellular and extracellular environment that increase the risk of HCC. This is also supported by studies on HBx transgenic mice showing that these animals were more sensitive than the controls to low levels of carcinogenic agents and by clinical data demonstrating that HCC incidence is particularly higher in special HBV populations such as patients with liver cirrhosis, family history of HCC, diabetes or exposure to aflatoxin B1<sup>95,96</sup>.

The mechanisms implicated in HBV-related carcinogenesis are summarized in **Fig. 5**.

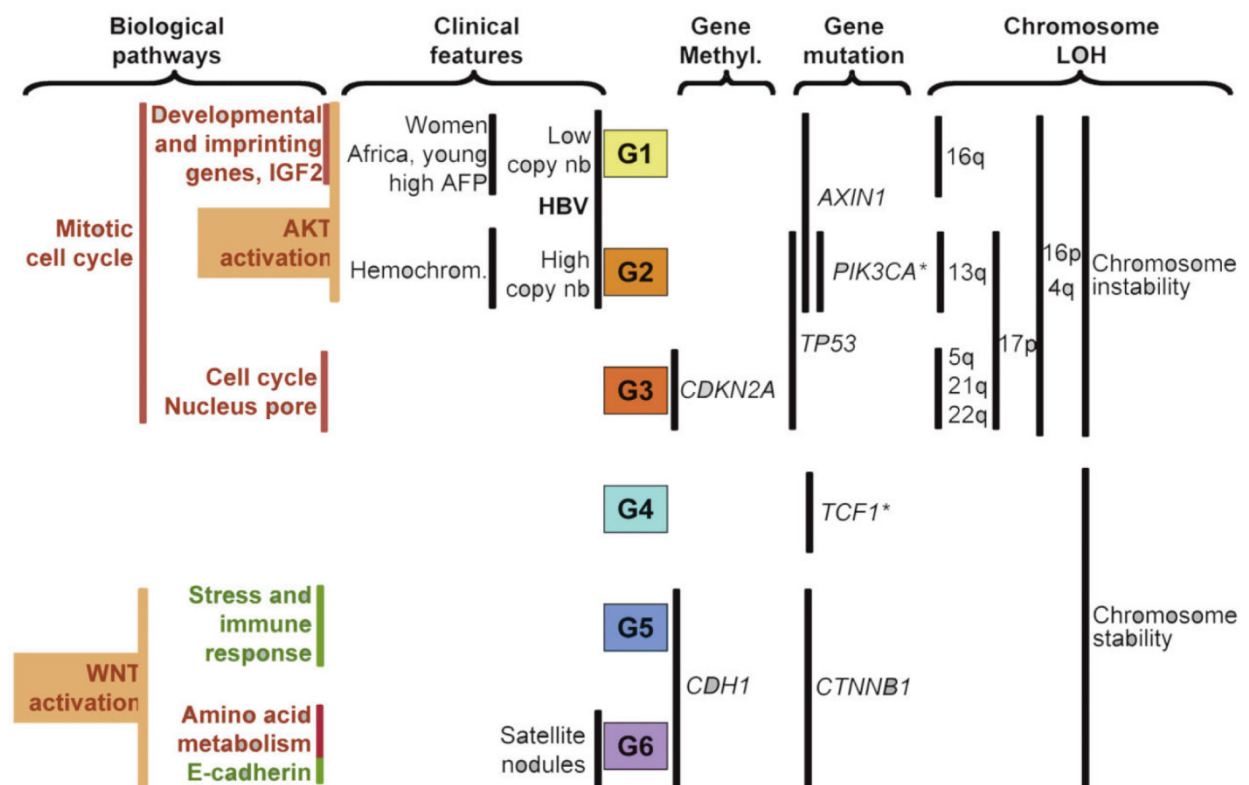


**Figure 5. Mechanisms of HBV-related hepatocarcinogenesis.** From Bouchard et al.<sup>61</sup>.

## 1.6 HETEROGENEITY, MOLECULAR AND HISTOLOGICAL HCC CLASSIFICATION

HCC is a highly heterogeneous cancer and several classification systems have been developed to describe histological or transcriptomic differences and predict patients' prognosis.

Boyault et al. identified two major distinct transcriptomic groups that can furtherly be divided in 3 subgroups (G1-G6) (Fig. 6)<sup>97</sup>.



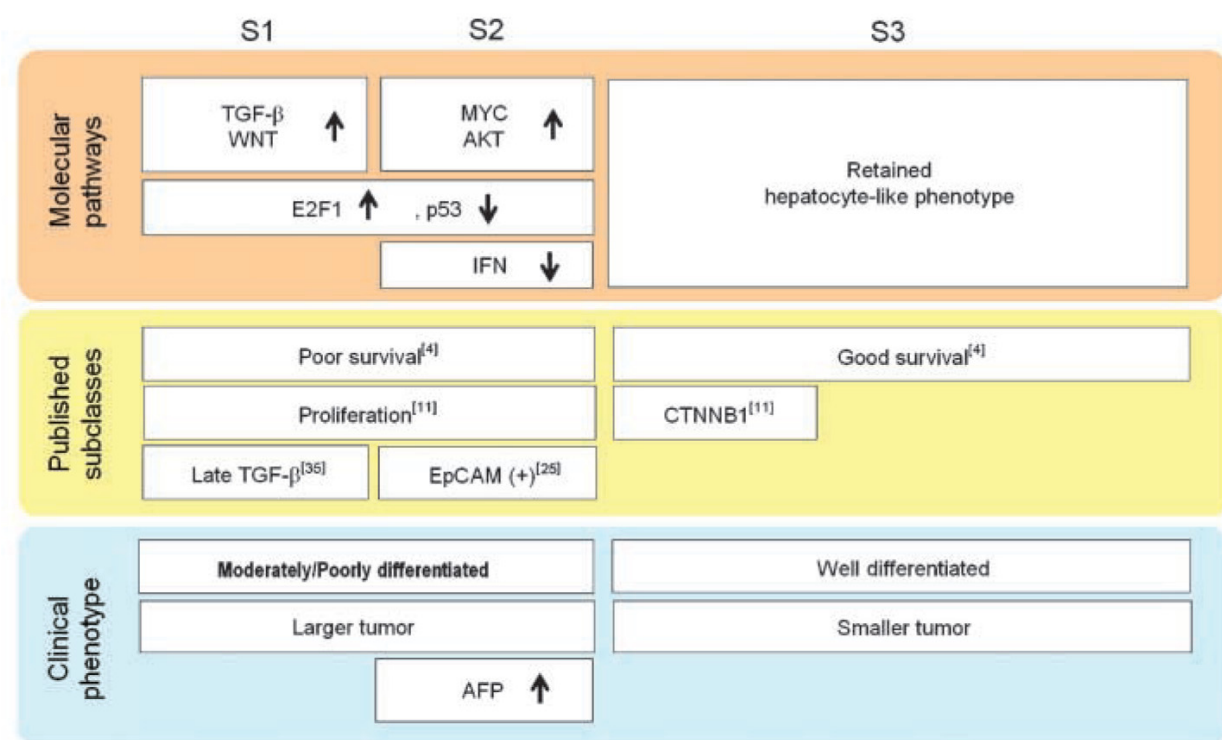
**Figure 6. G1-G6 classification of HCC. From Boyault et al.<sup>97</sup>.**

The first group (G1-G3) is enriched in proliferation, DNA metabolism and cell-cycle genes. In the G1 subgroup, parentally imprinting genes were overexpressed together with IGF2, patients were younger, native of Africa and had usually HBV infection with a low viremia and high alpha-fetoprotein (AFP) levels. The G2 subgroup included HBV infected patients with high HBV DNA levels and tumors with vascular invasion, HCC with *TP53* mutations but also a rare subgroup of tumors bearing a *PIK3CA* mutations. Both G1 and G2 subgroups show activation of AKT pathways. G3 subgroup is characterized by overexpression of genes controlling cell cycle such as *CCNA2*, *CDC6*, *MCM2*, genes encoding for proteins of the nuclear pore and tumors with hypermethylation of *CDKN2A* promoter.

The second group (G4-G6) is a low proliferation class that shows chromosomal stability and is associated with a well-differentiated phenotype<sup>97</sup>. G4 is a heterogeneous group with well-differentiated tumors closely resembling to normal liver tissues. Both G5 and G6 subgroups included tumors with  $\beta$ -catenin mutation and activation but G6 tumors had a downregulation of E-cadherin protein and a greater overexpression of  $\beta$ -catenin that is mainly localized in the cytoplasm and nucleus. G5 HCCs gene expression was enriched with genes involved in the immune and stress response.

A transcriptomic classification developed by Hoshida et al. divide HCC in 3 molecular subclasses (S1-S3) (**Fig. 7**)<sup>98</sup>. The hallmark of the S1 tumors is the TGF- $\beta$  and WNT/ $\beta$ -catenin pathways activation not directly linked to  $\beta$ -catenin (*CTNNB1*) mutations. Indeed,  $\beta$ -catenin mutations are more likely found in the S3 class. TGF- $\beta$  induces modification in the cellular  $\beta$ -catenin localization

that is moving from the plasmatic membrane to cytoplasm forming perinuclear aggregation. S2 subclass includes HCC with a relative suppression of IFN target genes, enrichment in MYC target genes, positivity for EPCAM, higher AFP levels and AKT activation. S3 HCCs are well-differentiated tumors characterized by a different activation of p53 and p21, higher expression of metabolic hepatocytes genes and *CTNNB1* mutations. The S1-S3 classification correlates also with HCC phenotypes at histology. S1-S2 tumors are larger, moderately/poorly differentiated, with cholangiocarcinoma or stem cell-like phenotype while S3 HCC are smaller and more differentiated<sup>99</sup>.



**Figure 7. HCC S1-S3 subclasses.** From Hoshida et al.<sup>98</sup>

At the histology, HCC is characterized by increased nucleus-cytoplasm ratio, cell density, abnormal vascularization with loss of sinusoids, pseudoglandular formation and stromal invasion. HCC differentiation is graded into 3 categories according to the WHO recommendations or into 4 categories according to Edmonson and Steiner classification<sup>100</sup>. HCC can also be classified according to the architectural structure (microtrabecular, macrotrabecular, pseudoglandular or compact) and cytological pattern (clear cell, fatty changes, cholestasis, pleiomorphic or spindle cells)<sup>101</sup>.

The *CTNNB1* mutated HCC are well-differentiated, microtrabecular, pseudoglandular tumors with intratumor cholestasis, low immune infiltration and retained expression of hepatocellular proteins such as *APOB*<sup>102-104</sup>.

The macrotrabecular-massive HCC (MTM-HCC) is a novel subtype that is more frequent in HBV-related HCC and in tumors with higher AFP. It is characterized by macrotrabecular architectural pattern (>6 cells thick) and frequently associated with satellite nodules and macrovascular and/or microvascular invasion<sup>102,105</sup>. Angiogenetic pathways are activated with overexpression of VEGFA and ANGPT2, chromosomal instability is frequently detected and *TP53* mutations and *FGF19* amplifications are frequently observed. On the transcriptomic level MTM HCC is frequently classified as G3<sup>102</sup>.

The steatohepatitic HCC is characterized by histological features of steatohepatitis with cell ballooning, peri-cellular fibrosis, lobular inflammation and Mallory-Denk bodies<sup>106</sup>. It is associated with the subclass G4 and its cancer-associated fibroblasts upregulate IL-6 and the JAK/STAT

pathway with overexpression of C-reactive protein (target gene of the JAK/STAT pathway)<sup>102,107</sup>.

This variant is more frequent in patients with NASH<sup>106</sup>.

The scirrhous HCC presents abundant and dense stroma embedding cancer cells and expresses several progenitor and cancer stem genes such as *KRT17*, *KRT19*, *THY1*, *PROM1*, activation of TGF- $\beta$  and EMT pathways and it is thought to represent an intermediate stage between HCC and cholangiocellular carcinoma<sup>101,108</sup>.

The progenitor HCC (PG-HCC) has poor prognosis, high expression of CK19 and frequently bears *TP53* mutation as well as overexpression of progenitor cell markers as EPCAM and PROM1 and it is associated with HCC G1-G3 subtype and S2 subclass<sup>109,110</sup>.

Fibrolamellar carcinoma (FLC) is a rare subtype (<1%) and is more frequent in young patients with no history of chronic liver disease and fibrosis<sup>101</sup>. Histologically, it is characterized by large eosinophilic neoplastic cells positive for CK7 and HepPar1 and a dense fibrotic stroma<sup>101</sup>. Importantly, FLC appearance can be present in the entire tumor specimen (pure FLC) or part of it (mixed FLC with areas resembling classical HCC). On the clinical and molecular levels, pure and mixed FLCs are different entities. Pure FLCs are more frequently characterized by the production of a DNAJB1-PRKACA chimera protein that is considered a key oncogenic driver of these tumors<sup>111-113</sup> and also has a unique transcriptomic profile with overexpression of ERBB2 and neuroendocrine genes<sup>114</sup>. Compared with mixed FLC, patients with pure FLC are also younger, have lower AFP levels, higher rates of lymph node involvement and better overall



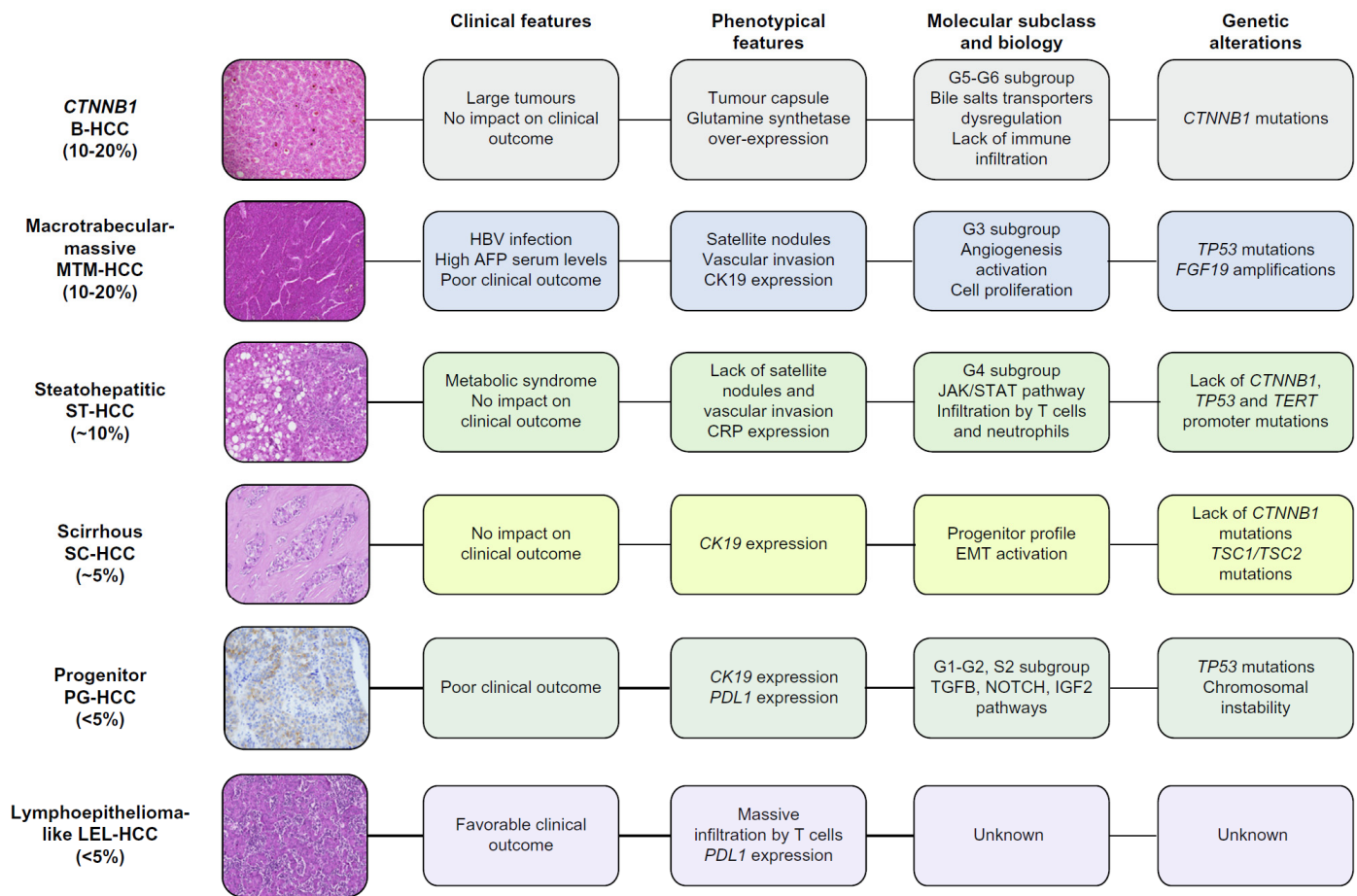
survival (OS) after resection<sup>115</sup>. The expression of both biliary, hepatocyte and pancreatic markers suggests that these tumors originated from a multipotent stem cell located in the biliary tree<sup>116</sup>.

Lymphoepithelioma-like carcinoma (LEL-C) resembles phenotypically to the nasopharyngeal lymphoepithelioma with a pronounced stroma and immune infiltration. It is a rare subtype with a better prognosis and the immune infiltrate is probably representing an antitumor response and shows a high prevalence of CD8+ T cells and increased PD1 and PDL1 expression<sup>117-119</sup>.

Combined hepatocellular-cholangiocarcinoma (cHCC-CC) is a rare tumor showing both hepatocyte and biliary features and it is considered a biliary-derived tumor representing very likely a variant of cholangiocarcinoma<sup>120</sup>.

Other rare HCC subtypes include the sarcomatoid HCC characterized by spindle cells and sarcoma-like features, the chromophobe with abrupt anaplasia HCC associated with alternative lengthening of a telomere and mutations in *ATRX*, *H3F3A* and *DAAX*<sup>121,122</sup> and granulocyte-colony stimulating factor (G-CSF) HCC defined by intratumor production of G-CSF and subsequent neutrophils infiltration.

An integrative view of the histological and transcriptomic classification of the main HCC subtypes is shown in **Fig. 8**.



**Figure 8. Integrative view of clinical and histological features, transcriptomic classification and genetic alterations of the main HCC subtypes. Adapted from Calderaro et al.<sup>101</sup>.**

The clinical relevance of all these classifications has yet to be determined. MTM-HCC, PG-HCC and sarcomatous HCC have worse prognosis and lower response to the surgical and ablative treatment. HCC with high immune infiltrate or the presence of intratumour tertiary lymphoid structures have demonstrated to be associated with a better prognosis and lower risk of recurrence after liver resection<sup>123,124</sup>. However, in clinical practice, histological or molecular HCC subtypes do not guide the treatment nor are used yet to define patient prognosis. In the era of

immunotherapy, the study of immune infiltrate in HCC will become more and more relevant. PDL1 expression is found in 75% of infiltrating immune cells and in 15% of the HCC cells<sup>119</sup>. The majority of PDL1-positive cells are macrophages and their presence is correlated with active intratumor immune response, higher prevalence of CD8 T cells and increased levels of granzyme, perforin and IFN $\gamma$ <sup>125</sup>.

Histological, molecular and genetic classification of HCC need to be implemented in the clinical setting and ideally used to assess patient's prognosis or to allocate treatment. A recent paper from Nault et al., has evaluated the clinical impact of genomic diversity in HCC in early and advanced cancer and shown that G1-G6 transcriptomic classification and a molecular prognostic 5-gene score are differently distributed among HCC stages and allocated HCC treatments. The 5-gene score is based on the combined expression of HN1, RAN, RAMP3, KRT19 and TAF9 and is associated with OS in patients treated by liver resection, ablation or having advanced HCC<sup>126,127</sup>. Late stage HCCs are more frequently G3 tumors and have poor prognostic score, increased proliferation and lower differentiation<sup>127</sup>.

Molecular characterization of tumors is already used to guide therapy in patients with colon and breast cancer. More data are needed to characterize heterogeneity and drivers of HCC and ultimately develop personalized approaches for HCC treatment.

## 1.7 TREATMENT OF LIVER FIBROSIS

Fibrosis is the result of an abnormal response to liver damage and is regulated by multiple cell types, pathways and complex cell-to-cell interactions. Currently, the only therapeutic approach for liver fibrosis consists in the treatment of the underlying chronic liver injury and/or liver transplantation.

Fibrosis is the most important risk factor for HCC development and anti-fibrotic treatments may, in theory, reduce HCC risk. Anti-fibrotic drugs were tested so far mainly in patients with NASH or primary sclerosing cholangitis (PSC) since no effective therapies are currently available for these diseases.

Apoptosis signal-regulating kinase 1 (ASK1) activates Raf-independent JNK and p38 pathways in response to various cellular stresses. Lysyl oxidase homolog 2 (LOXL2) is an enzyme responsible of the first steps in the formation of crosslinks that stabilize collagen and elastin in the extracellular matrix<sup>128</sup>. Selonsertib is a ASK1 inhibitor that was tested alone and in association with simtuzumab, a monoclonal IgG4 targeting LOXL2, for the treatment of NASH and liver fibrosis in randomized clinical trials (RCT). In a phase II randomized open-label trial involving patients with NASH stage 2 or 3, selonsertib 16 mg with or without simtuzumab for 24 weeks reduced liver fibrosis in 43% of patients and was superior to simtuzumab alone<sup>129</sup>. Phase III randomized placebo-controlled trials of selonsertib alone in patients with NASH at stage F3 (STELLAR-3) or F4 (STELLAR-4) did not meet the primary endpoint of  $\geq 1$ -stage histologic improvement in fibrosis without worsening of NASH after 48 weeks of treatment<sup>130</sup>. Phase III

randomized placebo-controlled trial of simtuzumab alone in patients with liver cirrhosis did not show any significant improvement of fibrosis stage, progression of fibrosis or liver-related clinical events<sup>131</sup>. Simtuzumab failed also to show any benefit in a randomized placebo-controlled phase IIb trial including patients with PSC<sup>132</sup>.

Cenicriviroc is an inhibitor of cytokines CCR2/CCR5 that are important players in liver fibrosis. The effect of cenicriviroc was studied in a phase IIb randomized placebo-controlled trial involving patients with NASH stages F1-3 and NAFLD activity score > 4 which showed that more patients improved in fibrosis by  $\geq 1$  stage after 1 year of cenicriviroc (20% vs. 10%)<sup>133</sup>. The primary endpoint of NAFLD activity score (NAS) improvement and resolution of NASH was not met. A phase III trial (AURORA, NCT03028740) is currently ongoing.

Peroxisome proliferator-activated receptors (PPARs) form a group of nuclear receptors that transcriptionally regulate various cell function and metabolic processes to maintain energy homeostasis<sup>134</sup>. A dual PPAR $\alpha/\delta$  agonist, elafibranor, was tested in phase IIb trial and, even though did not meet the pre-specified primary endpoint, a post-hoc analysis showed a resolution of NASH and reduction of fibrosis stage in patient with NAS $\geq 4$ <sup>135</sup>. A follow-up phase III trial is ongoing (RESOLVE-IT, NCT02704403). Elafibranor has also been investigated in patient with primary biliary cholangitis and an inadequate response to ursodeoxycholic acid and demonstrated to significantly reduce alkaline phosphatase levels and pruritus<sup>136</sup>.

Farnesoid X receptor (FXR) is a nuclear receptor activated by bile acids and it is involved in regulating immune responses, lipid and glucose metabolism as well as insulin signaling<sup>137,138</sup>. In

a phase II trial, the FXR agonist obethicolic acid (OCA) demonstrated to reduce NASH histological features and liver fibrosis<sup>139</sup>. A phase III trial (REGENERATE) is ongoing and preliminary results suggest that OCA treatment is associated with improvement of fibrosis by  $\geq 1$ -2 stages compared to placebo<sup>140</sup>. Moreover, OCA demonstrated to improve fibrosis in 46% of patients with primary biliary cholangitis with inadequate response to ursodeoxycholic acid.

Liver fibrosis treatments are a major medical need. Some of the NASH drugs in the pipeline showed a slight improvement of fibrosis in non-cirrhotic patients with NAS > 4. Novel therapeutic targets and treatment strategies are needed for patients with advanced fibrosis independently of the etiology.

## 1.8 HCC CHEMOPREVENTION

Aetiology-specific HCC prevention strategies rely on the treatment of the underlying liver disease.

Antiviral therapies showed to reduce the risk of HCC even though the risk is not completely eliminated and is higher in patients with liver cirrhosis<sup>2,12,141</sup>. For HBV, a vaccine is available and is an effective primary prevention measure to reduce HBV infection and incidence<sup>142</sup>.

Besides anti-fibrotic drugs already discussed in the previous paragraph - whose chemopreventive effect is not yet demonstrated - generic chemoprevention therapies such as metformin, statins, aspirin and non-steroidal anti-inflammatory drugs (NSAIDs) have been investigated.

Statins are drugs with pleiotropic effects. They inhibit oncogenic pathways such as Myc, Akt, NF- $\kappa$ B and TNF-mediated IL-6 production and induce p53-dependent apoptosis. Statins also demonstrated to reduce activation of stellate cells and portal hypertension via non-canonical hedgehog signaling<sup>143</sup>. Statin treatment in diabetic HBV, HCV or NAFLD patients was associated with lower progression to cirrhosis, liver decompensation, mortality and HCC development<sup>144,145</sup>. Since data are derived from retrospective cohorts, randomized controlled studies are needed before recommend statin for HCC chemoprevention.

Type 2 diabetes increases the risk of HCC<sup>146</sup>. Metformin improves insulin sensitivity, inhibits gluconeogenesis, mTOR pathways, angiogenesis and cell cycle progression and induces apoptosis<sup>143</sup>. Clinical data show that metformin was associated with a lower HCC incidence in diabetic patients with HBV or HCV infection, cirrhosis or obesity. Since no data are available in

non-diabetic populations, metformin use outside the frame of diabetes care could not yet be suggested<sup>147,148</sup>.

Aspirin and anti-inflammatory agents could reduce the risk of HCC acting on chronic inflammation. Cyclooxygenase-2 (COX2) dependent prostaglandins have been shown to be upregulated in chronic liver disease<sup>149</sup>. A pooled analysis of 2 prospective US cohorts showed that the risk of HCC decreased significantly with long-term regular aspirin use (>5 years) and no preventing effect was associated with NSAIDs<sup>150</sup>. Besides the irreversible inhibition of COX2, aspirin has also been shown to inhibit platelet thromboxane, subsequently leading to the inhibition of sphingosine-1-phosphate S1P, a lipid molecule that promotes HCC proliferation<sup>151</sup>. Recent preclinical data showed that platelet adhesion and activation play a role in NASH and HCC development and that the combination of aspirin plus clopidogrel is a potential strategy for NASH treatment and HCC chemoprevention<sup>152</sup>. Potential side effects are bleeding events that may limit the long-term use of these drugs.

PI3K/AKT/mTOR pathway is a key HCC driver and regulates cell proliferation<sup>153</sup>. Inhibitors of mTOR, such as sirolimus and everolimus, are immunosuppressive drugs that have been studied in the HCC prevention post liver transplantation. In an open-label phase III trial, the use of sirolimus as immunosuppressive agent after liver transplantation did not reduce HCC recurrence<sup>154</sup>. A subgroup analysis of this trial suggested that younger patients with lower tumor burden within Milan criteria could benefit from sirolimus<sup>154</sup>.



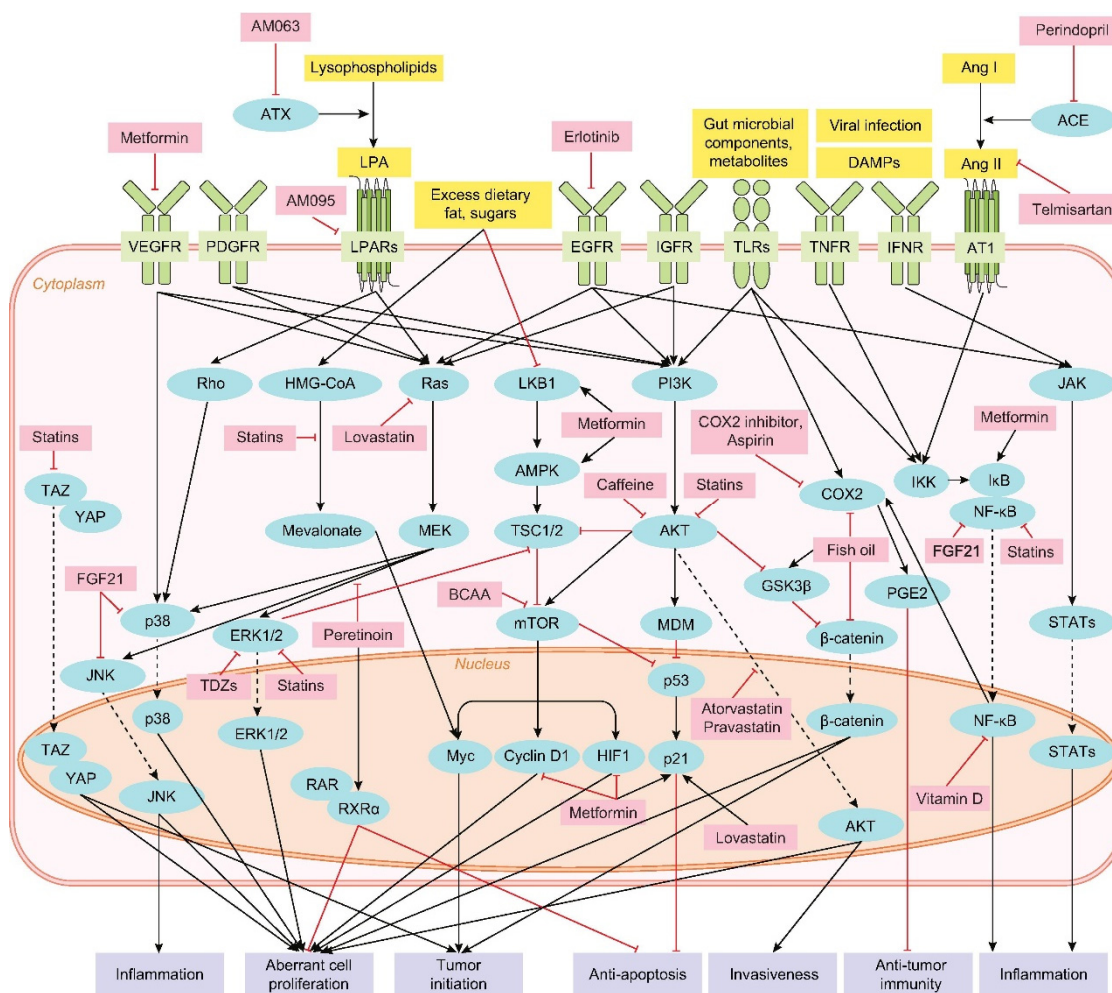
Lysophosphatidic acid (LPA) is upregulated in liver fibrosis and is involved in regulating cell survival and proliferation via G-protein-coupled receptor<sup>155</sup>. LPA receptor 1 (LPAR1) signaling was found to be a driver of HCC in cirrhotic livers and the inhibition of LPAR1 or ATX - an enzyme that converts lysophosphatidylcholine into LPA - improved liver fibrosis and reduced HCC burden in animal models<sup>153</sup>.

The renin-angiotensin system is a potential target for HCC chemoprevention. Angiotensin II promotes HSC survival<sup>156</sup> and, in animal models, ACE inhibitors and angiotensin II type 1 receptor blockers reduced liver fibrosis and HCC<sup>157</sup>. In the clinical setting, the use of ARBs was associated with lower fibrosis progression in NASH and in lower rates of HCC recurrence after curative HCC treatments<sup>158-160</sup>. In a prospective study, the ACE inhibitor perindopril administered together with vitamin K2 after curative treatments reduced serum VEGF levels and the cumulative HCC recurrence<sup>161</sup>.

Epidermal growth factor receptor (EGFR) activation in HSCs and macrophages induces liver fibrosis and HCC development and its inhibition by erlotinib suppressed HCC development in rodents<sup>162,163</sup>. Erlotinib is currently tested in a clinical trial as an adjuvant therapy after liver resection (NCT02273362).

Peretinoin, a vitamin A analogue, can inhibit WNT and PDGF signaling and showed to be an effective treatment for NASH and HCC prevention in mouse models. Clinical data from a phase III trial showed a trend toward lower HCC occurrence and higher overall survival rates in patients treated with peretinoin<sup>164</sup>.

Molecular targets for HCC chemoprevention therapies is summarized in **Fig. 9**.



**Figure 9. Molecular targets for HCC chemoprevention.** Solid line with arrowhead or bar indicates activation or inhibition, respectively. Dotted line with arrowhead indicates translocation between intracellular compartments. From Fujiwara N. et al.<sup>143</sup>.

## 1.9 HCC TREATMENT

HCC arises mainly in a context of liver cirrhosis. Hence, HCC treatment options vary according not only to the cancer stage but also to the underlying liver function. EASL and AASLD guidelines are based on the prognostic Barcelona Clinic Liver Cancer (BCLC) stage system to allocate treatment in HCC in liver cirrhosis (**Fig. 10, 11**).

Patients with very early stage BCLC 0 disease can benefit from liver resection or local ablation. Patients with early stage BCLC A can be treated with either liver resection, local ablation or liver transplantation according to the liver function and tumor burden. All these treatment options are considered curative and associated with good patient prognosis (> 5 years). Liver transplantation is also an option for patients with small tumors and end-stage liver disease or with high tumor burden disease that can be safely downstaged to meet transplant criteria.

Patients with more advanced disease can be managed with locoregional or systemic therapies with the aim to reduce disease progression and prolong survival. Intermediate stage BCLC B HCC should be treated by locoregional approaches such as chemoembolization. Advanced stage BCLC C patients can be treated with systemic therapies to improve OS of approximately 1 year. Sorafenib and lenvatinib are first-line systemic therapies. In second-line, regorafenib and cabozantinib can be used in Europe and US while nivolumab is only approved in US. The survival benefit of these drugs is limited and the side effects are significant. Sorafenib is a multi-tyrosine kinase inhibitor (TKI) and demonstrated to improve the median OS of approximately 3 months compared to placebo in a randomized controlled trial<sup>165</sup>.

Lenvatinib is a multi-TKI that targets VEGFR1-3, FGFR1-4, PDGFR $\alpha$ , RET and KIT<sup>166</sup>. Lenvatinib was tested in a non-inferiority open label trial against sorafenib showing comparable results (median OS lenvatinib vs. sorafenib, respectively, 13.6 months and 12.3 months)<sup>167</sup>.

Regorafenib is a multi-TKI targeting angiogenesis, tumor development and microenvironment<sup>168</sup>. In the RESORCE trial, a phase III RCT, sorafenib-resistant patients were treated with either placebo or regorafenib and the latter showed a survival benefit of 2.8 months (median OS in regorafenib group 10.6 months)<sup>169</sup>.

Cabozantinib is a MET, VEGFR2 and RET inhibitor already used for renal and thyroid tumors and demonstrated a median survival benefit of 2 months in second-line treatment<sup>170</sup>.

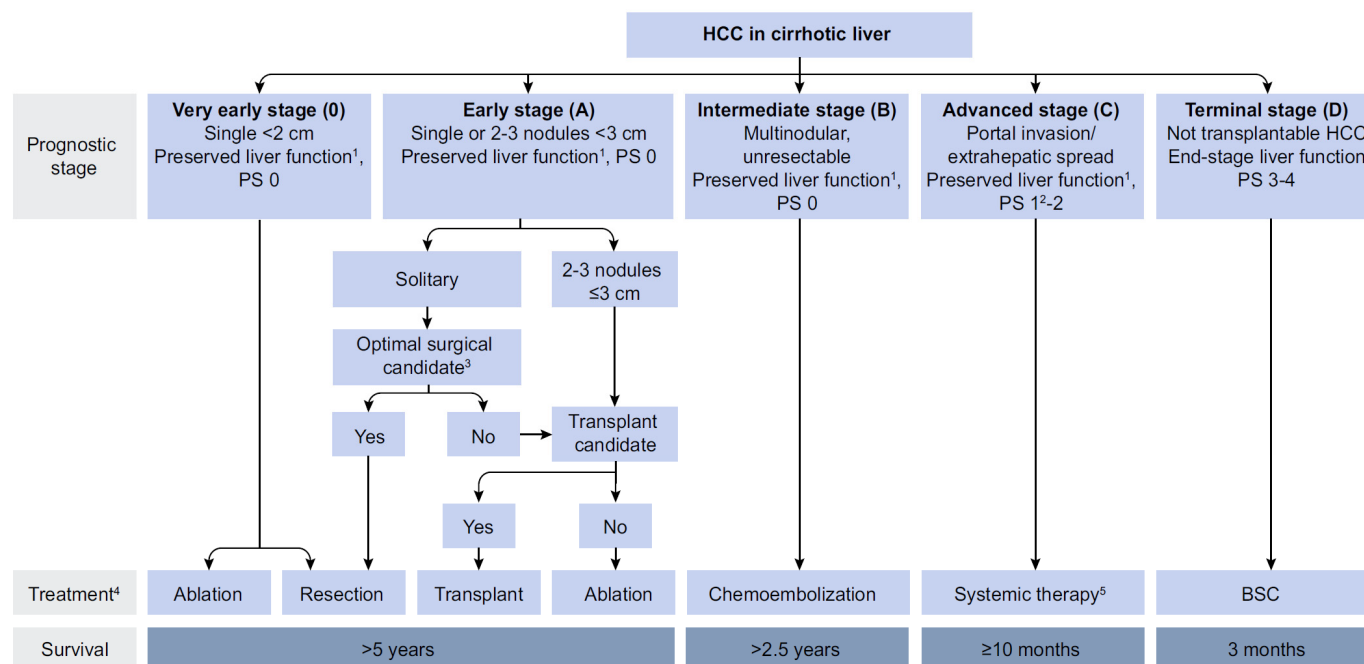
The safety profile of sorafenib, lenvatinib and regorafenib is comparable, with most patients experiencing hypertension, diarrhea, palmar-plantar erythrodysesthesia, fatigue, decreased weight and appetite<sup>165,167,169</sup>.

Nivolumab is an anti-PD1 monoclonal antibody that restores anti-tumor activity of exhausted T cells. Promising results from a phase I/II study showing an objective response rate of 20% and a median OS of 16 months in second-line granted FDA conditional approval<sup>171</sup>. However, a phase III trial (CheckMate-459) comparing nivolumab with sorafenib did not meet the primary endpoint<sup>172</sup>.

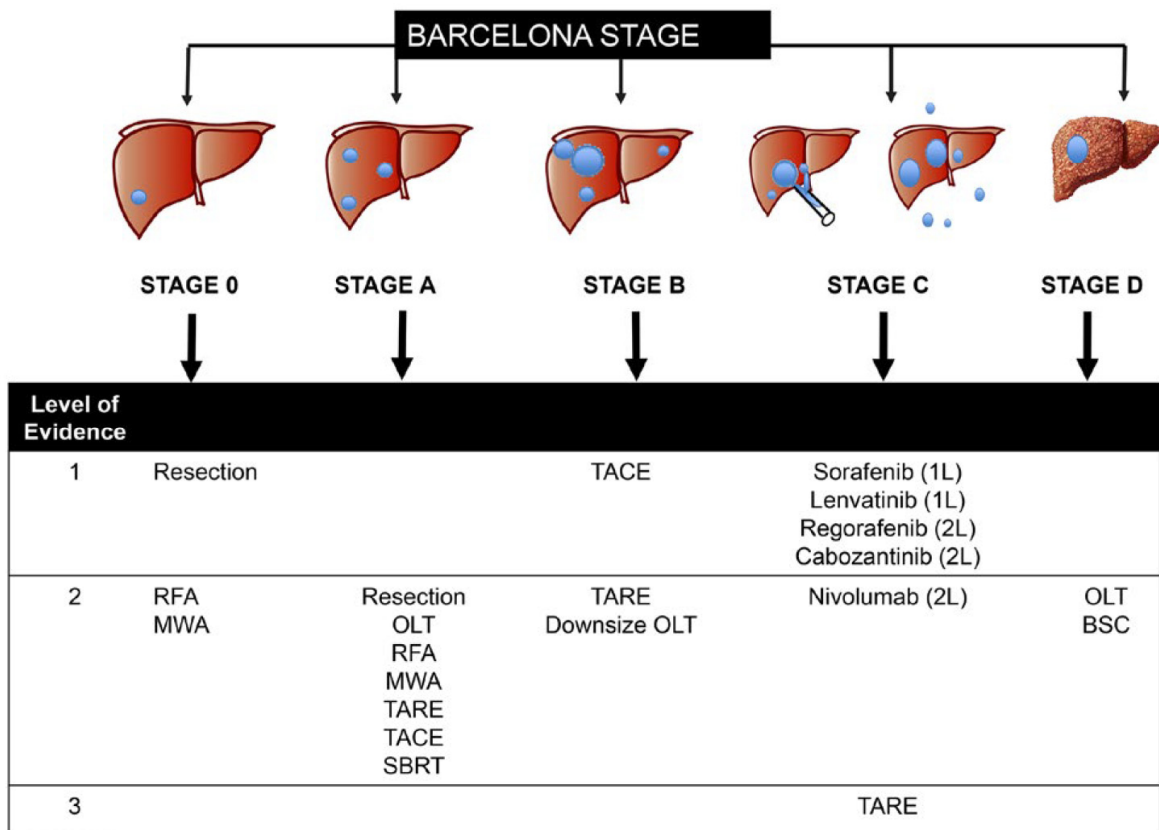
Nivolumab is currently evaluated in several trials as a single agent in adjuvant therapy (NCT03383458) and in combination with ipilimumab in previously treated patients (NCT01658878). The most common adverse events are pruritus, diarrhea, fatigue and liver enzyme elevation.

Finally, best supportive cares are indicated for patients with non-transplantable end-stage liver disease BCLC D.

A summary of EASL and AASLD treatment guidelines are reported in **Fig. 10** and **11**.



**Figure 10. EASL clinical practice guidelines – HCC treatment.** <sup>1</sup>“Preserved liver function” refers to Child-Pugh A without any ascites. <sup>2</sup>PS 1 refers to tumour induced (as per physician opinion) modification of performance capacity. <sup>3</sup>Optimal surgical candidacy is based on a multiparametric evaluation including compensated Child-Pugh class A liver function with MELD score <10, to be matched with grade of portal hypertension, acceptable amount of remaining parenchyma and possibility to adopt a laparoscopic/minimally invasive approach. The combination of the previous factors should lead to an expected perioperative mortality <3% and morbidity <20% including a postsurgical severe liver failure incidence <5%. From EASL Clinical Practice guidelines 2018<sup>173</sup>



**Figure 11. AASLD guidelines for HCC treatment.** BSC, best supportive care; MWA, microwave ablation; OLT, orthotopic liver transplantation; RFA, radiofrequency ablation; SBRT, stereotactic body radiation therapy; TACE, transarterial chemoembolization; TARE, transarterial radioembolization; 1L, first-line therapy; 2L, second-line therapy. Adapted from Heimbach et al. AASLD guidelines 2018<sup>174</sup>

## 1.10 SINGLE-CELL RNA SEQUENCING PRINCIPLES AND TECHNOLOGIES

Single-cell RNA-sequencing (scRNA-seq) allows genome-wide RNA profiling at individual cell level to study cell heterogeneity and diversity, stochastic gene expression, organ development and rare cell types.

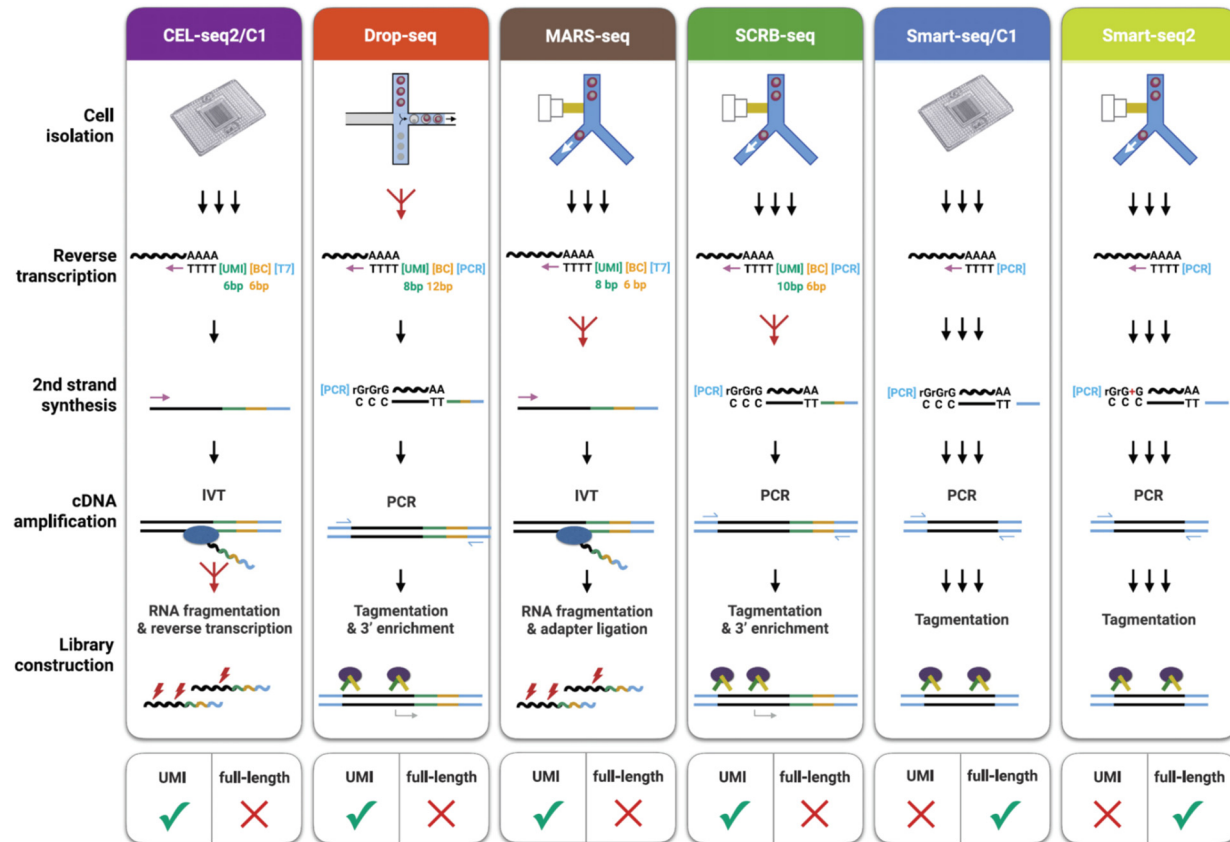
Briefly, scRNA-seq is based on the reverse transcription of RNA in complementary DNA (cDNA) and its subsequent PCR amplification or *in vitro* transcription (IVT) followed by deep sequencing.

The first steps consist in the tissues collection, dissociation and isolation of single cells. Single cell isolation can be obtained by micromanipulation, flow-activated cell sorting (FACS), laser microdissection, magnetic separation using ferrofluid nanoparticles coupled with specific antibodies (CellSearch) or microfluidic technologies<sup>175</sup>. FACS is one of the most used technique and allows the selection and enrichment of specific cell populations from heterogeneous tissues.

After cell isolation, scRNA-seq libraries are generated by cell lysis, reverse transcription into cDNA and cDNA amplification. These steps vary across the different scRNA-seq protocols (**Fig. 12**).

Following the reverse transcription of the first strand of the cDNA, the second strand is transcribed by either poly(A) tailing or template-switching methods to guarantee a non-strand specific

coverage<sup>176,177</sup>. cDNAs are then amplified by PCR or IVT and promoters and adapters are added to ensure massive parallel sequence. Usually an Illumina platform for sequencing is used.



**Figure 12. Overview of the most common scRNA-seq protocols.** Adapted from Ziegenhain et al.<sup>178</sup>.

Smart-seq and its improved version Smart-seq2 are two scRNA-seq protocols which allow the sequence of full-length transcripts independently from the strand, the splicing or the allelic origin using template-switching technologies for the reverse transcription and PCR technologies for the amplification<sup>179,180</sup>.



Since Smart-seq is an expensive procedure, different protocols have been set up to have a better ratio between coverage and costs. These protocols are based on the capture of the RNA polyA tail with the insertion of unique molecular identifiers (UMIs) and cell barcodes. Cell barcodes and UMIs allow to pool RNAs from different cells, amplify and sequence them together. The analysis of the cell barcode and the quantification of the UMIs allow to reconstruct the cell of origin and quantify gene expression.

In a study by Ziegenhain et al., the performance of the different protocols was assessed<sup>178</sup> (**Fig. 12**). Smart-seq2 and CEL-seq2 showed the highest sensitivity while Drop-seq technology is the most cost-effective method. Smart-seq2 is preferred when analyzing splicing, transcriptome annotations, viral integration and identifying sequence variants while Drop-seq technology is preferable for the analysis of a large number of cells at low coverage. CEL-seq2 would be the best compromise between running costs and cell coverage<sup>178</sup>.

ScRNA-seq comprises multiple technologies and the appropriate technique should be chosen taking into account the study design (e.g. need of cell population enrichment) and endpoints (e.g. study of rare cell types or lowly expressed genes, splicing variants analysis).

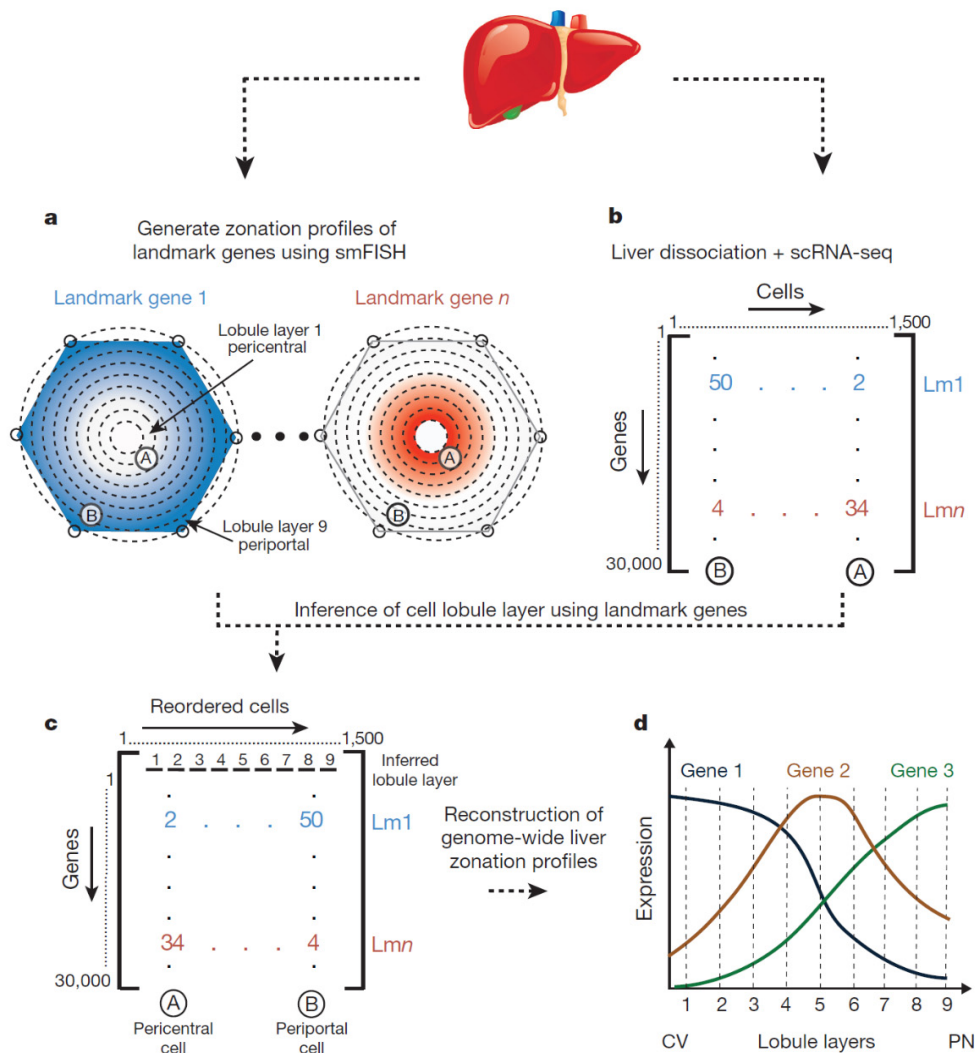
## 1.11 SINGLE-CELL RNA-SEQUENCING TO STUDY LIVER PHYSIOLOGY AND DISEASES

### 1.11.1 LIVER HETEROGENEITY AND ZONATION

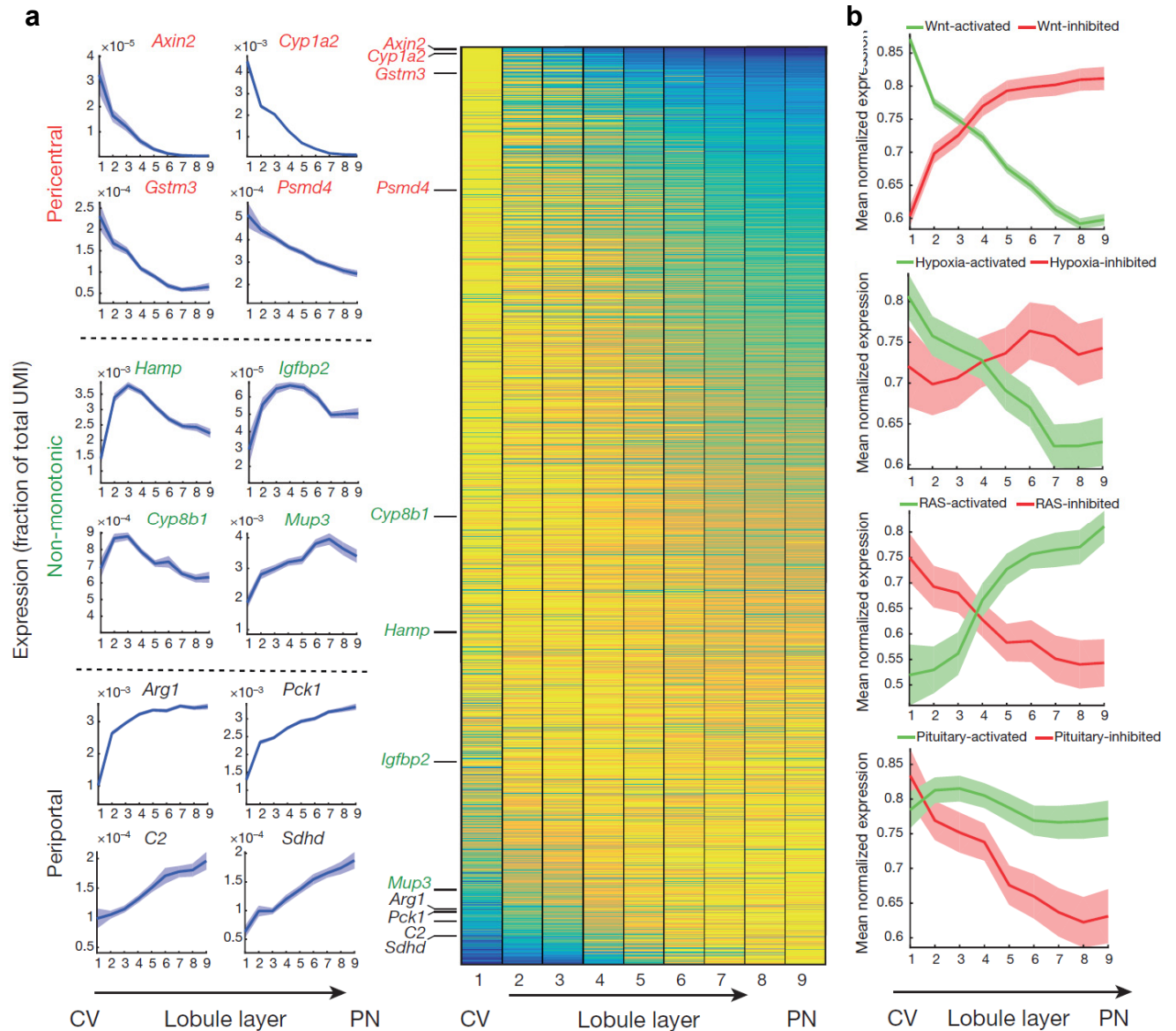
The mammalian liver is a complex organ receiving a double vascularization and subdivided in lobules furtherly organized along the porto-central venous axis. The porto-central organization and division of labor is also known as liver zonation.

ScRNA-seq coupled with single molecule RNA (smRNA) FISH has been used to investigate mouse liver zonation with a detailed genome-wide analysis. Mouse liver lobule was divided in 9 layers according to the smRNA FISH expression of already reported landmark genes<sup>181</sup>. ScRNA-seq of mouse liver hepatocytes was then performed and the zonation reconstructed according to the expression of landmark genes in the 9 layers previously identified by smRNA FISH<sup>181</sup> (**Fig. 13**). The authors observed that around 50% of mouse liver genes were significantly zoned (**Fig. 14a**). Among the genes with the highest degree of zonation there were the pericentral *Axin2* and the periportal *Sox9*. Interestingly, part of the hepatocyte zonation was determined by Wnt signaling, oxygen gradient and Ras signaling (**Fig. 14b**). On the functional level, periportal areas were enriched with genes encoding for liver-secreted proteins and ATP-demanding tasks that require high oxygen levels, while pericentral areas were enriched with genes dedicated to xenobiotic metabolism, glutathione metabolism, bile acid synthesis and proteasome. Non-monotonic genes and pathways that peaked in the mid-layer of liver lobule were also identified. Among them, *HAMP* and *HAMP2* encoding for hepcidin and *IGFBP2* encoding for a protein

regulating the levels of IGF1. IGF1 is highly expressed in the periportal layers, demonstrating the existence of an intralobular feedback loop. Finally, the entire bile acid synthesis chain was found to be spatially zonated with genes peaking in pericentral and mid-layer such as *Cyp8b1* (**Fig. 14**).



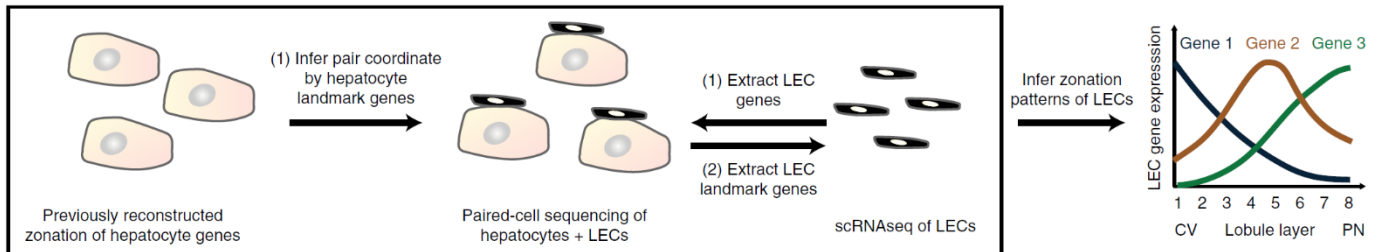
**Figure 13. Spatially resolved scRNA-seq of mouse liver.** *a*, Generation of spatial barcode defining liver lobule layers using known zoned landmark (*Lm*) genes and smFISH. *b*, ScRNA-seq of mouse liver hepatocytes. *c*, Inference of porto-central coordinates of each cell combining landmark genes. *d*, Reconstruction of spatial zonation profiles. *Lm*, landmark; CV, central vein; PN, portal node. Adapted from Halpern et al.<sup>181</sup>.



**Figure 14. Mouse liver zonation patterns of hepatocytes revealed by spatially resolved scRNA-seq.** *a*, Zonation profiles of spatially zoned genes. *b*, Signaling pathways determining liver zonation. CV, central vein; PN, portal node. Adapted from Halpern et al.<sup>181</sup>.

Recent data obtained by paired scRNA-seq demonstrated that also liver non-parenchymal cells are zoned. Halpern et al. performed a suboptimal dissociation of mouse liver tissues and

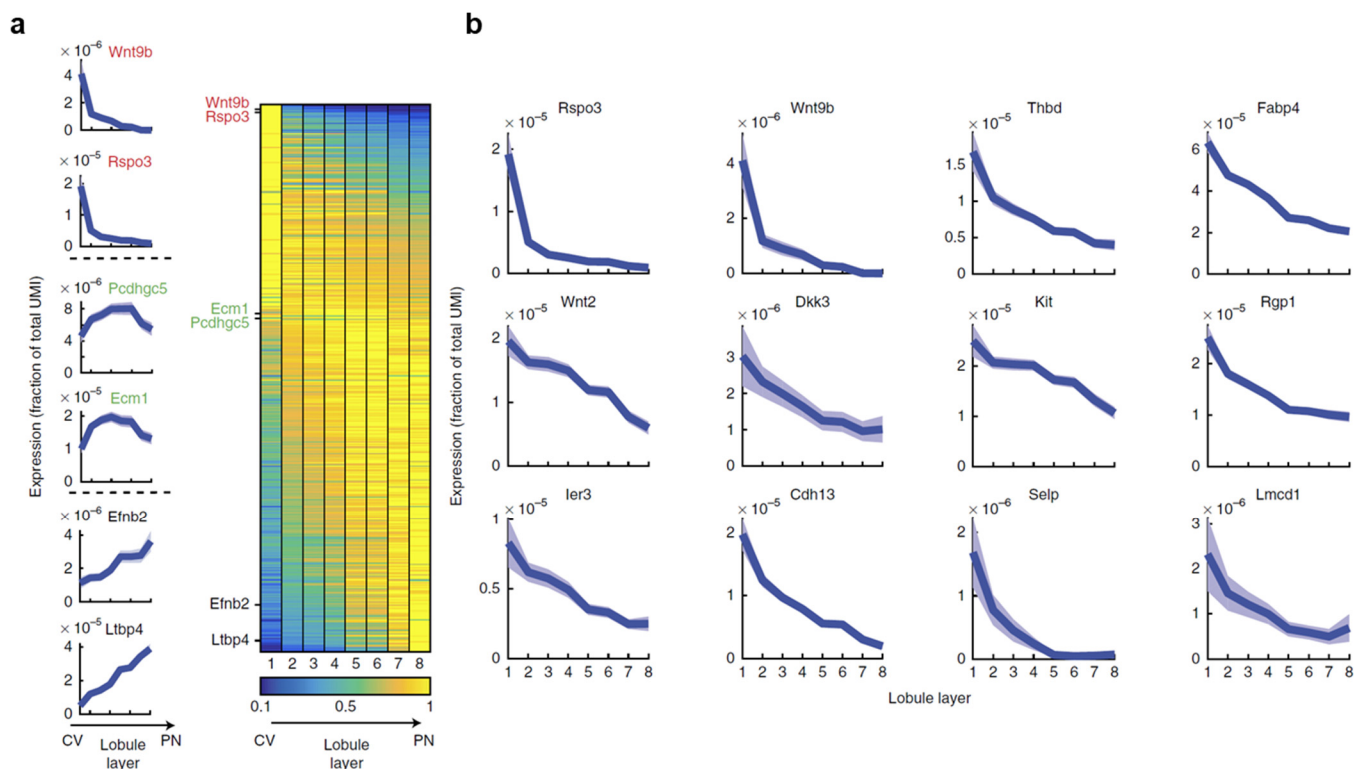
performed RNA-seq of pairs of liver endothelial cells (LECs) and hepatocytes and used the previously described hepatocyte zonation data to infer the zonation profile of LECs<sup>182</sup> (**Fig. 15**).



**Figure 15. Paired-cell scRNA-seq to infer mouse liver endothelial cells zonation.** Adapted from Halpern et al.<sup>182</sup>.

This spatial reconstruction showed that more than 30% of LECs genes are zoned. Pericentral LECs are enriched with WNT signaling genes (*WNT2*, *WNT9B*, *RSPO3*) and modulators (*DKK3*) that influence hepatocyte zonation<sup>182</sup> (**Fig. 14b** and **16**).

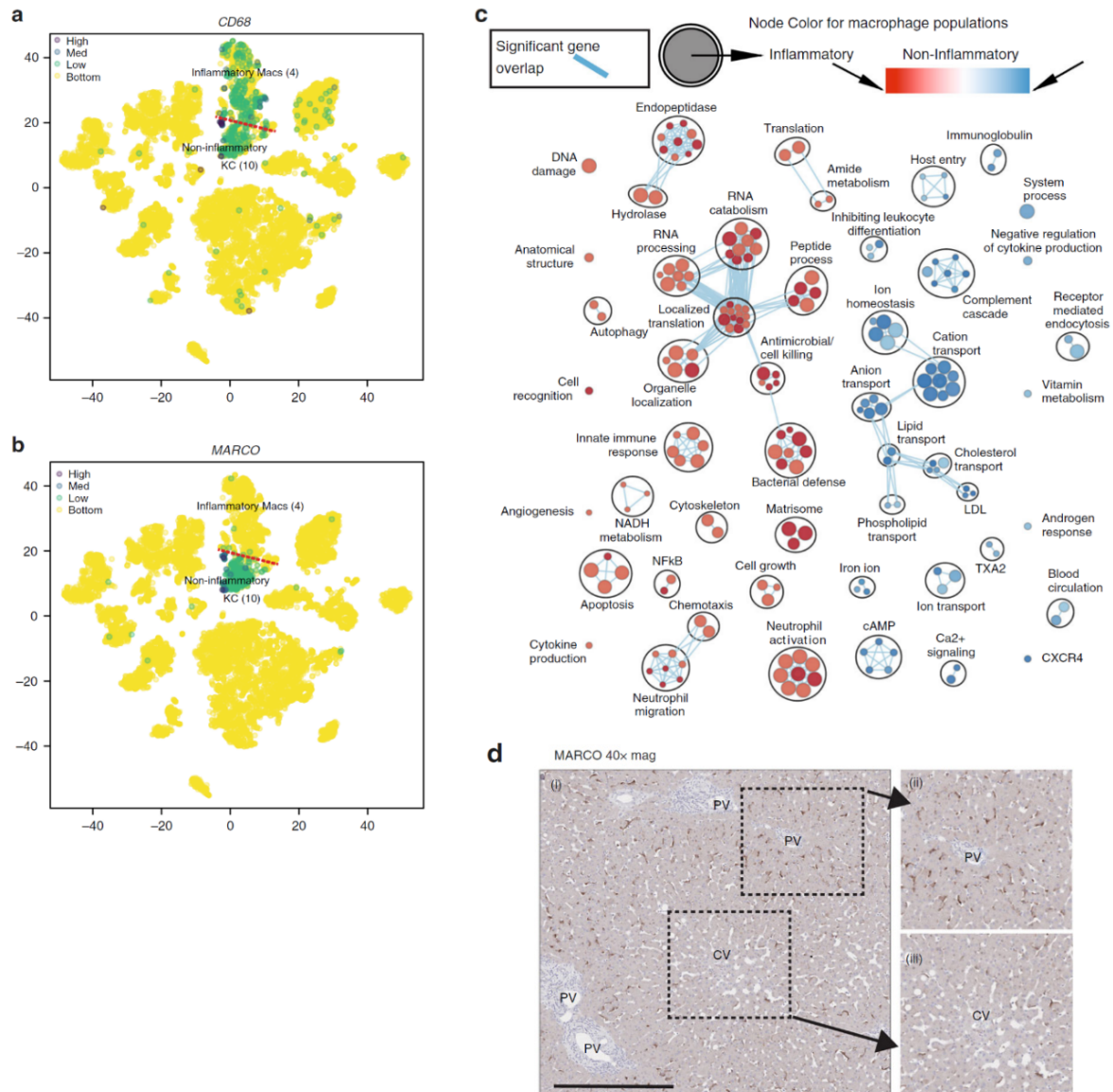
ScRNA-seq showed that in mouse liver, energetically demanding tasks, metabolites transporters and non-parenchymal cells are zoned, opposite tasks are segregated (gluconeogenesis is periportal zoned and glycolysis is pericentral), intermediate metabolites are transferred from one layer to another (e.g. bile acid biosynthesis) and spatial recycling is present (a process in which metabolites produced in periportal areas are uptaken by pericentral hepatocytes)<sup>183</sup>.



**Figure 16. Zonation of mouse liver endothelial cells obtained with spatially reconstructed paired-scRNA-seq. a, Profiles of significantly zoned genes in LECs. b, Zonation profile and signature of pericentral LECs enriched in WNT signaling genes. Adapted from Halpern et al.<sup>182</sup>**

More limited data are available on human liver. MacParland et al. build a human liver atlas using a droplet based system (10x Chromium)<sup>184</sup>. The authors compared their scRNA-seq data with the mouse liver data from Halpern et al<sup>181</sup> and found a partial correspondence between mouse and human zonation. Two distinct populations of liver macrophages were identified and characterized by either a pro-inflammatory (CD68<sup>+</sup> MARCO<sup>-</sup>) or immunoregulatory phenotype (CD68<sup>+</sup> MARCO<sup>+</sup> concentrated in the periportal areas) (**Fig. 17**). Three subsets/clusters of endothelial cells were also described: central venous liver sinusoidal cells (LSECs), periportal LSECs and portal endothelial cells. Even though MacParland et al. explored part of the human liver heterogeneity,

the low gene coverage and the absence of spatial reconstruction are important limiting factors and more detailed analysis is needed to have a full picture of human liver heterogeneity and zonation.



**Figure 17. ScRNA-seq on human liver tissue identified two distinct populations of human resident macrophages.** a, b t-SNE maps of CD68 and MARCO. c, Pairwise pathway enrichment analysis of the two macrophage subsets; pathways enriched in non-inflammatory macrophages are in blue while the ones enriched in the inflammatory macrophages are in red. Size of color nodes represent the number of genes included in the pathway. d, MARCO+ macrophages are localized in the portal areas. PV, portal vein, CV central vein. Adapted from MacParland et al.<sup>184</sup>.

### 1.11.2 HEPATOBILIARY PROGENITORS

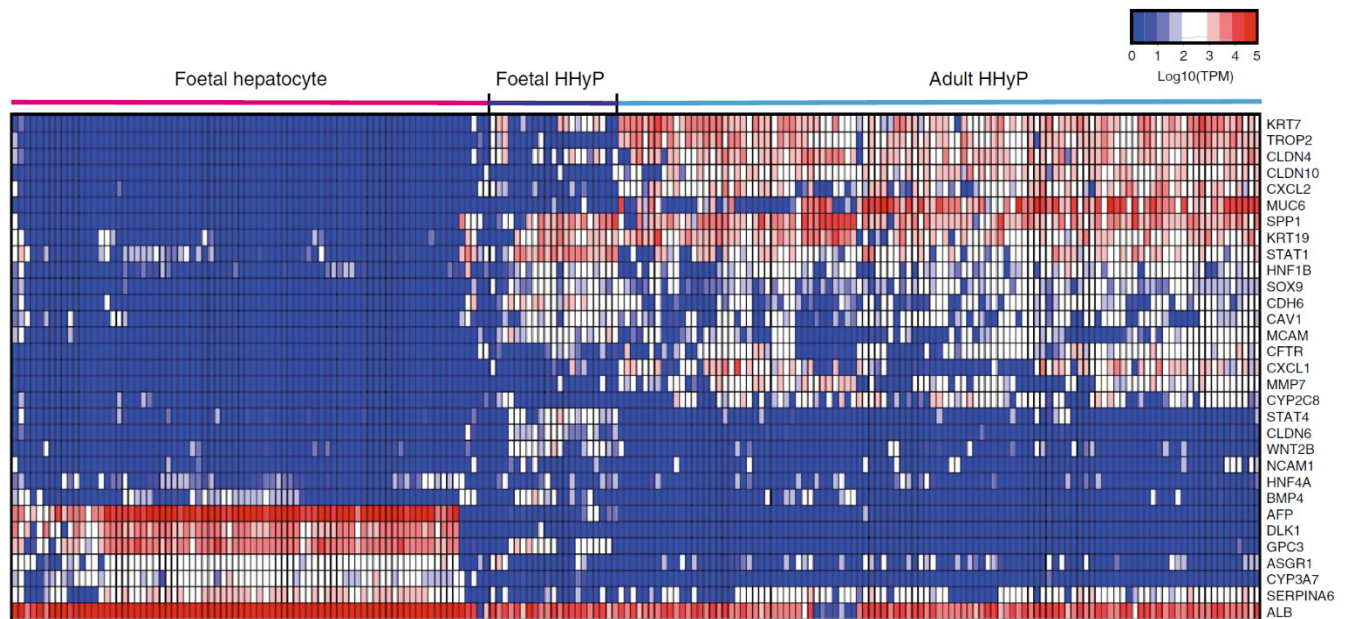
The heterogeneity of hepatobiliary precursor cells and the mechanisms of liver regeneration are poorly understood. Data from mice are conflicting and suggest either the existence of a biliary-like progenitor cell (oval cell) in the ductal areas or stem cells around the central vein or a bipotent progenitor cell that is able to differentiate into hepatocytes and cholangiocytes or a dedifferentiation process of hepatocyte and/or cholangiocyte into bipotent progenitor cells<sup>185-189</sup>. In humans, fetal progenitor cells are believed to be EPCAM<sup>+</sup> and reside in the ductal plate. In the adult liver, these cells are supposed to populate Canals of Herring and get activated upon injury<sup>190,191</sup>.

ScRNA-seq was used to capture the heterogeneity of hepatobiliary precursors in the human liver to better understand liver regeneration processes. Segal et al. studied human fetal and adult progenitors at the single-cell level and described a distinct hepatobiliary hybrid progenitor (HHyP) population restricted to the ductal plate of the fetal liver<sup>192</sup>. This HHyP belongs to the EpCAM<sup>+</sup>/NCAM<sup>+</sup> compartment and it is positive for cholangiocyte markers (*SOX9*, *HNF1B* and *KRT19*), hepatocyte markers (*ALB*, *APOE*, *TF*) and progenitor markers (*CD24*, *CD133*, *FGFR2*, *KRT7*, *SPP1*) (**Fig. 18**). An equivalent of HHyP was also identified in the adult liver in the EPCAM<sup>+</sup> compartment (**Fig. 18**) even though an experimental validation of its multipotency was not performed.

ScRNA-seq dissected mechanisms of liver regeneration in mouse<sup>193,194</sup>. The analysis of mouse EPCAM<sup>+</sup> biliary epithelial cells (BECs) showed that these cells are not homogenous and some



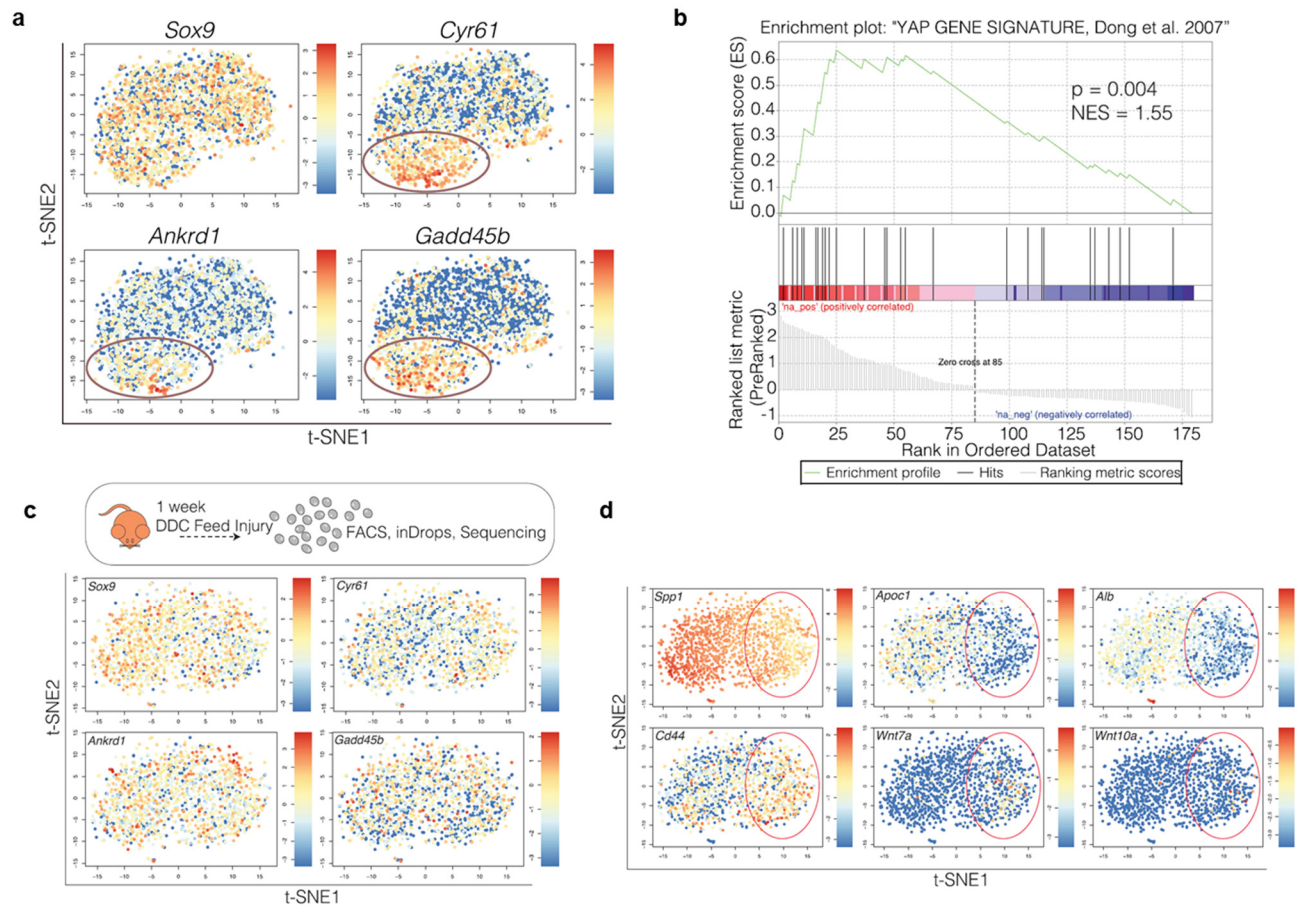
well-known markers such as *KRT19* and *HNF1B* are not uniformly expressed. One important source of the BECs heterogeneity is the different expression of YAP target genes (*CYR61*, *ANKRD1* and *GADD45B*) and of a YAP gene signature (**Fig. 19a and b**)<sup>193</sup>.



**Figure 18. ScRNA-seq comparison of fetal hepatocyte, fetal and adult HHyP cells. Adapted from Segal et al.<sup>192</sup>.**

ScRNA-seq of BECs from mice fed with 1 week of DDC diet (0.1% 3,5-diethoxycarbonyl-1,4-dihydrocollidine), a model of liver injury and biliary proliferation, showed that this YAP signature represents a dynamic inducible state<sup>195</sup>. Indeed, upon DDC, YAP-active BECs increased and YAP target genes are globally expressed (**Fig. 19c**). Moreover, in the DDC mice some BECs

expressed high level of Wnt-related genes and low hepatocyte-specific markers suggesting that these cells are in a proliferative state and are sustaining liver regeneration<sup>193</sup> (**Fig. 19d**).



**Figure 19. ScRNA-seq of mouse biliary epithelial cells.** a, t-SNE maps showing ScRNA-seq data of the EPCAM<sup>+</sup> compartment, the circle is showing a cluster with overexpression of YAP target genes. b, Gene set enrichment analysis of the YAP gene signature in the cluster highlighted in a. c-d, t-SNE maps showing scRNA-seq data of the EPCAM<sup>+</sup> cells upon DDC injury. c, YAP target genes are uniformly expressed among the cells. d, a cluster overexpressing WNT related genes and downregulating hepatocyte markers representing proliferative cells is circled. Adapted from Pepe-Mooney et al.<sup>193</sup>.

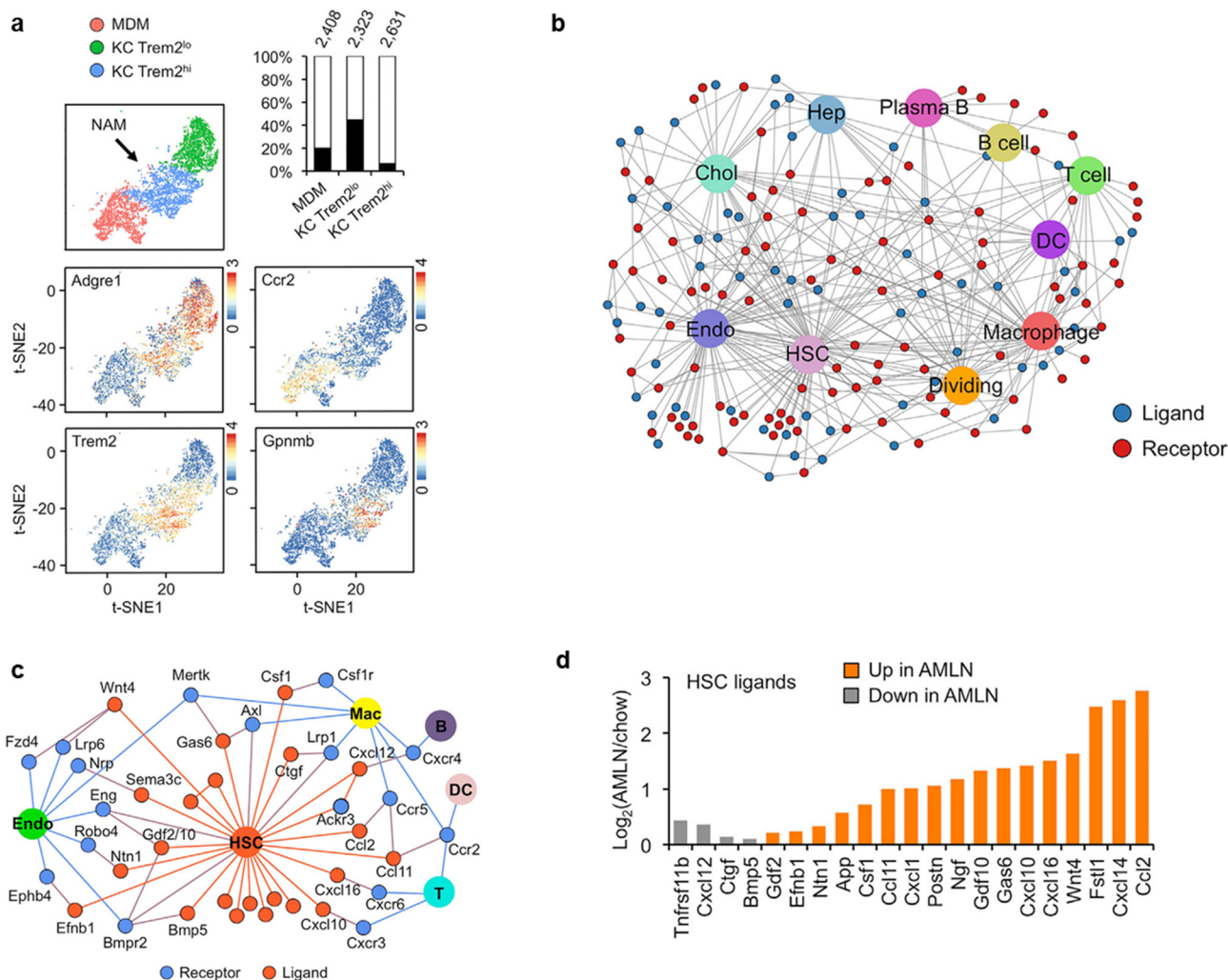
ScRNA-seq has demonstrated its ability to dissect heterogeneous cell compartments and cell states. Hitherto a complete analysis of the adult hepatobiliary progenitors and human liver regeneration processes after chronic and acute injury at the single-cell level are still needed to accurately characterize liver regeneration mechanisms, patterns and players.

### 1.11.3 NASH AND LIVER FIBROSIS

Currently, there are only few scRNA-seq data available on NASH, liver cirrhosis and fibrogenesis.

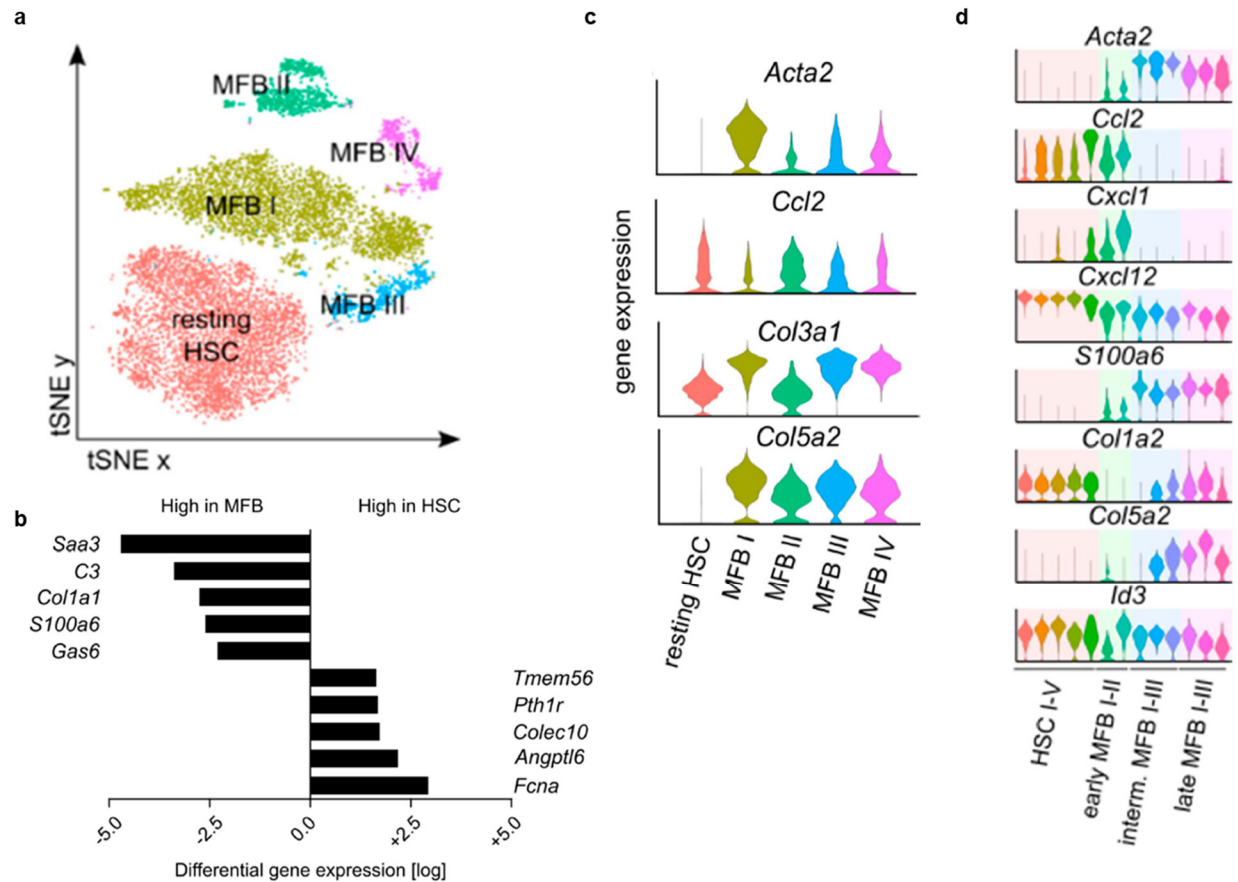
Xiong et al. analyzed at the single-cell level healthy and diet-induce NASH AMLN (amylin) mice<sup>196</sup> with a special focus on cell secretome and cell-to-cell interactions<sup>197</sup>. The authors found that Trem2<sup>+</sup> Kupffer cells are enriched during NASH and HSCs are a central hub in liver secretome network (**Fig. 20**).

Krenkel et al. performed scRNA-seq on liver MFBs from healthy mice, mice treated with repetitive injections of carbon tetrachloride (CCl<sub>4</sub>) for 3 weeks and *in vitro* activated MFBs<sup>198</sup> (**Fig. 21**). ScRNA-seq was able to distinguish between resting HSCs in healthy mouse liver and MFBs in CCl<sub>4</sub>-induced liver fibrosis and revealed the heterogeneity of these compartments. SA100A6 was identified as a new MFB marker and some clusters of MFBs were expressing chemokines or macrophages markers, suggesting macrophage transdifferentiation. The analysis of *in vitro* activated MFBs in a time-course experiment showed that only early MFBs express chemokines capable of recruiting myeloid cells such as *CXCL1* (neutrophils) and *CCL2* (monocytes)<sup>198</sup>.



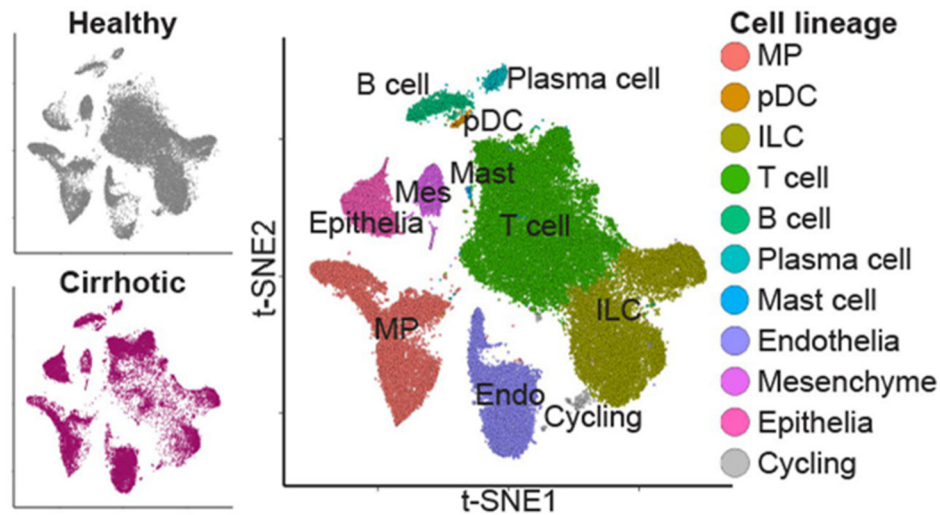
**Figure 20. Intercellular crosstalk in healthy and NASH mouse livers.** *a*, t-SNE of KCs with low (green) and high (blue) Trem2 levels. Barplot showing contributions of chow (filled) and AMLN (open) macrophages to each subpopulation and total cell counts. Feature plots of marker gene expression are shown at the bottom. *b*, Network visualization of ligand-receptor connectivity among different mouse liver cell types. *c*, The HSC secretome. Ligands exhibiting >3-fold enriched expression in the HSC cluster are shown in orange with their known receptors indicated in blue. The ligand-receptor pairs are shown when receptor expression was observed in at least one cluster (normalized UMI > 1.0) based on the scRNA-seq dataset. *d*, Regulation of stellate gene expression in NASH. Average expression values from chow and AMLN liver RNA-seq dataset were used. Adapted from Xiong et al.<sup>197</sup>





**Figure 21. ScRNA-seq data of mouse HSC and MFB and in vitro MFB.** a, ScRNA-seq of HSC from healthy mouse and MFB from CCl<sub>4</sub>-induced liver fibrotic mice showing heterogeneity in MFB population. b, Differentially expressed genes between HSC and MFB. c, MFB subpopulation are different in terms of collagen and chemokine production. d, ScRNA-seq analysis of in vitro activated MFB showed that early MFB express high level of chemokines such as CCL2, CXCL and CXCL12 and low levels of collagens genes; late MFBs have an opposite pattern. Adapted from Krenkel et al.<sup>198</sup>.

ScRNA-seq data on the fibrotic niche in human liver cirrhosis were recently published. Ramachandran et al. sequenced single cells from both healthy and cirrhotic livers and analyzed heterogeneity in fibrosis-associated non-parenchymal cells<sup>199</sup> (**Fig. 22**).

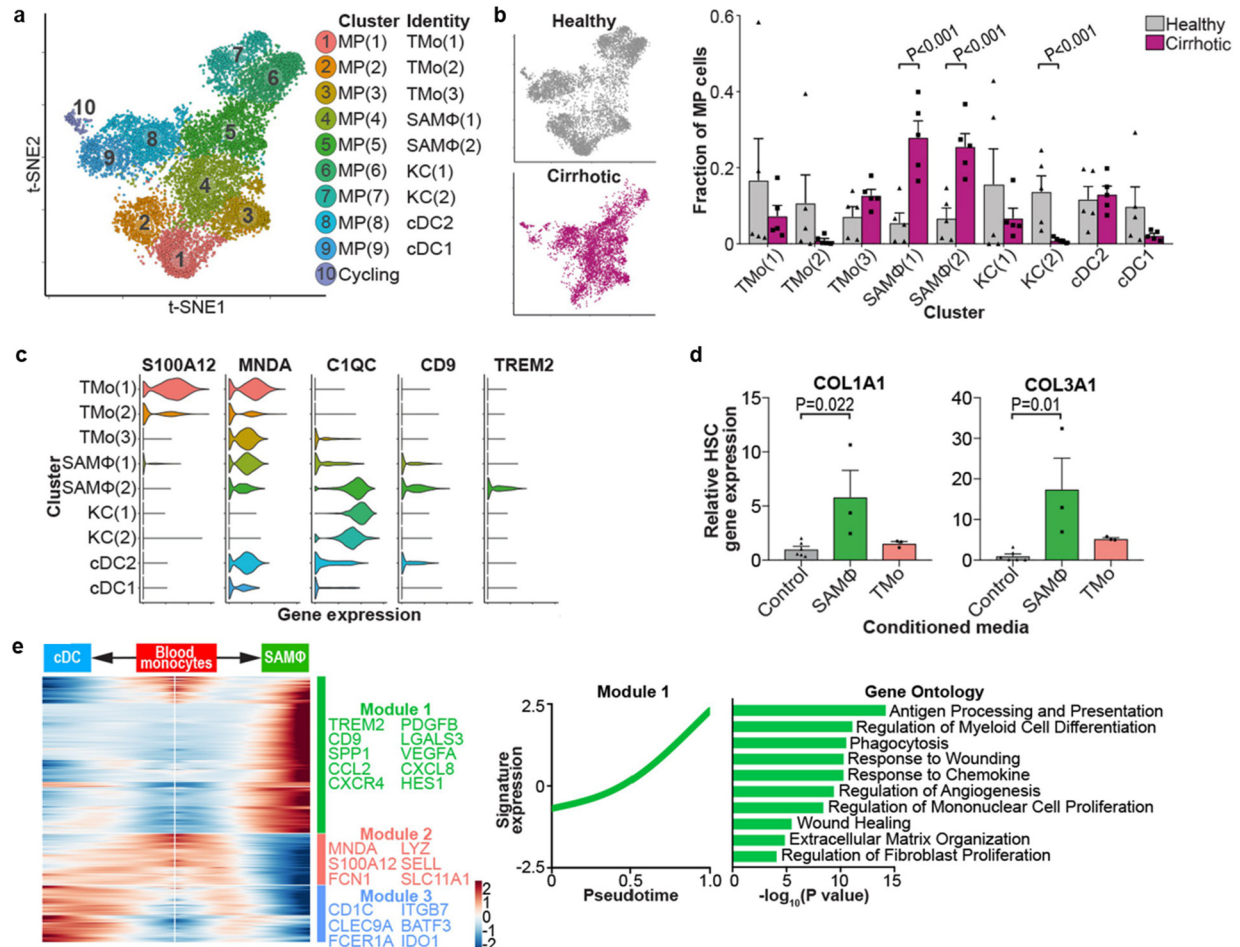


**Figure 22. ScRNA-seq of healthy and cirrhotic livers.** Adapted from Ramachandran et al.<sup>199</sup>.

The analysis of macrophages and KCs compartment revealed 10 different clusters of tissue monocyte-derived macrophages (TMO), KCs and conventional dendritic cells (cDC) (**Fig. 23a**). Some macrophages were more prevalent in cirrhotic tissues and were annotated as scar-associated macrophages (SAM $\Phi$ ) (**Fig. 23b**). The SAM $\Phi$  were characterized by the expression of TREM2 and CD9 and were able to activate HSCs (**Fig. 23c, d**).

Self-organizing mapping and pseudotime analysis revealed that SAM $\Phi$  and cDCs are derived from blood monocytes and the expression of genes implicated in antigen processing and

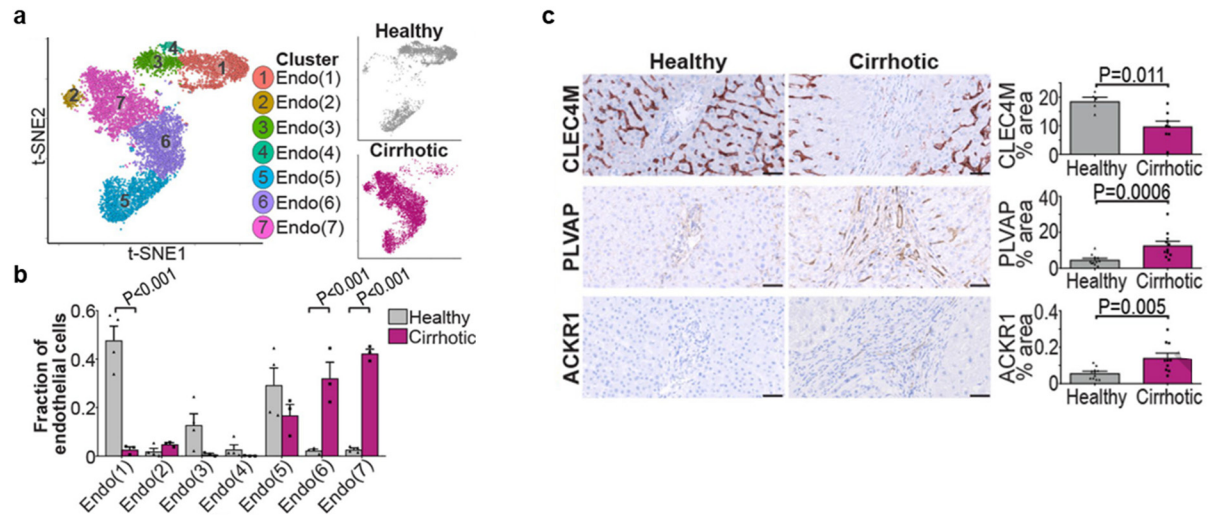
presentation, phagocytosis, chemokines, angiogenesis, production of extracellular matrix and wound healing were associated with a differentiation toward SAM $\Phi$  fate (**Fig. 23e**).



**Figure 23. Macrophage heterogeneity in liver cirrhosis.** *a*, Macrophage clusters and subset distribution in healthy and cirrhotic livers. *b*, Two subsets of macrophages were more prevalent in liver cirrhosis and were annotated as scar-associated macrophages (SAM $\Phi$ ). *c*, Gene expression profile of macrophages. *d*, Conditioned media from SAM $\Phi$  activate HSCs. *e*, Fate analysis of blood monocytes capable of differentiating into SAM $\Phi$  and conventional dendritic cells (cDCs). Adapted from Ramachandran et al.<sup>199</sup>.



Endothelial cell heterogeneity was also analyzed and 3 subsets were differently distributed in healthy and cirrhotic livers (**Fig. 24a, b**). Cirrhotic livers were characterized by a low prevalence of CLEC4M<sup>+</sup> endothelial cells and higher levels of PVALP<sup>+</sup> and ACKR1<sup>+</sup> cells (**Fig. 24c**).

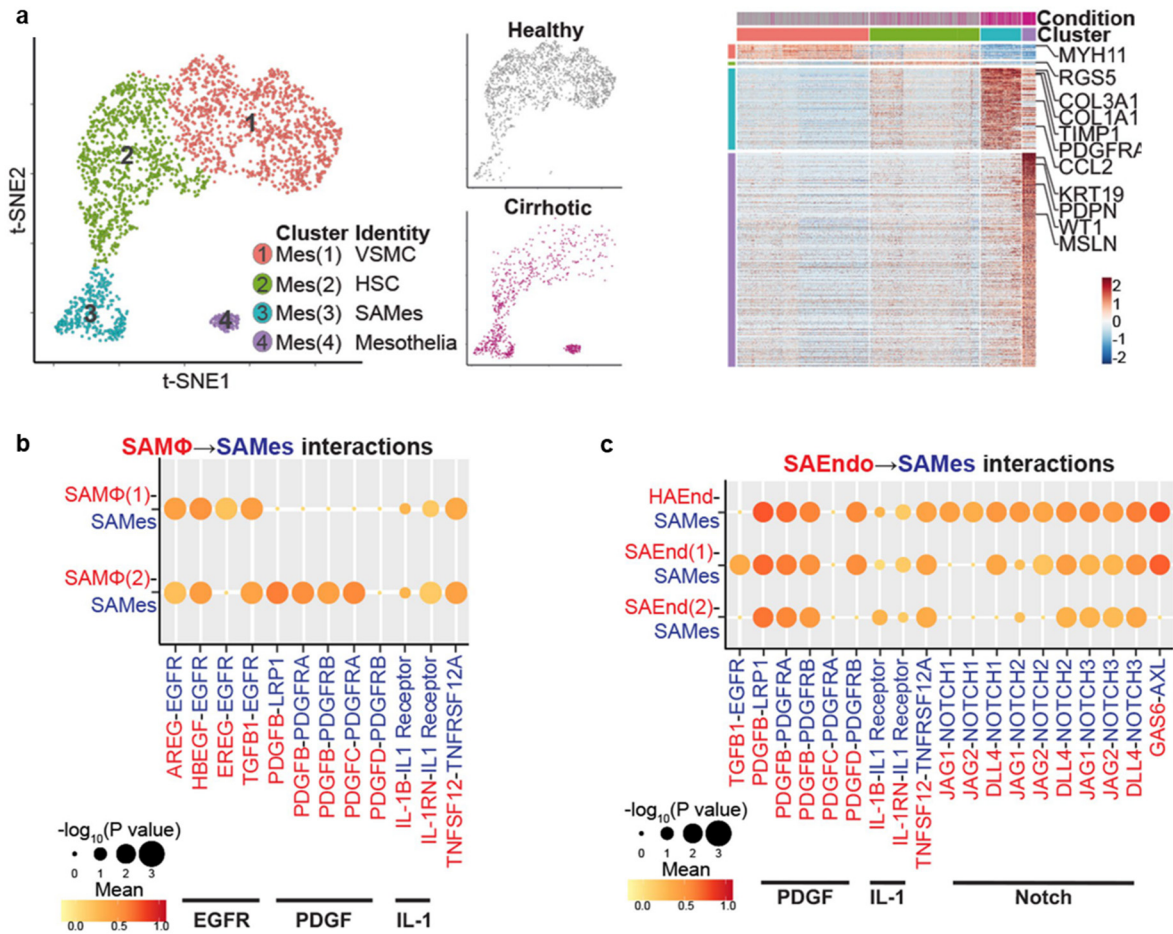


**Figure 24. Endothelial cell heterogeneity in healthy and cirrhotic livers.** *a, b*, Endothelial cells clusters and subset distribution in healthy and cirrhotic livers. *c*, Immunohistochemistry validation of the CLEC4M<sup>+</sup>, PLVAP<sup>+</sup> and ACKR1<sup>+</sup> subpopulation in healthy and cirrhotic livers. Adapted from Ramachandran et al.<sup>199</sup>.

The same type of analysis was carried on mesenchymal cells and 4 subpopulations were identified. The one enriched in cirrhotic livers and characterized by the expression of PDGFR $\alpha$  was annotated as scar-associated mesenchymal cells (SAMes, **Fig. 25a**). Using a database of receptor-ligand interactions (CellPhoneDB) and multiplex immunofluorescence, the two-way interactions between SAM $\Phi$ , scar-associated endothelial cells (SAEndo), arterial endothelial cells (HAEnd) and SAMes were described (**Fig. 25b, c**). SAM $\Phi$  express ligands for PDGFR, IL-1R, EGFR and TNFRSF12A expressed on SAMes and via these interactions they regulate SAMes

activation, proliferation and survival (**Fig. 25b**). SAMes and SAEndo interact via non-canonical Notch ligands *DLL4*, *JAG1*, *JAG2* with the receptor *NOTCH3* expressed on SAMes (**Fig. 25c**). These interactions were validated *ex vivo*. Conditioned media recovered from culturing primary SAM $\Phi$  activated human HSC. Primary endothelial cells from cirrhotic liver cultured together with HSC promoted collagen production, which was decreased upon treatment with the Notch-signaling inhibitor Dibenazepine.

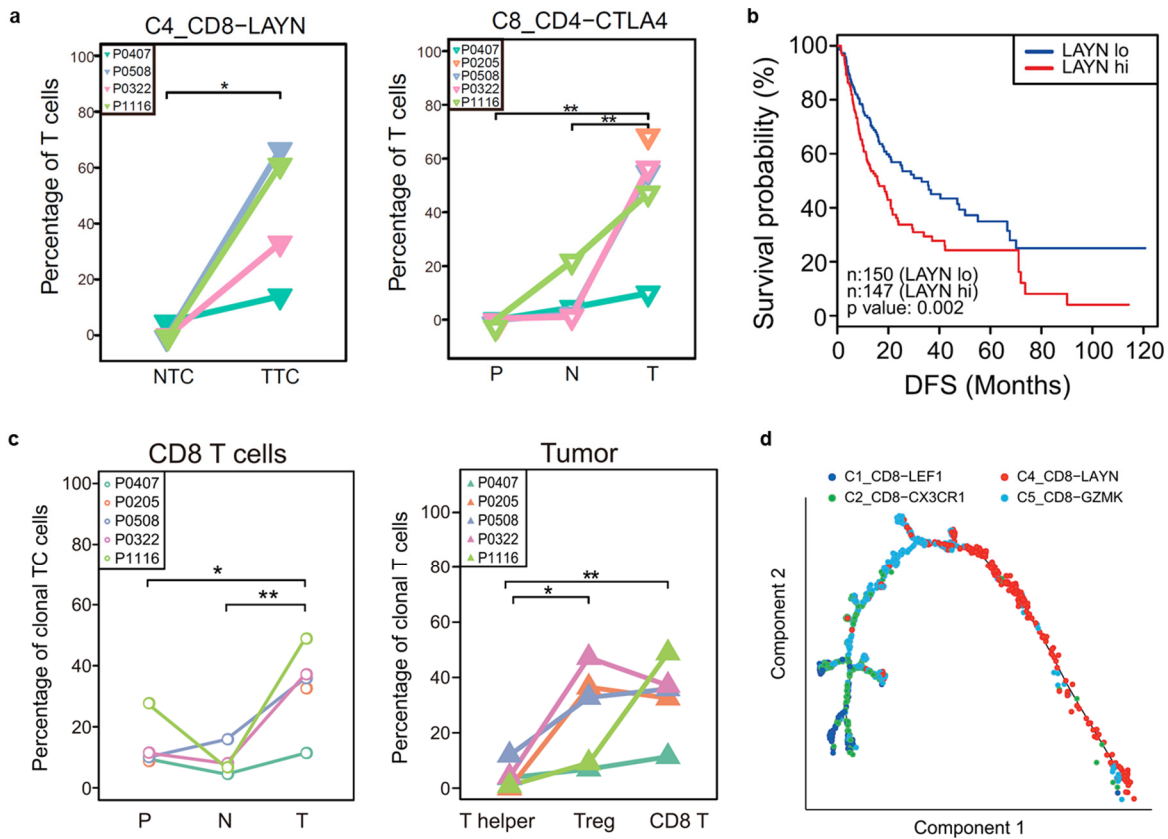
ScRNA-seq confirmed that fibrogenesis in liver cirrhosis is a process characterized by multiple cell interactions and cell state heterogeneity. Anti-fibrotic therapies need to take into account such complexity of the liver fibrotic niche.



**Figure 25. Mesenchymal cell heterogeneity in healthy and cirrhotic livers.** *a*, Mesenchymal cells clusters from healthy and cirrhotic livers and marker genes characterizing the different subsets. *b*, Dotplot of selected ligand-receptor interactions between SAM $\Phi$  and SAMes. *d*, Dotplot of selected ligand-receptor interactions between SAEndo and SAMes. Adapted from Ramachandran et al.<sup>199</sup>.

### 1.11.4 PRIMARY LIVER CANCER

In the era of immunotherapy, scRNA-seq has been used to study primarily immune cells composing HCC microenvironment. Zheng et al. investigated at single-cell level T cell composition in blood, non-tumor liver and tumor tissues from HCC patients and found that Tregs and exhausted CD8+ cells were clonally enriched in HCC (**Fig. 26**)<sup>200</sup>.



**Figure 26. T cell composing HCC microenvironment.** a, Exhausted CD8+ (CD8+LAYN+) and Treg cells (CD4+CTLA4+) are enriched in HCC (T, TTC) compared to plasma (P) and non-tumor tissue (N, NTC). b, LAYN expression is associated with disease free survival in TGCA database. c, CD8 T cells and Treg cells are clonally expanded in HCC. d, Pseudotime analysis of CD8 T cells TCR clonality showing that exhausted CD8 T cells derived from CD8+GZMK+ cells. Adapted from Zhen et al.<sup>200</sup>.

The authors identified 11 subsets of T cells in the blood and liver tissues (5 CD8+, 6 CD4+ and a unique subpopulation of regulatory-like CD8+FOXP3+ cells) and LAYN as a novel marker of T cell exhaustion associated with higher HCC recurrence<sup>200</sup> (**Fig. 26a, b**). Non-tumor tissues had high level of MAIT CD8 cells (CD8+ SCL4A10+) while HCCs had higher prevalence of exhausted CD8 cells (CD8+ LAYN+) and CD4 Treg cells (CD4+ CTLA4+), suggesting that HCC is characterized by an immunosuppressive microenvironment (**Fig. 26a**). The analysis at the single cell level of TCR clonality showed that in HCC the enrichment of CD8 exhausted and Treg cells is clonal and that CD8-LAYN+ cells probably derive from CD8-GZMK+ cells (**Fig. 26c, d**). This CD8+GZMK+ population could be a potential target for immunotherapy strategies aiming at preventing T cell exhaustion and the transition to CD8-LAYN+ cells<sup>200</sup>.

An integrated multiomic analysis of HCC using whole-exome sequencing, RNA-seq including scRNA-seq, mass spectrometry-based proteomics and metabolomics, and single cell mass cytometry techniques was also performed<sup>201</sup>. This analysis demonstrated the high heterogeneity of HCC cells and allowed a classification of HCC in three subtypes according to the immune state at the scRNA-seq level. HCC subtype 2 is characterized by reduced lymphocyte infiltration and higher prevalence of DCs and NK cells. HCC subtype 3 had higher frequency of immunosuppressive Treg cells, Breg cells and M2 macrophages while HCC subtype 1 had normal T cells infiltration. The three subtypes have also a different metabolomic profile and long-term patients' prognosis with subtype 1 patients showing the better survival<sup>201</sup>.

Ho et al. performed scRNA-seq in HCCs engrafted in mouse with a special focus on the EPCAM+ subpopulation to explore potential cancer stem cell and cancer stem cells markers; in doing so they were able to identify a rare CD24+/CD44+ subpopulation that can have a role in HCC proliferation<sup>202</sup>.

To study clonal evolution in HBV-related HCC, Duan et al. performed single-cell whole-genome sequencing in 96 tumor cells from 3 HCCs and showed that copy number variations occur early in HCC development and that by tracking HBV integration patterns it is possible to define the monoclonal or the polyclonal origin of a tumor<sup>203</sup>.

In a proof-of-concept study, D'Avola et al. combined imaging flow cytometry with scRNA-seq to analyze HCC-associated circulating tumor cells (CTCs) and found that scRNA-seq of CTCs can identify HCC driver genes and molecular HCC heterogeneity in tumors with vascular invasion<sup>204</sup>.

Multomics approaches to study DNA mutations, epigenetic changes and transcriptomic on the same cell are also under development. The scTRIO-seq technique allows the simultaneous recovery and analysis of single cell mRNA and DNA and it has been tested on few HCC cells showing the ability to catch cell heterogeneity on a multiomic level<sup>205</sup>.

While data at the single-cell level on cholangiocarcinoma are lacking, one study explored combined hepatocellular and intrahepatic cholangiocarcinoma (cHCC-ICC) using laser dissection and single nucleus sequencing. The authors found that combined and mixed type cHCC-ICCs are distinct subtypes, with the former showing strong ICC features and the latter Hoshida S2-like HCC characteristics and both having stem-like features and poor prognosis<sup>206</sup>. The clonality

tracing revealed that mixed cHCC-ICC can have both monoclonal and multiclonal origin while combined cHCC-ICC displayed always a monoclonal origin<sup>206</sup>.

ScRNA-seq is entering the cancer field with promising perspective in deepening our knowledge of cancer biology and hopefully in identifying new therapeutic targets for primary cancer treatment and prevention.

## 2. RESEARCH OBJECTIVES

Transcriptomic technologies have shown great potential for establishing tumor diagnosis, prognosis, and response to therapy in patients with advanced liver disease. However, cell-averaging (bulk) transcriptomic provides mix signals from the cell forming the tissue and cannot dissect cell heterogeneity, identify specific cell types contributing to disease or rare subpopulation hidden in the bulk signal noise. ScRNA-seq is a high-resolution technique that allows transcriptome-wide analyses of individual cells and represents the most advanced tool to study liver physiopathology, hepatocarcinogenesis and tumor microenvironment and track cell composition changes during liver disease.

My research focused on study liver physiopathology at single cell level and had two main objectives. The first aim was to characterize cell heterogeneity, zonation and progenitors in the normal human liver using mCEL-Seq2 – a high-resolution scRNA-seq technology - and build a human liver cell atlas to pave the way to characterize chronic liver diseases and cancer physiopathology at single cell-level and finally identify new therapeutic targets. In my second aim, scRNA-seq was used to study intratumor heterogeneity and virus host-interactions in HBV-related HCC to unravel HBV impact on cancer physiopathology and development.



## 3. RESULTS

### 3.1 A HUMAN LIVER CELL ATLAS REVEALS HETEROGENEITY AND EPITHELIAL PROGENITORS

The cellular composition of the liver remains poorly understood. The aim of the study was to performed scRNA-seq of normal human liver tissue to construct a human liver cell atlas. In collaboration with the Max-Planck Institute of Epigenetic and Immunology (Freiburg, Germany) we sequenced more than 10,000 cells by mCEL-Seq2 from nine human donors. Our analysis identified previously unknown subtypes of endothelial cells, KCs and hepatocytes and reconstructed the transcriptome-wide zonation of some of these populations. We found that around 41% of hepatocyte genes are significantly zoned, periportal hepatocytes are enriched in genes involved in biological oxidation as well as glycogen synthesis and that the human zonation is not monotonic with some pathways and genes highly expressed in the mid lobular zone. We discovered that the EPCAM<sup>+</sup> population is heterogeneous and includes hepatocyte-biased and cholangiocyte-biased populations as well as a TROP2<sup>int</sup> progenitor population with strong potential to form liver organoids. As a proof-of-principle, we used our atlas to unravel the transcriptomic changes that occur in HCC cells and in human hepatocytes and liver endothelial cells engrafted into a mouse liver. Our human liver cell atlas provides a powerful resource to enable the discovery of previously unknown cell types in normal and diseased livers and identify new therapeutic targets for chronic liver diseases and cancer.

# A human liver cell atlas reveals heterogeneity and epithelial progenitors

Nadim Aizarani<sup>1,2,3</sup>, Antonio Saviano<sup>4,5,6,8</sup>, Sagar<sup>1,8</sup>, Laurent Mailly<sup>4,5</sup>, Sarah Durand<sup>4,5</sup>, Josip S. Herman<sup>1,2,3</sup>, Patrick Pessaux<sup>4,5,6</sup>, Thomas F. Baumert<sup>4,5,6\*</sup> & Dominic Grün<sup>1,7\*</sup>

**The human liver is an essential multifunctional organ. The incidence of liver diseases is rising and there are limited treatment options. However, the cellular composition of the liver remains poorly understood. Here we performed single-cell RNA sequencing of about 10,000 cells from normal liver tissue from nine human donors to construct a human liver cell atlas. Our analysis identified previously unknown subtypes of endothelial cells, Kupffer cells, and hepatocytes, with transcriptome-wide zonation of some of these populations. We show that the EPCAM<sup>+</sup> population is heterogeneous, comprising hepatocyte-biased and cholangiocyte populations as well as a TROP2<sup>int</sup> progenitor population with strong potential to form bipotent liver organoids. As a proof-of-principle, we used our atlas to unravel the phenotypic changes that occur in hepatocellular carcinoma cells and in human hepatocytes and liver endothelial cells engrafted into a mouse liver. Our human liver cell atlas provides a powerful resource to enable the discovery of previously unknown cell types in normal and diseased livers.**

The liver serves as a central metabolic coordinator with a wide array of essential functions, including the regulation of glucose and lipid metabolism, protein synthesis, and bile synthesis. Furthermore, the liver is a visceral organ that is capable of remarkable natural regeneration after tissue loss<sup>1</sup>. However, the prevalence of liver diseases and mortality associated with them have risen markedly within recent decades<sup>2</sup>. The liver cellular landscape has barely been explored at single-cell resolution, which limits our molecular understanding of liver function and disease biology. The recent emergence of sensitive single-cell RNA sequencing (scRNA-seq) methods<sup>3</sup> allows us to investigate cell types in healthy and diseased tissue.

To characterize the human liver at single-cell resolution, we developed a robust pipeline for scRNA-seq of cryopreserved and freshly isolated samples of patient-derived human liver and assembled an atlas consisting of 10,372 cells from nine donors. We performed in-depth analysis of all liver cell types with a focus on epithelial liver cell progenitors.

## scRNA-seq of the human liver

We used mCEL-Seq<sup>24</sup> for scRNA-seq of non-diseased liver tissue from nine patients who underwent liver resections for colorectal cancer metastasis or cholangiocarcinoma without history of chronic liver disease (Fig. 1a, see Methods). We sorted and sequenced viable cells both in an unbiased fashion and by enriching specific cell populations on the basis of cell surface marker expression (Extended Data Fig. 1, see Methods). Because fresh liver tissue material is scarce and difficult to preserve, and biobanks with cryopreserved liver samples represent rich resources, we generated scRNA-seq data from cryopreserved cells as well as single-cell suspensions from freshly prepared liver samples (see Methods). We then used RaceID3 for the identification of cell types<sup>4,5</sup> (see Methods).

Cells from different patients, isolated from freshly prepared or cryopreserved single-cell suspensions, co-clustered (Extended Data Fig. 1).

Furthermore, fresh and cryopreserved cells from the same patient did not have markedly different gene signatures (Extended Data Fig. 1e–h). However, there were compositional differences both between fresh and cryopreserved samples derived from the same patient and among different fresh (or cryopreserved) samples. We attribute these differences to variability in cell viability and cell type composition across samples.

As scRNA-seq of randomly sampled populations yielded almost exclusively hepatocytes and immune cells (Extended Data Fig. 1i), we applied additional sorting strategies to enrich for endothelial cells (Extended Data Fig. 1a–c) and EPCAM<sup>+</sup> cells (see below).

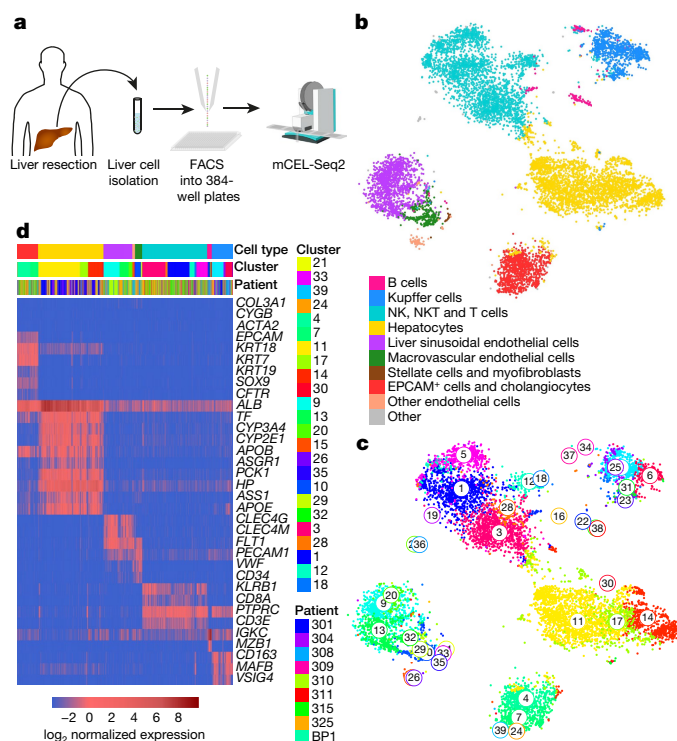
Our atlas comprises all the main liver cell types defined by the expression of marker genes, including hepatocytes, EPCAM<sup>+</sup> bile duct cells (cholangiocytes), CLEC4G<sup>+</sup> liver sinusoidal endothelial cells (LSECs), CD34<sup>+</sup>PECAM<sup>high</sup> macrovascular endothelial cells (MaVECs), hepatic stellate cells and myofibroblasts, Kupffer cells, and immune cells (Fig. 1b–d, Supplementary Table 1). To facilitate interactive exploration of our human liver cell atlas, we created a web interface: <http://human-liver-cell-atlas.ie-freiburg.mpg.de/>.

## Zonation of human liver cell types

Hepatocytes are spatially heterogeneous and zoned along the portal–central axis of the liver lobule<sup>6–8</sup>. According to metabolic sub-specialization, the liver lobule has been divided into the periportal zone surrounding the portal triad (portal vein, hepatic artery and bile duct), the central zone nearest to the central vein, and the remaining mid zone<sup>6–8</sup>. Whereas previous observations have suggested that non-parenchymal cells such as LSECs and Kupffer cells have specialized subtypes<sup>6</sup>, it has been hard to demonstrate heterogeneity of these cell types, and most studies have been carried out in rodents.

We were able to directly compare the signatures of MaVECs and LSECs, and identified several previously unknown subpopulations (Extended Data Fig. 2, Supplementary Note 1).

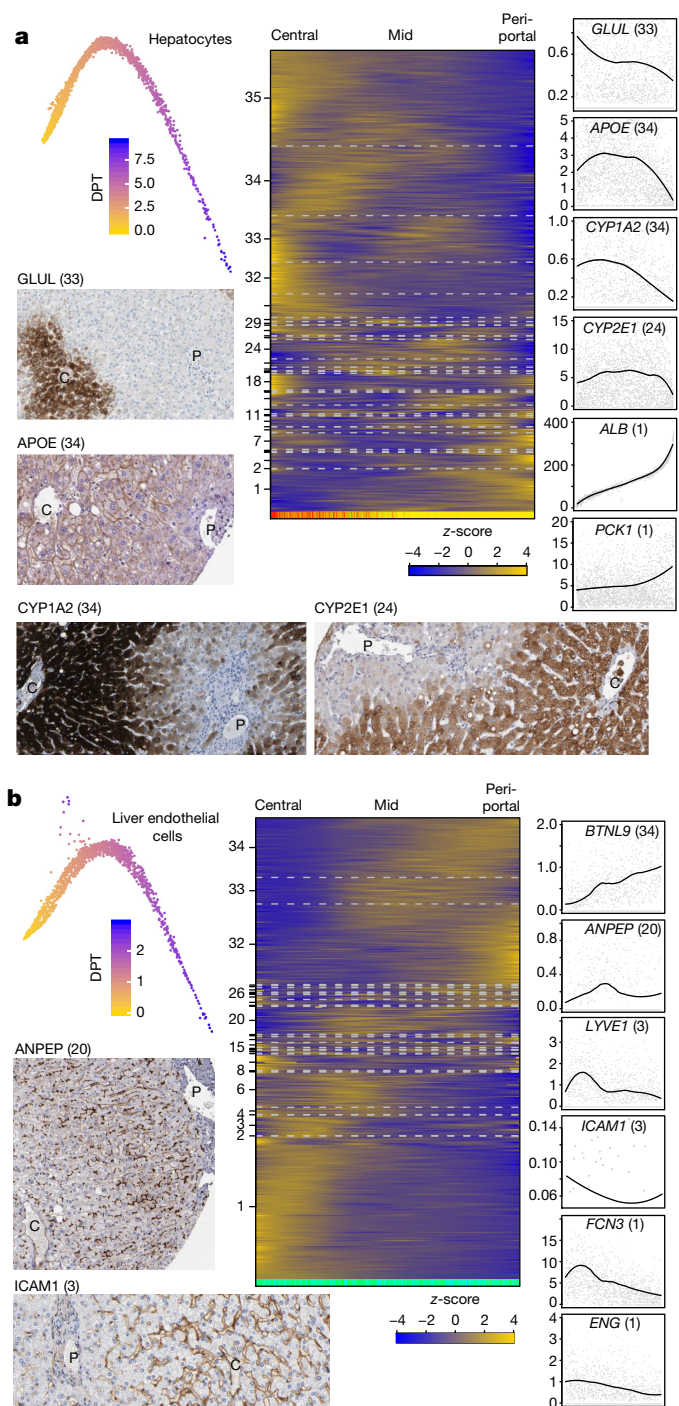
<sup>1</sup>Max-Planck-Institute of Immunobiology and Epigenetics, Freiburg, Germany. <sup>2</sup>Faculty of Biology, University of Freiburg, Freiburg, Germany. <sup>3</sup>International Max Planck Research School for Molecular and Cellular Biology (IMPRS-MCB), Freiburg, Germany. <sup>4</sup>Institut National de la Santé et de la Recherche Médicale, Unité 1110, Institut de Recherche sur les Maladies Virales et Hépatiques, Strasbourg, France. <sup>5</sup>Université de Strasbourg, Strasbourg, France. <sup>6</sup>Pôle Hepato-digestif, Institut Hospitalo-universitaire, Hôpitaux Universitaires, Strasbourg, France. <sup>7</sup>CIBSS—Centre for Integrative Biological Signaling Studies, University of Freiburg, Freiburg, Germany. <sup>8</sup>These authors contributed equally: Antonio Saviano, Sagar. \*e-mail: [thomas.baumert@unistra.fr](mailto:thomas.baumert@unistra.fr); [gruen@ie-freiburg.mpg.de](mailto:gruen@ie-freiburg.mpg.de)



**Fig. 1 | scRNA-seq reveals cell types in the adult human liver.** **a**, Outline of the protocol used for scRNA-seq of human liver cells. Samples from liver resections were digested to prepare single-cell suspensions. Cells were sorted into 384-well plates and processed according to the mCEL-Seq2 protocol. **b**, *t*-SNE map of single-cell transcriptomes from normal liver tissue from nine donors highlighting the main liver cell compartments. 'Other' denotes various small populations comprising 22 red blood cells and 46 cells that cannot be unambiguously annotated. 'Other endothelial cells' cannot be unambiguously classified as LSECs or MaVECs. **c**, *t*-SNE map of single-cell transcriptomes highlighting RaceID3 clusters, which reveals subtype heterogeneity in all major cell populations of the human liver. Numbers denote clusters. **d**, Heat map showing the expression of established marker genes for each cell compartment. Colour bars indicate patient, major cell type, and RaceID3 cluster. Scale bar,  $\log_2$ -transformed normalized expression. **b**, **c**,  $n = 10,372$  cells.

scRNA-seq has been highly informative on hepatocyte zonation in mouse<sup>9</sup>, and the first single-cell analysis of human hepatocyte and endothelial cell zonation at limited resolution was done recently<sup>10</sup>. To infer continuous transcriptome-wide zonation, we reasoned that the major axis of variability for a cell type could reflect gene expression changes associated with zonation. Hence, we ordered LSECs and hepatocytes by diffusion pseudo-time (dpt)<sup>11</sup>, here interpreted as pseudo-space, along this axis and applied self-organizing maps (SOMs) to infer co-expression modules (Fig. 2, see Methods).

We first validated our strategy by recovering the previously characterized zonation of mouse hepatocytes<sup>9</sup> (Extended Data Fig. 3a–d). For our human hepatocytes, this approach recovered zoned expression patterns of landmark genes: for example, *ALB* and *PCK1* (periportal module 1), *CYP1A2* and *CYP2E1* (central/midzonal modules 34 and 24, respectively), and *GLUL* (central module 33)<sup>7,9</sup> (Fig. 2a, Extended Data Fig. 3e–g, Supplementary Tables 2, 3). In total, 1,384 out of 3,395 expressed genes (41%) included in the hepatocyte analysis exhibited significant zonation (Benjamini–Hochberg corrected ANOVA,  $P < 0.01$ ). Pathway enrichment analysis revealed that periportal hepatocyte modules are enriched in genes involved in biological oxidation, consistent with an oxygen gradient that peaks in the periportal zone<sup>6–8</sup>, and in the glycogen synthesis pathway (Extended Data Fig. 3h). In accordance with its zonation in mouse hepatocytes, the urea cycle enzyme *CPS1* peaks in periportal hepatocytes (Extended Data Fig. 3g). Midzonal hepatocyte modules are enriched in, for example, metabolism



**Fig. 2 | Heterogeneity and zonation of hepatocytes and endothelial cells.** **a**, Diffusion maps (left) and SOMs (middle) of single-cell transcriptome-derived zonation profiles for hepatocytes ( $n = 2,534$  cells). DPT indicates diffusion pseudo-time and is interpreted here as a spatial zonation coordinate. Right, zonation profiles of *GLUL* (central), *APOE* (midzonal), *CYP1A2* and *CYP2E1* (central/midzonal) and *ALB* and *PCK1* (periportal); bottom left, immunostaining for *GLUL*, *APOE*, *CYP1A2*, and *CYP2E1* from the Human Protein Atlas<sup>31</sup>. See Extended Data Fig. 3g for additional images. **b**, Diffusion maps (left) and SOMs (middle) of single-cell transcriptome-derived zonation profiles for endothelial cells ( $n = 1,361$  cells). Right, zonation profiles of *BTNL9* and *ANPEP* (periportal), *LYVE1* and *FCN3* (midzonal), and *ICAM1*, *FCN3* and *ENG* (central); bottom left, immunostaining for *ICAM1* and *ANPEP* from the Human Protein Atlas. P, portal tracts; C, central. Colour bars at the bottom of the SOMs show RaceID3 cluster as in Fig. 1. The y axis of the zonation profiles indicates normalized expression.



of xenobiotics by cytochrome P450. Immunostainings for selected genes validate the predicted zonation at the protein level (Fig. 2a).

LYVE1 and CD14 have been identified as markers that distinguish midzonal and central LSECs from periportal LSECs<sup>12</sup>. Analysis of LSEC zonation showed that 806 out of 1,198 expressed genes (67%) exhibited significant zonation (Benjamini–Hochberg corrected ANOVA,  $P < 0.01$ ) (Fig. 2b, Extended Data Fig. 3i, Supplementary Tables 4, 5). Central and midzonal endothelial cells (modules 1 and 3) exhibited peaked expression of *LYVE1* and *FCN3*, which encodes a ficolin protein that can switch on the lectin pathway of complement activation. Notably, pathway enrichment analysis of the central and midzonal endothelial modules recovered pathways, such as binding and uptake of ligands by scavenger receptors, that are shared with midzonal hepatocytes (Extended Data Fig. 3j). Together with a more detailed gene expression analysis (Supplementary Note 2) this observation suggests that genes and functions are co-zonated across hepatocytes and endothelial cells.

Finally, a comparison between mouse<sup>9,13</sup> and human cells revealed only limited evolutionary conservation of gene expression zonation (Supplementary Note 3, Extended Data Fig. 3k, l, Supplementary Tables 6, 7), reflecting widespread evolutionary changes.

### Human liver immune cell populations

A detailed analysis of the *CD163*<sup>+</sup>*VSIG4*<sup>+</sup> Kupffer cell compartment revealed subpopulations with distinct gene expression signatures (Supplementary Note 4, Extended Data Fig. 4), in agreement with a recent study<sup>10</sup>. Moreover, we detected shared gene expression and pathways between Kupffer cell subsets and endothelial cells (Supplementary Note 4, Extended Data Fig. 4), providing further evidence that different cell types show functional co-operation.

We identified an *MS4A1*<sup>+</sup>*CD37*<sup>+</sup> subset of B cells, which corresponds to circulating B cells with upregulated MHC class II components, and a liver-resident *MZB1*<sup>+</sup> subset of B cells that expresses *DERL3*, *SSR4* and *IGHG4* (Extended Data Fig. 5).

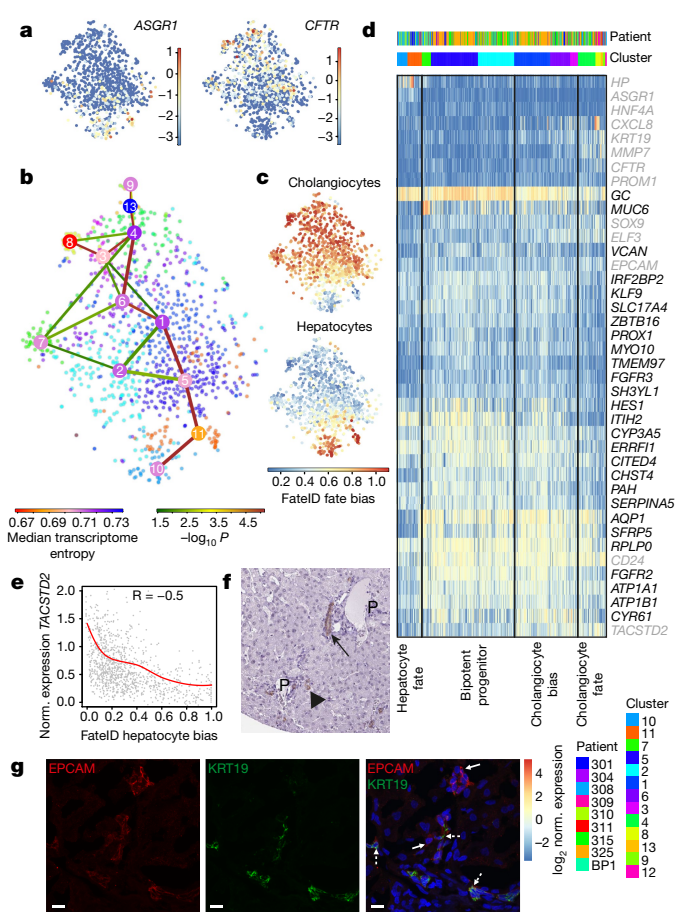
Finally, we recovered a population of *CD56*<sup>+</sup> (also known as *NCAM1*<sup>+</sup>) natural killer (NK) cells (cluster 5), as well as *CD56*<sup>−</sup> (cluster 3) and *CD56*<sup>+</sup> (cluster 1) *CD8A*<sup>+</sup> NKT cells, which expressed different combinations of chemokine ligands, granzymes, and killer cell lectin-like receptor genes (Extended Data Fig. 6). In clusters 12 and 18, a number of heat-shock genes are upregulated. These observations demonstrate an unexpected variety of immune cell subtypes in the human liver.

### Putative bipotent epithelial progenitors

Liver regeneration after tissue damage involves the replication of several types of liver cells, including hepatocytes and cholangiocytes. Furthermore, different types of liver damage lead to specific mechanisms of liver regeneration<sup>14,15</sup>. However, the existence of a population of naive adult stem cells in the human liver and its contribution to turnover and regeneration remains controversial. Rare *EPCAM*<sup>+</sup> cells have been termed hepatic stem cells<sup>16</sup>; these can form dense round colonies when cultured and are bipotent progenitors of hepatoblasts, which differentiate into cholangiocytes or hepatocytes both in vitro and in vivo<sup>16,17</sup>.

To search for genuine liver progenitor cells, we sorted and sequenced single *EPCAM*<sup>+</sup> cells from adult human livers. We identified biliary and potential liver progenitor cell surface marker genes that correlated with *EPCAM* or *TROP1* expression; these included *TACSTD2* (also known as *TROP2*), *FGFR2*, *TM4SF4* and *CLDN1*. Immunohistochemistry confirmed the expression of predicted markers such as ANXA4 and the transcriptional co-activator *WWTR1* (Extended Data Fig. 7a).

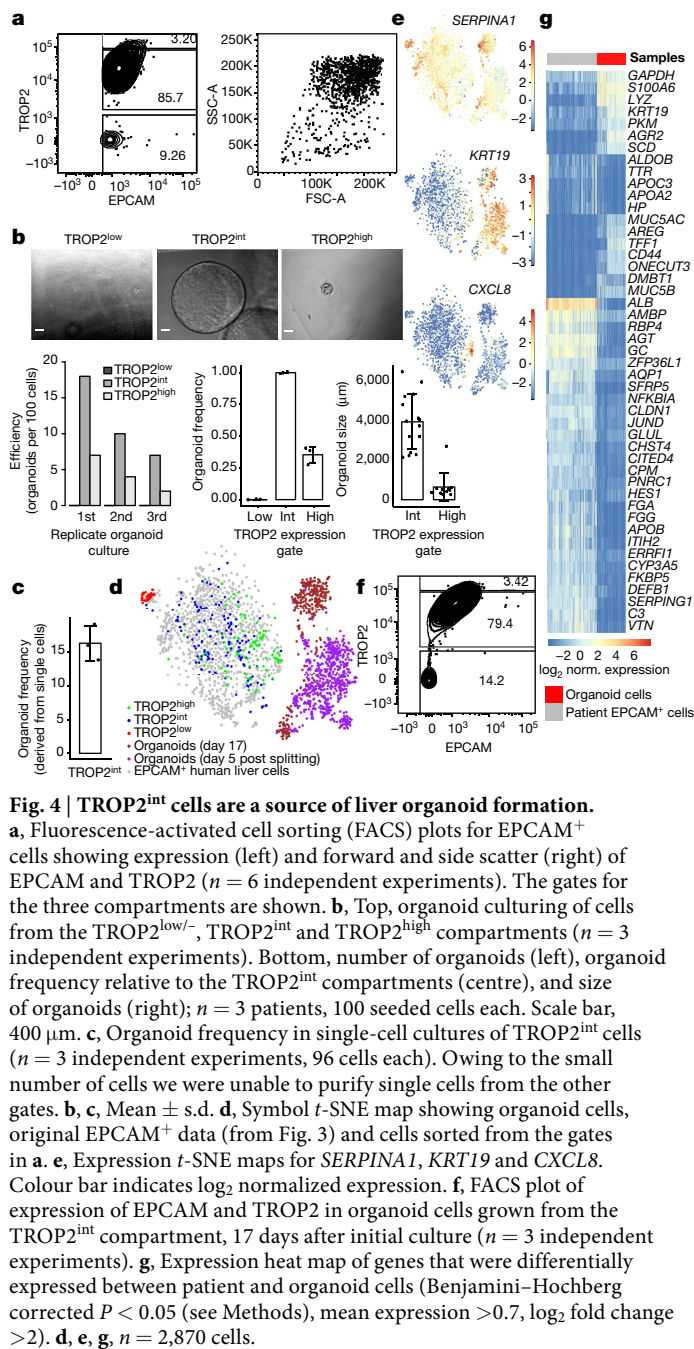
A focused analysis revealed that the *EPCAM*<sup>+</sup> compartment is transcriptionally heterogeneous and consists of an *ASGR1*<sup>+</sup> hepatocyte-biased population, *KRT19*<sup>high</sup>*CFTR*<sup>high</sup>*ALB*<sup>low</sup> cholangiocyte populations, and a remaining population of putative naive progenitor cells (Fig. 3a, Extended Data Fig. 7b, c). The *EPCAM*<sup>+</sup> population exhibits only stochastic expression of the proliferation markers *MKI67* and



**Fig. 3 | Identification of a putative progenitor population in the adult human liver.** **a**, Expression *t*-SNE maps of *ASGR1* and *CFTR* for the *EPCAM*<sup>+</sup> compartment only. The colour bar indicates log<sub>2</sub> normalized expression. **b**, StemID2<sup>18</sup> analysis of the *EPCAM*<sup>+</sup> compartment. Shown are links with StemID2  $P < 0.05$ . Node colour denotes transcriptome entropy. **c**, FateID analysis of the *EPCAM*<sup>+</sup> compartment highlights populations that are preferentially biased towards hepatocyte progenitors and cholangiocytes, respectively, and reveals similar bias towards both lineages in the central population (clusters 1, 2, 5, 6 and 7). Colour bar indicates lineage probability. **d**, Expression heatmap of selected hepatocyte marker genes (*HP*, *ASGR1*), mature cholangiocyte genes (*KRT19*, *CFTR*, *CXCL8*, *MMP7*), additional progenitor markers (grey), and all genes upregulated in the central population (clusters 1, 2, 5, 6 and 7) within the *EPCAM*<sup>+</sup> compartment (Benjamini–Hochberg corrected  $P < 0.01$ ; fold change  $> 1.33$ ; see Methods). Four compartments are indicated, resolving the predicted fate bias (Extended Data Fig. 8). **e**, Correlation of nearest-neighbour-imputed ( $k = 5$ ) expression (using RaceID3) of *TACSTD2* and hepatocyte bias predicted by FateID. Red line, loess regression. **R**, Spearman's rank correlation. **a–e**,  $n = 1,087$  cells. **f**, Immunostaining for *TROP2* from the Human Protein Atlas ( $n = 3$  biologically independent samples). Arrow, bile duct; arrowhead, bile ductule. **g**, Immunofluorescence labelling of *EPCAM* and *KRT19*. *EPCAM*<sup>+</sup>*KRT19*<sup>low/−</sup> (solid arrow) and *EPCAM*<sup>+</sup>*KRT19*<sup>+</sup> (broken arrow) cells are indicated. Nuclei are stained with DAPI. Images are maximum *z*-stack projections of 6  $\mu\text{m}$ . Scale bar, 10  $\mu\text{m}$ . ( $n = 3$  independent experiments).

*PCNA* and is negative for the hepatoblast marker *AFP* (Extended Data Fig. 7d). Hence, the transcriptional heterogeneity of this population is unlikely to arise as a result of proliferation, and the observed subtypes reside in the normal human liver.

To explore the relatedness of these subpopulations, we reanalysed the *EPCAM*<sup>+</sup> population with RaceID3 and used StemID2 for lineage reconstruction<sup>4,18</sup> (Fig. 3b, see Methods). This analysis showed that the population in the centre of the *t*-distributed stochastic neighbour embedding (*t*-SNE) map (clusters 1, 2, 5, 6, 7) bifurcates into hepatocyte



progenitors and cholangiocytes. To provide further evidence for continuous differentiation trajectories connecting naive EPCAM<sup>+</sup> progenitors to cholangiocytes and mature hepatocytes, we performed StemID2 and diffusion map analyses on the combined population of mature hepatocytes and EPCAM<sup>+</sup> cells (Extended Data Fig. 8a–c).

To better understand the emergence of fate bias towards the two lineages, we used FateID to infer lineage probabilities in each cell<sup>4</sup>. Consistently, FateID inferred similar probabilities that the central population would differentiate towards hepatocytes and cholangiocytes (Fig. 3c). The fate bias predictions are supported by a differential gene expression analysis revealing upregulation of common genes that encode several signalling pathway components (*HES1*, *SFRP5*, *FGFR2*, *FGFR3*) in the central population (Fig. 3d), and gradual upregulation of distinct gene sets towards the hepatocyte-biased and cholangiocyte populations (Extended Data Fig. 8e). The expression of *TROP2* was negatively correlated with hepatocyte fate bias, exhibiting a gradient that ranged from high expression in mature cholangiocytes to very

low expression in the hepatocyte-biased population (Fig. 3e, Extended Data Fig. 7c). Immunostaining for TROP2 in normal human liver tissue showed specific expression in cells of the bile ducts and bile ductules (Fig. 3f). Notably, TROP2 expression has been found in amplifying oval cells in injured mouse livers<sup>19</sup>.

The central TROP2<sup>int</sup> population is in itself heterogeneous and contains a MUC6<sup>high</sup> population (cluster 7) (Extended Data Fig. 7c). MUC6 is highly expressed by pancreatic progenitors and multi-potent bile duct tree stem cells<sup>20</sup>, which have been proposed to be the origin of the EPCAM<sup>+</sup> hepatic stem cells. The TROP2<sup>high</sup> cholangiocyte clusters comprise a CXCL8<sup>+</sup> population (cluster 8) and an MMP7<sup>+</sup> population (clusters 4 and 13) (Extended Data Figs. 7c, 8e, f), whereas TROP2<sup>low</sup> clusters show upregulation of hepatocyte markers such as *ALB*, *HP*, *HNF4A* and *ASGR1* (Fig. 3d, Extended Data Figs. 7c, 8e, f).

The central TROP2<sup>int</sup> population that was stratified as bipotent on the basis of FateID-predicted bias expresses genes that encode early developmental transcription factors such as *HES1*, which is essential for tubular bile duct formation<sup>21</sup>, and *PROX1*, an early specification marker for the developing liver in the mammalian foregut endoderm that is required for hepatocyte proliferation and migration during development<sup>22</sup> (Fig. 3d). Furthermore, this population showed lower expression of hepatocyte genes such as *HNF4A*, *HP* and *ALB* and of cholangiocyte genes such as *KRT19* and *CFTR* compared to the hepatocyte-biased and mature cholangiocyte populations, respectively (Fig. 3d, Extended Data Figs. 7c, 8f). We speculate that we enriched for the TROP2<sup>int</sup> KRT19<sup>low/-</sup> immature population during cell isolation, as mature bile duct cells require a harsher digestion for their isolation, which can negatively affect other liver cell types. Thus, the actual fraction of KRT19<sup>high</sup> cells in the tissue is presumably higher. We validated the existence of EPCAM<sup>+</sup>KRT19<sup>low/-</sup> cells in addition to EPCAM<sup>+</sup>KRT19<sup>high/+</sup> cells in situ by immunofluorescence (Fig. 3g, Extended Data Fig. 7e).

Consistent with our scRNA-seq data, flow cytometry profiles of EPCAM and TROP2 displayed a gradient of TROP2 expression in EPCAM<sup>+</sup> cells, and EPCAM expression correlated with TROP2 expression (Fig. 4a). Moreover, forward and side-scatter profiles of EPCAM<sup>+</sup> cells indicated that the compartment is heterogeneous and consists of populations with different sizes and morphologies (Fig. 4a). On the basis of the distribution of TROP2 expression, we compartmentalized EPCAM<sup>+</sup> cells into three compartments: TROP2<sup>low/-</sup>, TROP2<sup>int</sup>, and TROP2<sup>high</sup> (Fig. 4a). To confirm that the TROP2<sup>int</sup> population harbours the progenitor population, we attempted to culture bipotent organoids<sup>23</sup> from each compartment. In agreement with our prediction, TROP2<sup>int</sup> cells exhibited the highest organoid-forming capacity, whereas TROP2<sup>low/-</sup> cells did not form organoids, and TROP2<sup>high</sup> cells gave rise to much smaller organoids at a strongly reduced frequency compared to TROP2<sup>int</sup> cells (Fig. 4b). Single-cell culture of TROP2<sup>int</sup> cells demonstrated the organoid-forming capacity of individual cells from this gate, providing evidence for bipotency at the clonal level (Fig. 4c). As expected, scRNA-seq of the input populations for organoid culture from each compartment showed a marked enrichment of the respective compartment in the original EPCAM<sup>+</sup> data (Fig. 4d, e, Extended Data Fig. 8g, h). Notably, flow cytometry profiles of EPCAM and TROP2 for organoid cells grown from the TROP2<sup>int</sup> compartment recovered TROP2<sup>low/-</sup>, TROP2<sup>int</sup> and TROP2<sup>high</sup> populations in the organoids (Fig. 4f).

To elucidate the cell type composition of the organoids in depth, we performed scRNA-seq. Co-analysis of organoid cells and EPCAM<sup>+</sup> cells sequenced directly from patients demonstrated marked transcriptome differences (Fig. 4e). Although EPCAM and CD24 were expressed in cells from both organoids and patients, organoid cells showed lower expression of various genes such as *AQP1* and the WNT signalling modulator *SFRP5*, and higher expression of others, such as the proliferation marker *MKI67*<sup>+</sup>, reflected by differential enrichment of the corresponding pathways (Fig. 4g, Extended Data Fig. 8i–k). We observed several subpopulations within the organoids, including a non-dividing hepatocyte-biased *SERPINA1*<sup>high</sup> population and a



non-dividing *KRT19*<sup>high</sup> cholangiocyte-biased population, consistent with the signature of the EPCAM<sup>+</sup> cells recovered from the patients (Fig. 4e). This further supports the claim that the TROP2<sup>int</sup> compartment harbours a bipotent progenitor population, which can give rise to hepatocyte and cholangiocyte populations.

In contrast to patient cells, organoid cells showed strong downregulation of *ALB* but expressed *AGR2* and other mucin family genes such as *MUC5AC* and *MUC5B*, which are normally expressed, for example, in intestinal cells and gastrointestinal cancers<sup>24,25</sup> (Fig. 4g, Extended Data Fig. 8j). These observations reflect that organoid cells express genes that are expressed in other systems, acquire a more proliferative state, and appear to upregulate stem cell-related pathways such as WNT signalling.

In light of these functional validation experiments, the observed gene signature of TROP2<sup>int</sup> cells, and the in situ location of these cells, our data strongly suggest that the putative liver progenitor population can be defined as a subpopulation of bile duct cells.

### Perturbed cell states in liver cancer

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer<sup>26</sup>. To demonstrate the value of our atlas as a reference for comparisons with diseased liver cells, we sequenced CD45<sup>+</sup> and CD45<sup>-</sup> cells from HCC tissue from three patients (Extended Data Fig. 9a, b, see Methods).

We recovered several cell types from the tumours, including cancer cells, endothelial cells, Kupffer cells, NKT cells and NK cells (Fig. 5a, Extended Data Fig. 9c) and compared them to the normal liver cell atlas. Differential gene expression analysis and immunohistochemistry revealed that cancer cells lose the expression of cytochrome P450 genes such as *CYP2E1* and *CYP2C8* and the periportal zoned gene *CPS1* (Fig. 5b, Extended Data Fig. 9d, e) as well as the metabolic signature of normal hepatocytes (Fig. 5c). They show increased expression of *AKR1B10*, a known biomarker of HCC with potential involvement in hepatocellular carcinogenesis<sup>27</sup> (Extended Data Fig. 9d). Moreover, immunohistochemistry confirmed that IL32, a pro-inflammatory TNF $\alpha$ -inducing cytokine, is highly upregulated in cancer cells (Fig. 5b). Overall, cancer cells show upregulation of WNT and Hedgehog signalling pathways, highlighting similarities between EPCAM<sup>+</sup> normal liver progenitors and the observed cancer cell population (Fig. 5c).

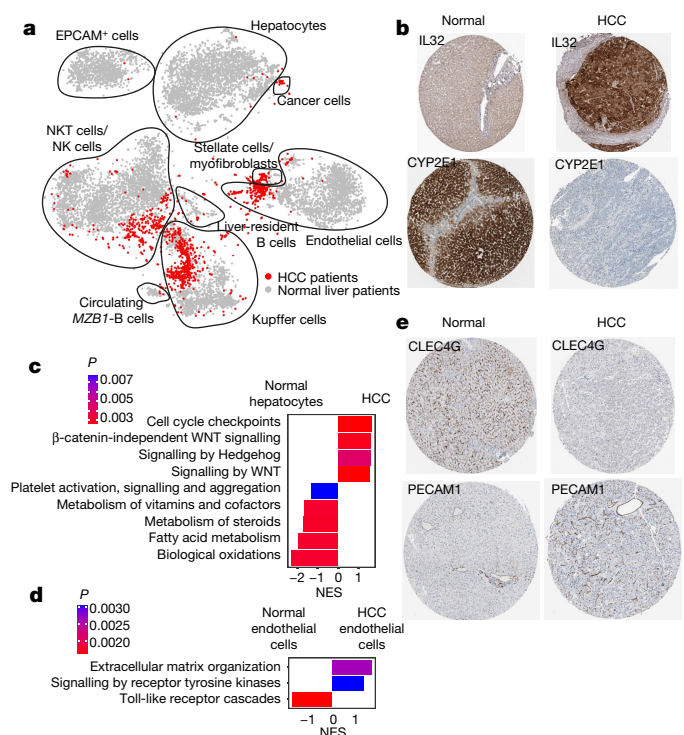
Endothelial cells from tumours show upregulation of, for example, extracellular matrix organization genes such as *COL4A2* and *SPARC* (Fig. 5d, Extended Data Fig. 9f). Strikingly, they do not express LSEC marker genes such as *CLEC4G* but do express MaVEC marker genes such as *PECAM1*, *AQP1* and *CD34* (Fig. 5e, Extended Data Fig. 9f, g). Moreover, HCC LSECs show increased expression of PLVAP, which makes them less permeable and could potentially restrict the access of lymphocytes and soluble antigens<sup>28</sup> to the tumour (Supplementary Note 5, Extended Data Fig. 9f, g).

We conclude that the comparison of scRNA-seq data between the cell populations of HCC and the liver cell atlas allows the inference of perturbed gene expression signatures, biomarkers and modulated functions across cell types.

### A human liver chimaeric mouse model

Mice harbouring patient-derived xenografted liver cells are a powerful tool for studying human liver cells and diseases in vivo<sup>29</sup>. To correctly interpret such experiments, it is crucial to understand the differences between cells taken directly from the human liver and human cells that have been transplanted into the mouse liver.

To address this issue, we transplanted human liver cells from patient-derived hepatocyte and non-parenchymal cell fractions into FRG-NOD (*Fah*<sup>-/-</sup> *Rag2*<sup>-/-</sup> *Il2rg*<sup>-/-</sup> non-obese diabetic) mice<sup>30</sup> (HMouse); after engraftment, we sorted single human cells in an unbiased fashion and on the basis of hepatocyte and endothelial cell markers for scRNA-seq (Fig. 6a, Extended Data Fig. 10a). We then compared engrafted cells to our reference atlas and observed that we



**Fig. 5 | scRNA-seq of patient-derived HCC reveals cancer-specific gene signatures and perturbed cellular phenotypes. a**, Symbol *t*-SNE map highlighting normal liver cells and cells from HCC. *n* = 11,654 cells, *n* = 3 patients with HCC. **b**, Immunostaining for IL32 and CYP2E1 in normal liver and HCC tissue. **c**, GSEA for genes that were differentially expressed between cancer cells from HCC and normal hepatocytes (*n* = 15,442 genes). **d**, GSEA for genes that were differentially expressed between normal endothelial cells and endothelial cells from HCC (*n* = 15,442 genes). **c**, **d**, Benjamini–Hochberg corrected *P* < 0.01; NES, normalized enrichment score; see Methods. **e**, Immunostaining for CLEC4G and PECAM1 in normal liver tissue and HCC tissue. All staining images are from the Human Protein Atlas<sup>31</sup>.

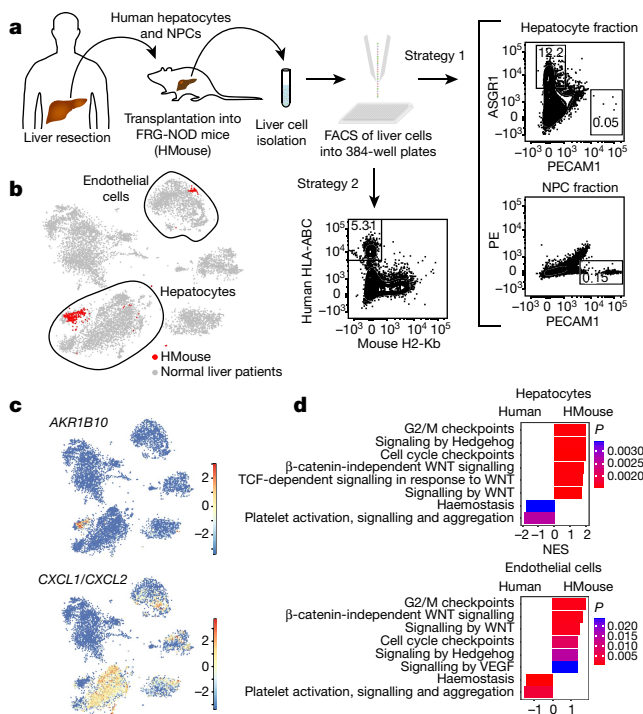
had successfully transplanted both human hepatocytes and endothelial cells (Fig. 6b, Extended Data Fig. 10b, c), which had maintained their fundamental gene signatures, such as expression of *ALB* or *PCK1* and *CLEC4G*, *PECAM1* or *CD34*, respectively (Extended Data Fig. 10b–f). Nevertheless, many genes were differentially expressed in engrafted cells compared to non-engrafted human liver cells; for example, *AKR1B10*, which was also expressed by cancer cells from HCC, was expressed in engrafted cells but not non-engrafted cells (Fig. 6c, Extended Data Fig. 10g). Gene set enrichment analysis (GSEA) of differentially expressed genes revealed that HMouse hepatocytes and endothelial cells showed downregulation of pathways such as haemostasis, and upregulation of WNT and Hedgehog signalling as well as cell cycle genes (Fig. 6d), akin to what we observed in HCC cells and cells from liver organoids.

### Discussion

We have established a human liver cell atlas, revealing heterogeneity within major liver cell populations and the existence of an epithelial progenitor in the adult human liver.

Our atlas reveals transcriptome-wide zonation of hepatocytes and endothelial cells, and suggests that different liver cell types may cooperate to carry out essential functions. Although we could validate predicted zonation profiles with antibody staining, it will be essential to perform more large scale in situ gene expression analysis.

The EPCAM<sup>+</sup>TROP2<sup>int</sup> population is a strong candidate for potential involvement in homeostatic turnover, liver regeneration, disease pathogenesis and tumour formation. Although our in silico analysis and in vitro organoid culture experiments provide evidence that this



**Fig. 6 | Exploring the gene expression signature of human liver cells in a humanized mouse model.** **a**, Outline of the transplantation of human liver cells (hepatocytes and non-parenchymal cells) into the FRG-NOD mouse and the two strategies for sorting engrafted human cells from the mouse liver. **b**, Symbol *t*-SNE map highlighting normal liver cells and cells from the humanized mouse model. The main engrafted cell types (hepatocytes and endothelial cells) are circled. **c**, Expression *t*-SNE maps of *AKR1B10* and *CXCL1/CXCL2*. Colour bar indicates  $\log_2$  normalized expression.  $n = 10,683$  cells. **d**, GSEA of genes that were differentially expressed between hepatocytes and endothelial cells from humanized mouse (HMouse) and patients (Human).  $n = 13,614$  genes; Benjamini–Hochberg corrected  $P < 0.01$ ; see Methods.

population is bipotent, its lineage potential remains to be demonstrated in vivo.

As demonstrated by our HCC analysis, the atlas provides a key reference for the investigation of liver diseases and will contribute to the development of urgently needed human liver models, including organoids and humanized liver chimaeric mouse models.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1373-2>.

Received: 14 May 2018; Accepted: 5 June 2019;  
Published online: 10 July 2019

1. Michalopoulos, G. K. & DeFrances, M. C. Liver regeneration. *Science* **276**, 60–66 (1997).

2. Ryerson, A. B. et al. Annual report to the nation on the status of cancer, 1975–2012, featuring the increasing incidence of liver cancer. *Cancer* **122**, 1312–1337 (2016).
3. Grün, D. & van Oudenaarden, A. Design and analysis of single-cell sequencing experiments. *Cell* **163**, 799–810 (2015).
4. Herman, J. S., Sagar & Grün, D. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat. Methods* **15**, 379–386 (2018).
5. Grün, D. et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**, 251–255 (2015).
6. Jungermann, K. & Kietzmann, T. Zonation of parenchymal and nonparenchymal metabolism in liver. *Annu. Rev. Nutr.* **16**, 179–203 (1996).
7. Gebhardt, R. Metabolic zonation of the liver: regulation and implications for liver function. *Pharmacol. Ther.* **53**, 275–354 (1992).
8. Kietzmann, T. Metabolic zonation of the liver: the oxygen gradient revisited. *Redox Biol.* **11**, 622–630 (2017).
9. Halpern, K. B. et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352–356 (2017).
10. MacParland, S. A. et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat. Commun.* **9**, 4383 (2018).
11. Haghverdi, L., Büttner, M., Wolf, F. A., Büttner, F. & Theis, F. J. Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–848 (2016).
12. Strauss, O., Phillips, A., Ruggiero, K., Bartlett, A. & Dunbar, P. R. Immunofluorescence identifies distinct subsets of endothelial cells in the human liver. *Sci. Rep.* **7**, 44356 (2017).
13. Halpern, K. B. et al. Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. *Nat. Biotechnol.* **36**, 962–970 (2018).
14. Raven, A. et al. Cholangiocytes act as facultative liver stem cells during impaired hepatocyte regeneration. *Nature* **547**, 350–354 (2017).
15. Michalopoulos, G. K., Barua, L. & Bowen, W. C. Transdifferentiation of rat hepatocytes into biliary cells after bile duct ligation and toxic biliary injury. *Hepatology* **41**, 535–544 (2005).
16. Schmelzer, E. et al. Human hepatic stem cells from fetal and postnatal donors. *J. Exp. Med.* **204**, 1973–1987 (2007).
17. Turner, R. et al. Human hepatic stem cell and maturational liver lineage biology. *Hepatology* **53**, 1035–1045 (2011).
18. Grün, D. et al. De novo prediction of stem cell identity using single-cell transcriptome data. *Cell Stem Cell* **19**, 266–277 (2016).
19. Okabe, M. et al. Potential hepatic stem cells reside in EpCAM<sup>+</sup> cells of normal and injured mouse liver. *Development* **136**, 1951–1960 (2009).
20. Cardinale, V. et al. Multipotent stem/progenitor cells in human biliary tree give rise to hepatocytes, cholangiocytes, and pancreatic islets. *Hepatology* **54**, 2159–2172 (2011).
21. Kodama, Y. et al. Hes1 is required for the development of intrahepatic bile ducts. *Gastroenterology* **124**, A123 (2003).
22. Sosa-Pineda, B., Wigle, J. T. & Oliver, G. Hepatocyte migration during liver development requires Prox1. *Nat. Genet.* **25**, 254–255 (2000).
23. Huch, M. et al. Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell* **160**, 299–312 (2015).
24. Betge, J. et al. MUC1, MUC2, MUC5AC, and MUC6 in colorectal cancer: expression profiles and clinical significance. *Virchows Arch.* **469**, 255–265 (2016).
25. Park, S.-W. et al. The protein disulfide isomerase AGR2 is essential for production of intestinal mucus. *Proc. Natl Acad. Sci. USA* **106**, 6950–6955 (2009).
26. Forner, A., Reig, M. & Bruix, J. Hepatocellular carcinoma. *Lancet* **391**, 1301–1314 (2018).
27. Matkowskyj, K. A. et al. Aldoketoreductase family 1B10 (*AKR1B10*) as a biomarker to distinguish hepatocellular carcinoma from benign liver lesions. *Hum. Pathol.* **45**, 834–843 (2014).
28. Rantakari, P. et al. The endothelial protein PLVAP in lymphatics controls the entry of lymphocytes and antigens into lymph nodes. *Nat. Immunol.* **16**, 386–396 (2015).
29. Grompe, M. & Strom, S. Mice with human livers. *Gastroenterology* **145**, 1209–1214 (2013).
30. Azuma, H. et al. Robust expansion of human hepatocytes in *Fah*<sup>−/−</sup>/*Rag2*<sup>−/−</sup>/*Il2rg*<sup>−/−</sup> mice. *Nat. Biotechnol.* **25**, 903–910 (2007).
31. Uhlén, M. et al. Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## METHODS

**Human liver samples.** Human liver tissue samples were obtained from patients who had undergone liver resections between 2014 and 2018 at the Center for Digestive and Liver Disease (Pôle Hépatodigestif) at the Strasbourg University Hospitals, University of Strasbourg, France. For the human liver cell atlas, samples were acquired from patients without chronic liver disease (defined as liver damage lasting over a period of at least six months), genetic haemochromatosis with homozygote C282Y mutation, active alcohol consumption ( $>20$  g/d in women and  $>30$  g/d in men), active infectious disease, pregnancy or any contraindication for liver resection. All patients provided written informed consent. The protocols followed the ethical principles of the declaration of Helsinki and were approved by the local Ethics Committee of the University of Strasbourg Hospitals and by the French Ministry of Education and Research (CPP 10-17, Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche; approval number DC-2016-2616). Data protection was performed according to EU legislation regarding privacy and confidentiality during personal data collection and processing (Directive 95/46/EC of the European Parliament and of the Council of the 24 October 1995). Samples (BP1) and tissue blocks were obtained from Biopredic International.

**Tissue dissociation and preparation of single-cell suspensions.** Human liver specimens obtained from resections were perfused for 15 min with calcium-free 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid buffer containing 0.5 mM ethylene glycol tetraacetic acid (Fluka) followed by perfusion with 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid containing 0.5 mg/ml collagenase (Sigma-Aldrich) and 0.075%  $\text{CaCl}_2$  at 37°C for 15 min as previously described<sup>32</sup>. Then the cells were washed with phosphate-buffered saline (PBS) and nonviable cells were removed by Percoll (Sigma-Aldrich) gradient centrifugation. Part of the isolated cells was further separated into primary human hepatocytes (PHHs) and non-parenchymal cells (NPCs) by an additional centrifugation step at 50g for 5 min at 4°C. The isolated cells were frozen in liquid nitrogen using the CryoStor CS10 solution (Sigma-Aldrich). Human HCC tissues were dissociated using the gentleMACS Dissociator (Miltenyi Biotec) according to the manufacturer's protocol.

**Transplantation of human cells into *Fah*<sup>-/-</sup>/*Rag2*<sup>-/-</sup>/*Il2rg*<sup>-/-</sup> mice.** *Fah*<sup>-/-</sup>/*Rag2*<sup>-/-</sup>/*Il2rg*<sup>-/-</sup> non-obese diabetic (ERG-NOD) breeding mice were kept at the Inserm Unit 1110 SPF animal facility and maintained with 16 mg/l of 2-(2-nitro-4-trifluoro-methyl-benzoyl)-1,3-cyclohexanedione (NTBC; Swedish Orphan Biovitrum) in drinking water. Six-week-old male and female mice were intravenously injected with  $1.5 \times 10^9$  plaque-forming units (pfu) of an adenoviral vector encoding the secreted form of the human urokinase-like plasminogen activator (Ad-uPA)<sup>33</sup>. Forty-eight hours later,  $10^6$  PHHs and  $2 \times 10^5$  NPCs from the same liver donor, isolated as previously described, were injected intrasplenically via a 27-gauge needle. For the procedure, mice were kept under gaseous isoflurane anaesthesia and received a subcutaneous injection of buprenorphine (0.1 mg/kg). After transplantation, the NTBC was gradually decreased and completely withdrawn in 7 days. The success of the transplantation was evaluated 2 months after the procedure by dosing human albumin in mouse serum as previously described<sup>34</sup>. This procedure was approved by the local ethics committee and authorized by the French ministry of higher education and research (authorization number #4485-201603115352125 v3). All procedures are consistent with the guidelines set by the Panel on Euthanasia (AVMA) and the NIH Guide for the Care and Use of Laboratory Animals as well as the Declaration of Helsinki in its latest version, and the Convention of the Council of Europe on Human Rights and Biomedicine. The animal research was performed within the regulations and conventions protecting animals used for research purposes (Directive 86/609/EEC), as well as with European and national laws regarding work with genetically modified organs. The animal facility at the University of Strasbourg, Inserm U1110 has been approved by the regional government (Préfecture) and granted authorization number E67-482-7, 2017/08/24.

**Mouse liver cell isolation.** The anaesthetized animal was restrained and the skin sprayed with 70% ethanol. The liver and other inner organs were revealed by cutting through the skin and peritoneum. A 24G needle was carefully inserted into the inferior vena cava and secured with a clamp, and chelating solution (0.05 M HEPES pH 7.2, 10 mM EGTA in HBSS without  $\text{CaCl}_2$  and  $\text{MgCl}_2$ ) was run at a low speed (1.5–2 ml/min). The portal vein was then cut and perfusion speed was increased to a flow rate of 7 ml/min. After that, the diaphragm was cut and the anterior vena cava clamped. The chelating perfusion was run for 7 min and then switched to collagenase solution (0.05 M HEPES pH 7.2, 4.7 mM  $\text{CaCl}_2$ , 20 µg/ml Liberase, Sigma LIBTM-RO) at a flow rate of 7 ml/min for 7 min. The liver was then removed and passed through a 70-µm cell strainer with 10 ml ice-cold PBS without  $\text{CaCl}_2$  and  $\text{MgCl}_2$ . The resulting single-cell suspension was centrifuged at 300g for 5 min at 4°C and washed twice with ice-cold PBS.

**FACS.** Liver cells were sorted from mixed, hepatocyte, and non-parenchymal cell fractions on an Aria Fusion I using a 100-µm nozzle. Cells from the HCC samples were not fractionated and were sorted directly after tissue digestion. Zombie Green

(Biolegend) was used as a viability dye. Cells were stained with human-specific antibodies against CD45 (Biolegend, cat. no. 304023), PECAM1 (Biolegend, cat. no. 303111), CD34 (Biolegend, cat. no. 343609), CLEC4G (R&D systems, cat. no. FAB2947A), ASGR1 (BD Biosciences, cat. no. 563655), EPCAM (R&D systems, cat. no. FAB960R), and TROP2 (Biolegend, cat. no. 363803). Organoids were stained with antibodies against EPCAM and TROP2. For the humanized mouse samples, cells were stained either with antibodies against ASGR1 and PECAM1 or with human HLA-ABC (BD Biosciences, cat. no. 740407) and mouse H2-Kb (BD Biosciences, cat. no. 553570). Viable cells were sorted in an unbiased fashion or from specific populations based on the expression of markers into the wells of 384-well plates containing lysis buffer.

**Single-cell RNA amplification and library preparation.** Single-cell RNA sequencing was performed according to the mCEL-Seq2 protocol<sup>4,35</sup>. Viable liver cells were sorted into 384-well plates containing 240 nl primer mix and 1.2 µl PCR encapsulation barrier, Vapour-Lock (QIAGEN) or mineral oil (Sigma-Aldrich). Sorted plates were centrifuged at 2,200g for a few minutes at 4°C, snap-frozen in liquid nitrogen and stored at -80°C until they were processed. We used 160 nl reverse transcription reaction mix and 2.2 µl second-strand reaction mix to convert RNA into cDNA. cDNA from 96 cells was pooled together before clean up and in vitro transcription, generating four libraries from one 384-well plate. We used 0.8 µl AMPure/RNAClean XP beads (Beckman Coulter) per 1 µl sample during all purification steps including library cleanup. Other steps were performed as described in the protocol<sup>4,35</sup>. Libraries were sequenced on an Illumina HiSeq 2500 and 3000 sequencing system (paired-end multiplexing run, high output mode) at a depth of ~150,000–200,000 reads per cell.

**Quantification of transcript abundance.** Paired-end reads were aligned to the transcriptome using bwa (version 0.6.2-r126) with default parameters<sup>36</sup>. The transcriptome contained all gene models based on the human whole genome ENCODE V24 release. All isoforms of the same gene were merged to a single gene locus. Subsequently, gene loci with  $>75\%$  sequence overlap were merged. The right mate of each read pair was mapped to the ensemble of all gene loci and to the set of 92 ERCC spike-ins in the sense direction. Reads mapping to multiple loci were discarded. The left read contains the barcode information: the first six bases corresponded to the unique molecular identifier (UMI) followed by six bases representing the cell-specific barcode. The remainder of the left read contains a polyT stretch. The left read was not used for quantification. For each cell barcode, the number of UMIs per transcript was counted and aggregated across all transcripts derived from the same gene locus. The number of observed UMIs was converted into transcript counts using binomial statistics<sup>37</sup>.

**Single-cell RNA sequencing data analysis.** Overall, 10,372 cells passed the quality control threshold of  $>1,000$  transcripts (Poisson-corrected UMIs<sup>37</sup>) for the normal human liver cell atlas. For cells from the organoids, 1,052 cells passed the quality control thresholds. For cells from HCC, 1,282 cells passed the quality control threshold of  $>1,000$  transcripts. For cells from the humanized mouse, 311 cells passed the quality control threshold of  $>1,000$  transcripts. All the datasets were analysed using RaceID3<sup>4</sup>. For normalization, the total transcript counts in each cell were normalized to 1 and multiplied by the minimum total transcript count across all cells that passed the quality control threshold ( $>1,000$  transcripts per cell). Prior to normalization, cells expressing  $>2\%$  of *KCNQ1OT1* transcripts, a previously identified marker of low quality cells<sup>18</sup>, were removed from the analysis. Moreover, transcripts correlating to *KCNQ1OT1* with a Pearson's correlation coefficient of  $>0.4$  were also removed. RaceID3 was run with the following parameters: mintotal = 1000, minexpr = 2, minnumber = 10, outmnc = 2, cln = 15.

**Diffusion pseudo-time analysis and self-organizing maps.** Diffusion pseudotime (dpt) analysis<sup>11</sup> was implemented and diffusion maps generated using the destiny R package. The number of nearest neighbours,  $k$ , was set to 100. SOMs were generated using the FateID package on the basis of the ordering computed by dpt as input. Only genes with  $>2$  counts after size normalization in at least a single cell were included for the SOM analysis. In brief, smooth zonation profiles were derived by applying local regression on normalized transcript counts after ordering cells by dpt. Next, a one-dimensional SOM with 200 nodes was computed on these profiles after z-transformation. Neighbouring nodes were merged if the Pearson's correlation coefficient of the average profiles of these nodes exceeded 0.85. The remaining aggregated nodes represent the gene modules shown in the SOM figures.

$P$  values for the significance of zonation were derived by binning dpt-ordered profiles into three equally sized bins to perform ANOVA. The resulting  $P$  values were corrected for multiple testing using the Benjamini–Hochberg method. Increasing the number of bins produced similar results.

**Conservation of zonation between human and mouse.** Expression data from Halpern et al.<sup>9</sup> (GEO accession code GSE84498) were used for analysing the evolutionary conservation of hepatocyte zonation between human and mouse. The transcript count data were analysed using RaceID3 to determine cell types, with parameter mintotal = 1,000 and cln = 6. A subgroup of clusters was identified as hepatocytes on the basis of marker gene expression and used for dpt and SOM



analysis, as was done for the human data. To obtain a similar number of genes, only genes with at least 1.5 counts after size normalization in at least a single cell were included. To identify orthologues between human and mouse for the references used in this study and by Halpern et al.<sup>9</sup> as provided by the authors, we first identified pairs of orthologues based on identical gene identifiers upon capitalization of all letters. We further computed mutual blastn (run with default) best hits. The final list comprises 16,670 pairs of orthologues.

Conservation of zonation was assessed using Pearson's correlation of zoned expression profiles after binning the human data into nine equally sized bins, akin to the nine zones derived in Halpern et al.<sup>9</sup>. Conservation of zonation of endothelial cells was evaluated based on published mouse data from Halpern et al.<sup>13</sup> using classification into four spatially stratified populations. To calculate Pearson's correlation coefficient between human and mouse endothelial cells, a diffusion-pseudotime analysis was performed for all human cells mapping to endothelial cell clusters and these profiles were discretized into four equally sized bins.

**Lineage analysis of the EPCAM<sup>+</sup> compartment.** For a separate analysis of the EPCAM<sup>+</sup> population, all cells from clusters 4, 7, 24 and 39 were extracted and reanalysed using RaceID3<sup>4</sup> with the parameters mintotal = 1000 and minexpr = 2, minnumber = 10 outmnc = 2, and default parameters otherwise. StemID2<sup>4</sup> was run on these clusters with cthr = 10, nmode = TRUE and knn = 3. FateID<sup>4</sup> was run on the filtered and feature-selected expression matrix from RaceID3, with target clusters inferred by FateID using *ASGR1* plus *ALB* and *CXCL8* plus *MMP7* as markers for hepatocyte and cholangiocyte lineage target clusters, respectively. Using *KRT19* and *CFTR* as mature cholangiocyte markers yields highly similar results.

**Differential gene expression analysis.** Differential gene expression analysis between cells and clusters was performed using the diffexpnb function from the RaceID package. First, negative binomial distributions reflecting the gene expression variability within each subgroup were inferred on the basis of the background model for the expected transcript count variability computed by RaceID3. Using these distributions, a *P* value for the observed difference in transcript counts between the two subgroups was calculated and corrected for multiple testing using the Benjamini–Hochberg method as described<sup>38</sup>.

**Pathway enrichment analysis and gene set enrichment analysis.** Symbol gene IDs were first converted to Entrez gene IDs using the clusterProfiler<sup>39</sup> package. Pathway enrichment analysis and GSEA<sup>40,41</sup> were implemented using the ReactomePA<sup>42</sup> package. Pathway enrichment analysis was done on genes taken from the different modules in the SOMs. GSEA was done using the differentially expressed genes inferred by the diffexpnb function from the RaceID package.

**Validation of protein expression using the Human Protein Atlas.** Immunostaining images were collected from the Human Protein Atlas<sup>31</sup> (<https://www.proteinatlas.org>).

**Immunofluorescence.** Human liver tissue was fixed overnight in 3.7% formaldehyde (Fig. 3g) or cryosectioned and fixed in 2.5% paraformaldehyde for 20 min (Extended Data Fig. 7e). The tissue was embedded in OCT and stored at –80°C. The tissue was cryosectioned into 7-μm sections. The tissue was washed twice for 5 min in 0.025% Triton 1 × PBS. The tissue was then blocked in 10% FBS with 1% BSA in 1 × PBS for 2 h at room temperature. The dilution used for the anti-human KRT19 (HPA002465, Sigma, Fig. 3g; MA5-12663, Invitrogen, Extended Data Fig. 7e) and EPCAM (SAB4200704, Sigma, Fig. 3g; PA5-19832, Invitrogen, Extended Data Fig. 7e) antibodies was 1:100 in 100 μl 1 × PBS with 1% BSA. The antibodies were incubated overnight at 4°C in the dark. The tissue was washed twice with 0.025% Triton 1 × PBS and then incubated with secondary antibodies donkey anti-rabbit IgG-AF488 ((A21206, Thermo Fisher Scientific), Fig. 3g) and goat anti-mouse IgG-AF568 ((A11019, Thermo Fisher Scientific), Fig. 3g) or sheep anti-mouse IgG-AF488 ((515-545-062, Jackson ImmunoResearch), Extended Data Fig. 7e) at 1:200 dilution and donkey anti-rabbit IgG-RRX ((711-295-152, Jackson ImmunoResearch), Extended Data Fig. 7e) at 1:100 dilution in 1 × PBS with 1% BSA for 1 h at room temperature. The tissue was then washed twice with 0.025% Triton 1 × PBS. DAPI Fluoromount-G (Southern Biotech) was added to the tissue and a coverslip placed on top. Imaging was done using a Zeiss confocal microscope LSM780 (Fig. 3g) or ZEISS Axio Vert.A1 (Extended Data Fig. 7e). Images were taken at 63× magnification.

**Organoid culturing.** Organoid culturing was done as previously described<sup>43</sup>. The cell populations from the EPCAM<sup>+</sup> compartment were sorted on an Aria Fusion I using a 100-μm nozzle into tubes containing culture medium supplemented with 10 μM ROCK inhibitor (Y-27632) (Sigma-Aldrich). After sorting, cells were centrifuged in order to remove the medium and then resuspended in 25 μl Matrigel. Droplets of the Matrigel solution containing the cells were added to the wells of a 24-well suspension plate and incubated for 5–10 min at 37°C until the Matrigel solidified. Droplets were overlaid with 250 μl liver isolation medium and then incubated at 37°C, 5% CO<sub>2</sub>. After 3–4 days, the liver isolation medium was replaced with liver expansion medium. For the single-cell culture, from each patient, single cells from the TROP2<sup>int</sup> gate were sorted into the wells of a

non-tissue-culture-treated 96-well plate containing medium with 5% Matrigel. Organoids were passaged 14 days after isolation and then passaged multiple times 5–7 days after splitting. For FACS, single-cell suspensions were prepared from the organoids by mechanical dissociation followed by TrypLE (Life Technologies) digestion as previously described<sup>43</sup>. Organoid cells were sequenced 5 days after splitting and 17 days after initially sorting the cells for the culture.

**Step-by-step protocol.** A detailed protocol for scRNA-seq of cryopreserved human liver cells is available at Protocol Exchange<sup>44</sup>.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

Data generated during this study have been deposited in the Gene Expression Omnibus (GEO) with the accession code GSE124395. The human liver cell atlas can be interactively explored at <http://human-liver-cell-atlas.ie-freiburg.mpg.de/>.

- Krieger, S. E. et al. Inhibition of hepatitis C virus infection by anti-claudin-1 antibodies is mediated by neutralization of E2-CD81-claudin-1 associations. *Hepatology* **51**, 1144–1157 (2010).
- Lieber, A., Peeters, M. J., Gown, A., Perkins, J. & Kay, M. A. A modified urokinase plasminogen activator induces liver regeneration without bleeding. *Hum. Gene Ther.* **6**, 1029–1037 (1995).
- Mailly, L. et al. Clearance of persistent hepatitis C virus infection in humanized mice using a claudin-1-targeting monoclonal antibody. *Nat. Biotechnol.* **33**, 549–554 (2015).
- Hashimshony, T. et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-seq. *Genome Biol.* **17**, 77 (2016).
- Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
- Grün, D., Kester, L. & van Oudenaarden, A. Validation of noise models for single-cell transcriptomics. *Nat. Methods* **11**, 637–640 (2014).
- Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
- Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
- Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
- Mootha, V. K. et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
- Yu, G. & He, Q. Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.* **12**, 477–479 (2016).
- Broutier, L. et al. Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation. *Nat. Protocols* **11**, 1724–1743 (2016).
- Aizarani, N. et al. Protocol for single-cell RNA-sequencing of cryopreserved human liver cells. *Protoc. Exch.* <https://doi.org/10.21203/rs.2.9620/v1> (2019).

**Acknowledgements** This study was supported by the Max Planck Society, the German Research Foundation (DFG) (SPP1937 GR4980/1-1, GR4980/3-1, and GRK2344 MelnBio), the DFG under Germany's Excellence Strategy (CIBSS – EXC-2189 – Project ID 390939984), and the Behrens-Weise-Foundation (all to D.G.); and by ARC, Paris and Institut Hospital-Universitaire, Strasbourg (TheraHCC2.0 IHUARC IHU201301187 and IHUARC2019 IHU201901299 to T.F.B.), the European Union (ERC-AdG-2014-671231-HEPCIR, EU H2020-667273-HEPCAR and ERC-PoC-2016-PRELICAN to T.F.B.), Agence nationale de recherches sur le sida et les hépatites virales (ANRS, ECT35076) and the Foundation of the University of Strasbourg. This work was done under the framework of the LABEX ANR-10-LABX-0028, HEPSYS and Inserm Plan Cancer (Plan Cancer 2014-2019, Action 13.1, appel à projets 2018) and benefits from funding from the state managed by the French National Research Agency as part of the Investments for the future. Institut Universitaire France (IUF) to T.F.B. We thank S. Hobitz, K. Schuldes (MPI-IE FACS facility) and U. Bönsch (MPI-IE Deep Sequencing facility); the CRB (Centre de Ressources Biologiques-Biological Resource Centre of the Strasbourg University Hospitals) for the management of regulatory requirements of patient-derived liver tissue; C. Fauvel and L. Heydmann for their contributions to the initial single-cell isolations; F. Juehling, F. H. T. Duong and C. Schuster for helpful discussions; the patients for providing informed consent to participate in the study; and the nurses, technicians and medical doctors of the hepatobiliary surgery and pathology services of Strasbourg University Hospitals for their support. This publication is part of the Human Cell Atlas (<https://www.humancellatlas.org/publications>).

**Author contributions** T.F.B. and D.G. conceived the study. N.A. designed, optimized and performed cell sorting, scRNA-seq, organoid culture and immunofluorescence, provided validation using the Human Protein Atlas, and performed computational analysis and interpretation of the data. A.S. managed the supply of patient material, isolated single cells from patient

tissue, performed animal experiments and carried out immunofluorescence. S. contributed to scRNA-seq analyses and performed scRNA-seq experiments. L.M. performed animal experiments. J.S.H. created the web interface. S.D. isolated single cells from patient tissues. P.P. performed liver resections and provided patient liver samples. T.F.B. established the liver tissue supply pipeline and supervised the animal experiments. D.G. analysed and interpreted the data and supervised experiments and analysis by N.A., S. and J.S.H. D.G., N.A. and T.F.B. coordinated and led the study. N.A. and D.G. wrote the manuscript with input from S., A.S. and T.F.B.

**Competing interests** The authors declare no competing interests.

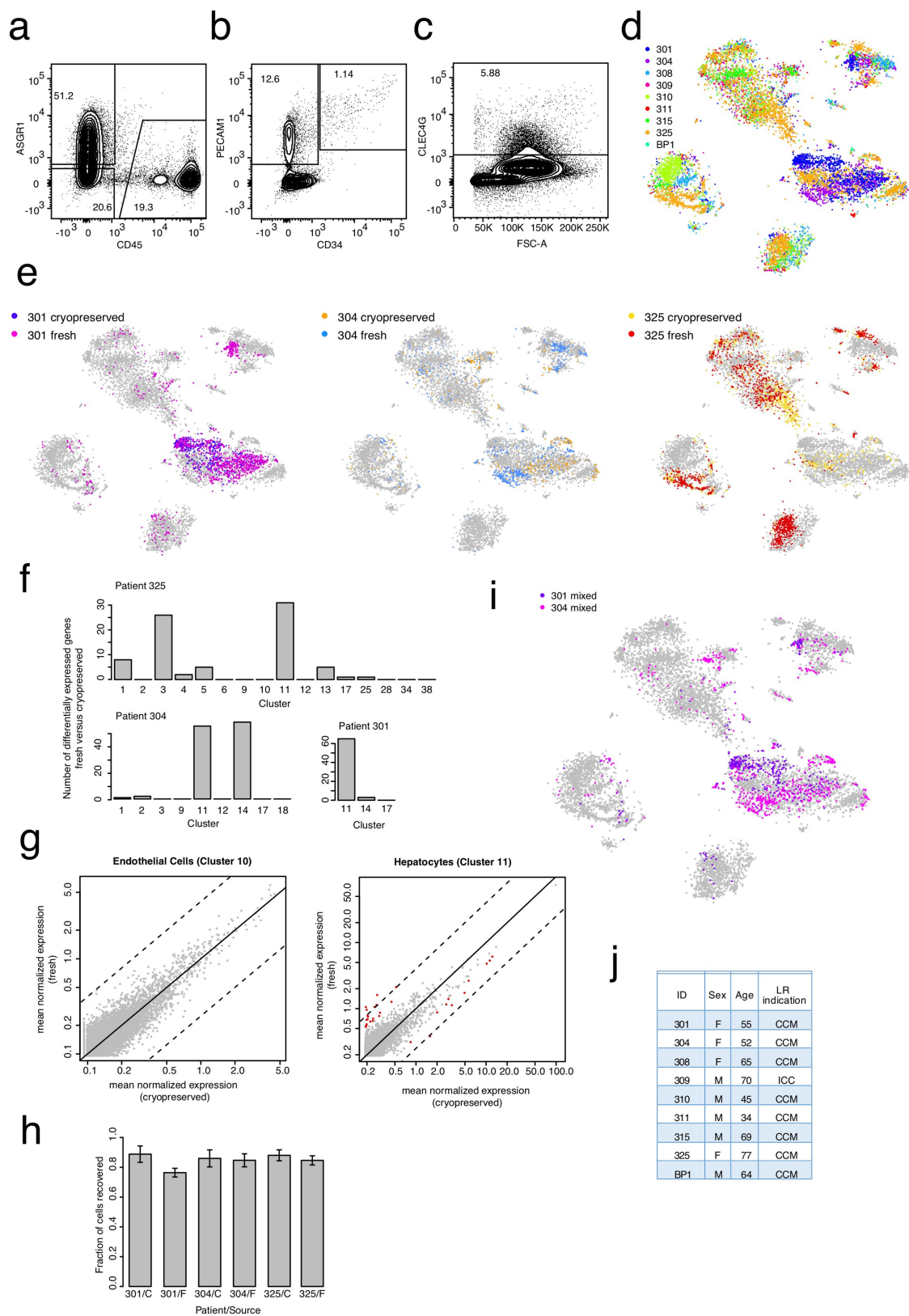
#### **Additional information**

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1373-2>.

**Correspondence and requests for materials** should be addressed to T.F.B. or D.G.

**Peer review information** *Nature* thanks Meritxell Huch, Shalev Itzkovitz and the other anonymous reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

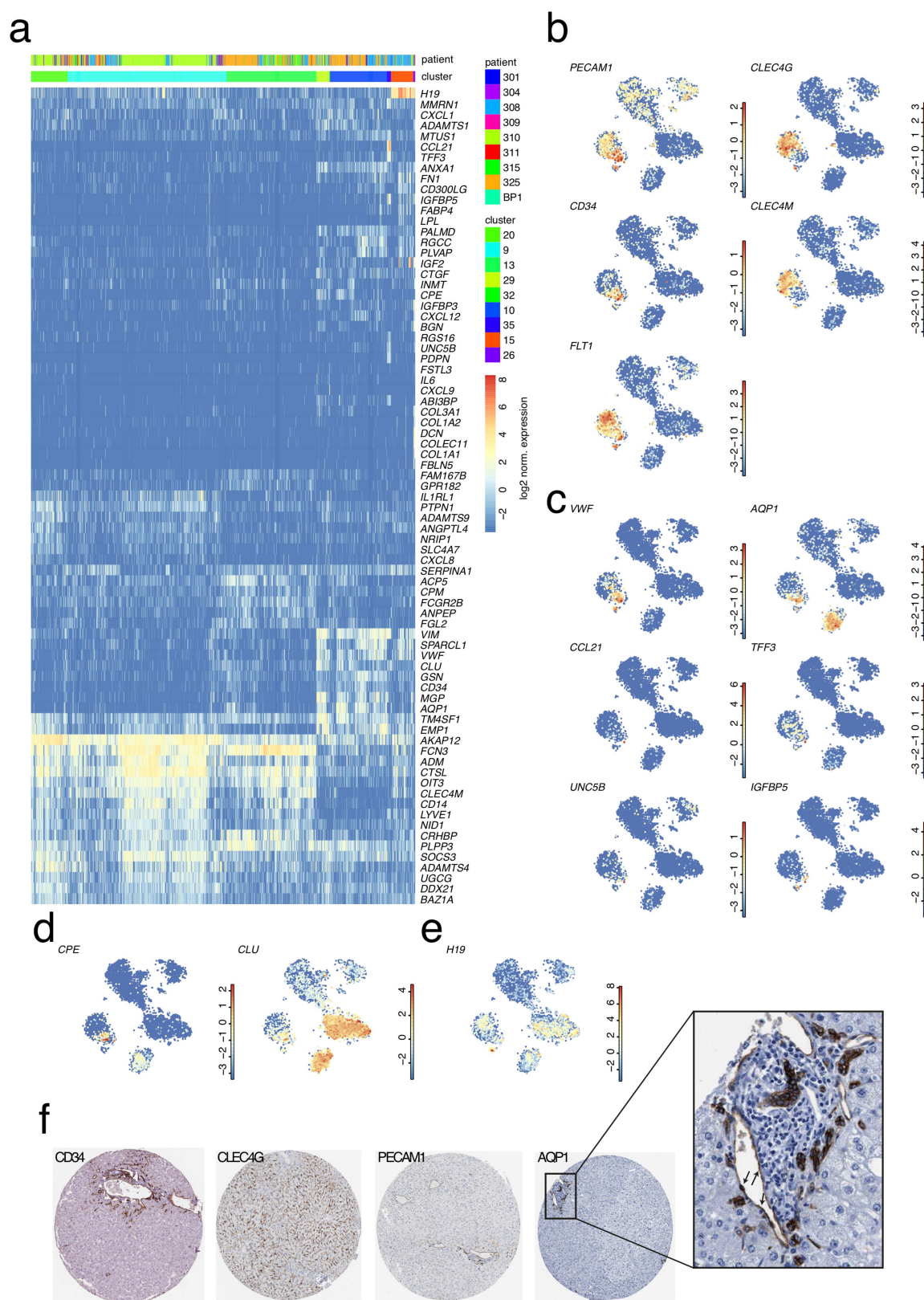


Extended Data Fig. 1 | See next page for caption.

**Extended Data Fig. 1 | scRNA-seq analysis of normal liver resection specimens from nine adult patients.** **a**, FACS plot for CD45 and ASGR1 staining from a mixed fraction (hepatocyte and non-parenchymal cells). **b**, FACS plot for PECAM1 and CD34 staining from a mixed fraction. **c**, FACS plot for CLEC4G staining from a mixed fraction. **a–c**,  $n = 6$  independent experiments. **d**,  $t$ -SNE map showing the IDs of the nine patients from whom the cells were sequenced. Cells were sequenced from freshly prepared single-cell suspensions for patients 301, 304, 325 and BP1, and from cryopreserved single-cell suspensions for patients 301, 304, 308, 309, 310, 311, 315 and 325. Cells were sorted and sequenced mainly in an unbiased fashion from non-parenchymal cell, hepatocyte and mixed fractions for patients 301 and 304. Non-parenchymal and mixed fractions were used to sort specific populations on the basis of markers. CD45<sup>−</sup> and CD45<sup>+</sup> cells were sorted from all patients. CLEC4G<sup>+</sup> LSECs were sorted by FACS from patients 308, 310, 315 and 325. EPCAM<sup>+</sup> cells were sorted by FACS from patients 308, 309, 310, 311, 315 and 325. **e**,  $t$ -SNE map highlighting data for fresh and cryopreserved cells from patients 301, 304 and 325. Although minor shifts in frequencies within cell populations are visible, transcriptomes of fresh and cryopreserved cells co-cluster. Differential gene expression analysis of fresh versus cryopreserved cells, for example, for endothelial cells of patient 325 in cluster 10 (**f**), did not reveal any differentially expressed genes. **d, e**,  $n = 10,372$  cells. **f**, Bar plot showing the number of differentially

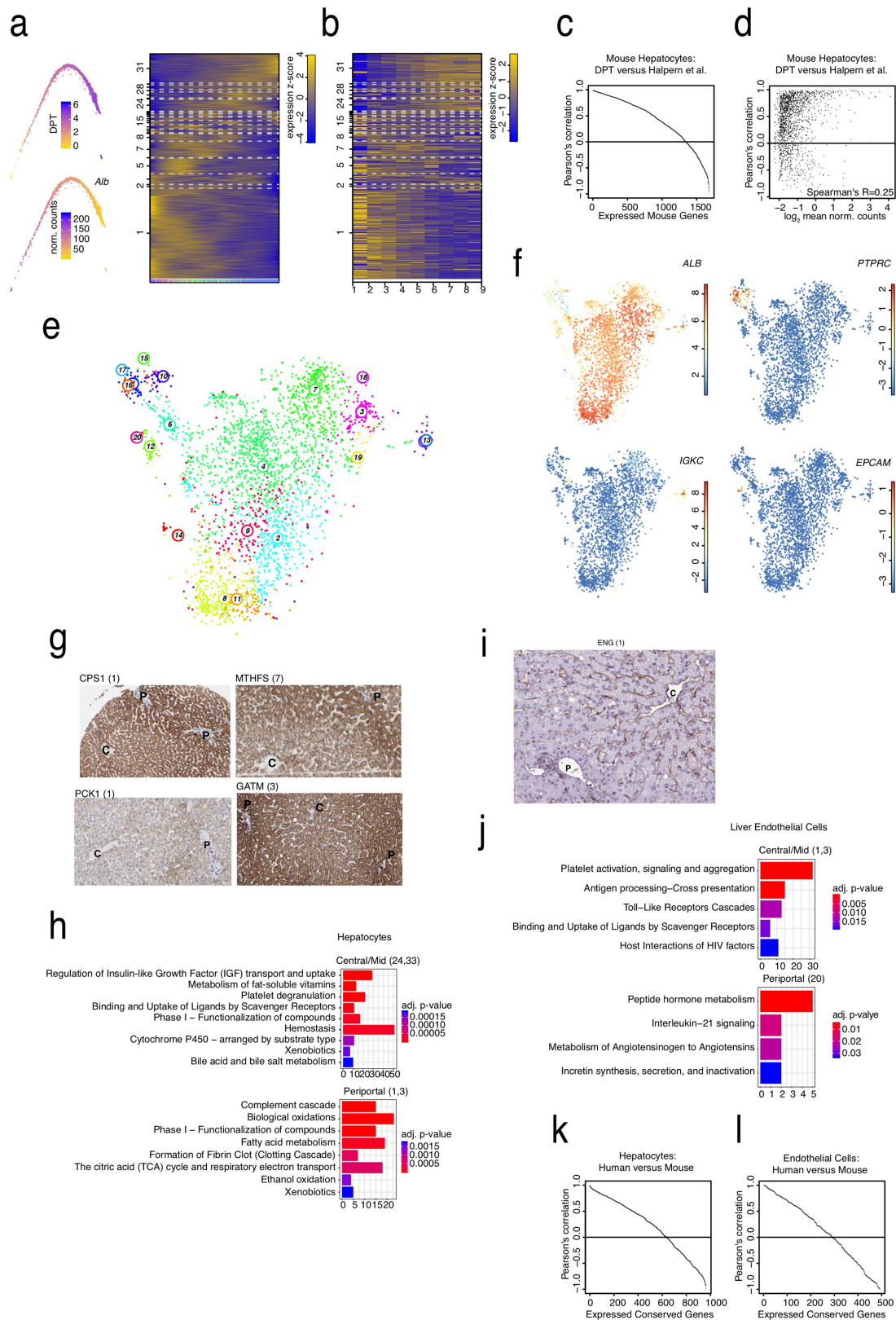
expressed genes (Benjamini–Hochberg corrected  $P < 0.01$ ; see Methods) between fresh and cryopreserved cells within each cluster for patient 325 (top;  $n = 2,248$  cells) and patients 304 ( $n = 959$  cells) and 301 ( $n = 1,329$  cells) (bottom). Only clusters with more than five cells from fresh and cryopreserved samples were included for the computation. **g**, Scatter plot of mean normalized expression across fresh and cryopreserved cells from patient 325 in endothelial cells of cluster 10 (no differentially expressed genes, left;  $n = 101$  cells) and cluster 11 (maximal number of differentially expressed genes across all clusters, right;  $n = 272$  cells). Red dots indicate differentially expressed genes (Benjamini–Hochberg corrected  $P < 0.01$ ; see Methods). Diagonal (solid black line) and  $\log_2$  fold changes of 4 (broken black lines) are indicated. Almost all differentially expressed genes for cluster 11 exhibit  $\log_2$  fold changes of less than 4. **h**, Bar plot showing the fraction of sorted cells which passed quality filtering (see Methods) after scRNA-seq. Error bars are derived from the sampling error assuming binomial counting statistics. F, fresh samples; C, cryopreserved samples. **i**,  $t$ -SNE map highlighting cells sequenced from mixed plates representing unbiased samples for patients 301 and 304. Without any enrichment strategy, hepatocytes and immune cells strongly dominate and endothelial cells and EPCAM<sup>+</sup> cells are rarely sequenced. **j**, Table of patient information. CCM, colon cancer metastasis; ICC, intrahepatic cholangiocarcinoma; LR, liver resection.





**Extended Data Fig. 2 | The endothelial cell compartment is a heterogeneous mixture of subpopulations. a**, Expression heat map of genes upregulated in endothelial cell clusters (Benjamini–Hochberg corrected  $P < 0.01$ ;  $n = 1,830$  cells; see Methods). For each cluster the top ten upregulated genes were extracted and expression of the joint set is shown in the heat map across all endothelial cell clusters. Genes were ordered by hierarchical clustering. **b**, Expression  $t$ -SNE maps for the LSEC and MaVEC marker genes *PECAM1*, *CLEC4G*, *CD34*, *CLEC4M* and

*FLT1*. **c**, Expression  $t$ -SNE maps for *VWF*, *AQP1*, *CCL21*, *TFF3*, *UNC5B* and *IGFBP5*. **d**, Expression  $t$ -SNE maps for *CPE* and *CLU*. **e**, Expression  $t$ -SNE map for *H19*. **b–e**, Colour bars indicate  $\log_2$  normalized expression.  $n = 10,372$  cells. **f**, Immunostaining of *CD34*, *CLEC4G*, *PECAM1* and *AQP1* in normal liver tissue from the Human Protein Atlas. The portal area for *AQP1* is enlarged to show positive staining of both bile duct cells and portal MaVECs (black arrows).



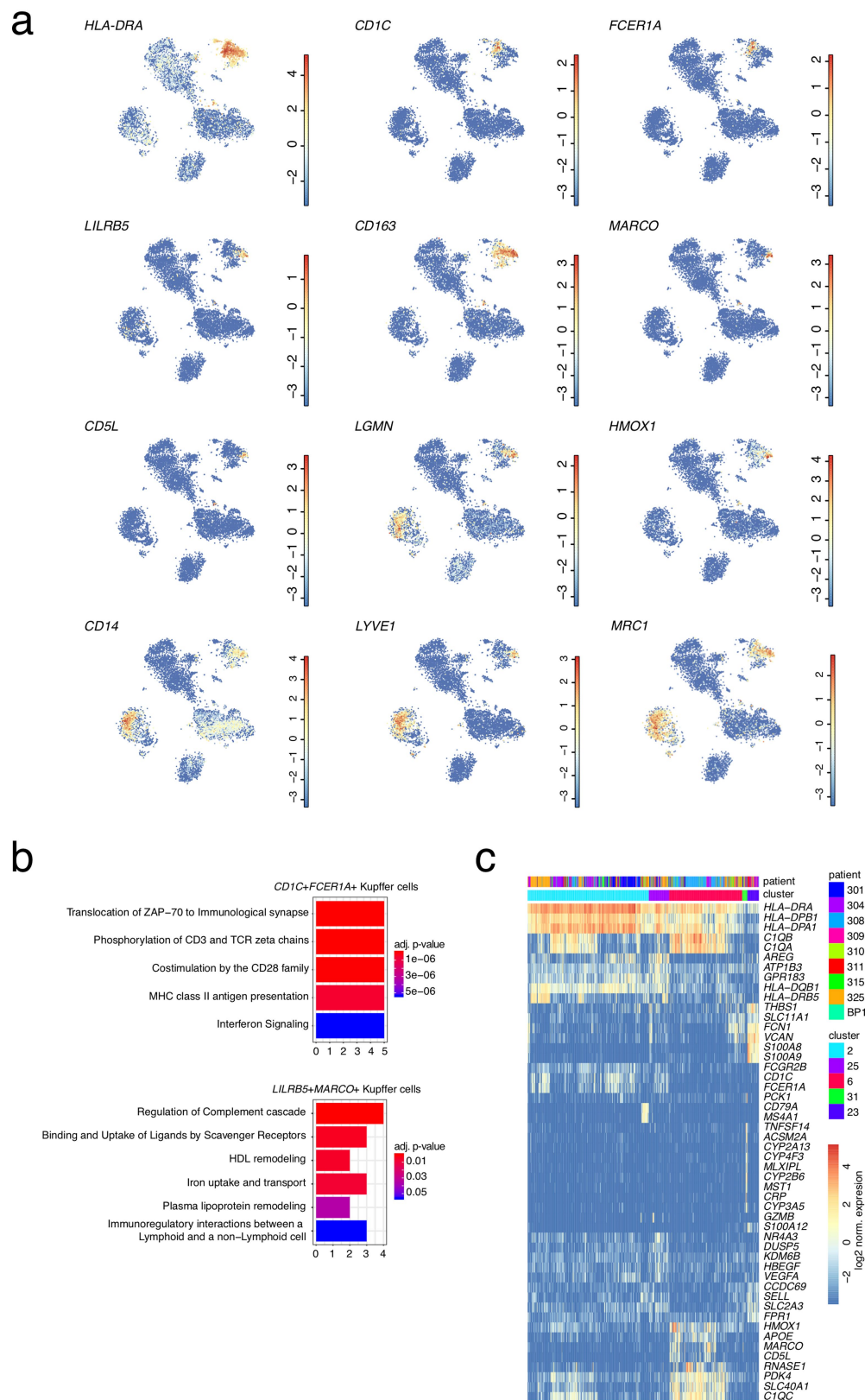
Extended Data Fig. 3 | See next page for caption.

**Extended Data Fig. 3 | Evolutionary conservation of zonation profiles.**

**a**, Diffusion maps highlighting inferred dpt and *Alb* expression (left), and a self-organizing map for mouse hepatocyte single-cell RNA-seq data<sup>9</sup> (right; see Methods). See Fig. 2 for details. **b**, Heat map showing the spatial hepatocyte zonation profiles (nine zones) inferred by Halpern et al.<sup>9</sup> using the same ordering of genes as in **a**. **c**, Pearson's correlation coefficient of zonation profiles inferred by Halpern et al.<sup>9</sup> and our dpt approach after discretizing dpt-inferred zonation profiles into nine equally sized bins. We found that 1,347 out of 1,684 genes (80%) above the expression cutoff exhibited a positive correlation between the two methods. **d**, Pearson's correlation coefficient as a function of average normalized expression. Negative correlations are enriched at low expression, and Pearson's correlation of zonation profiles positively correlates with expression (Spearman's  $R = 0.25$ ;  $n = 1,684$  genes). **e**, *t*-SNE map of single-cell transcriptomes highlighting the clusters generated by RaceID3, run separately on hepatocytes (clusters 11, 14, and 17 in Fig. 1c). The map reveals a major group of hepatocyte clusters and a number of small clusters that co-express T cell-related genes, B cell-related genes or progenitor genes. **f**, *t*-SNE maps highlighting the expression of *ALB*, the immune cell marker gene *PTPRC*, the B cell marker gene *IGKC*, and the progenitor marker gene *EPCAM*. The colour bar indicates  $\log_2$  normalized expression. Co-expression of hepatocyte and immune cell markers could

either indicate the presence of doublets or be due to spillover of highly expressed genes such as *ALB* between cells during library preparation. For the zonation analysis (Fig. 2), only cells in clusters 3, 7, 19, 4, 2, 9, 8 and 11 from the map in **e** were included. **e**, **f**,  $n = 3,040$  cells. **g**, Immunostaining for the periportal markers CPS1, PCK1, MTHFS, and GATM from the Human Protein Atlas<sup>31</sup>. The zonation module containing each gene in the SOM (Fig. 2a) is indicated in parentheses. P, portal tracts; C, central veins. **h**, Pathways enriched for genes in hepatocyte central/mid modules 24 and 33 (top;  $n = 659$  genes) and periportal modules 1 and 3 (bottom;  $n = 422$  genes) (compare with Fig. 2a). **i**, Immunostaining of the central marker ENG from the Human Protein Atlas<sup>31</sup>. The zonation module in the SOM (Fig. 2b) is indicated in parentheses. **j**, Pathways enriched for genes in endothelial central/mid modules 1 and 3 (top;  $n = 422$  genes) and periportal module 20 (bottom;  $n = 73$  genes) (compare with Fig. 2b). **h**, **j**, *P* values in the pathway enrichment analysis were calculated using a hypergeometric model and adjusted using the Benjamini–Hochberg method (see Methods). **k**, Pearson's correlation coefficient of hepatocyte zonation profiles of orthologue pairs of human and mouse genes. Mouse data are from Halpern et al.<sup>9</sup> ( $n = 967$  genes). **l**, Pearson's correlation coefficient of endothelial cell (including MVECs and LSECs) zonation profiles of orthologue pairs of human and mouse genes ( $n = 977$  genes). Mouse data are from Halpern et al.<sup>13</sup>. See Methods for details.

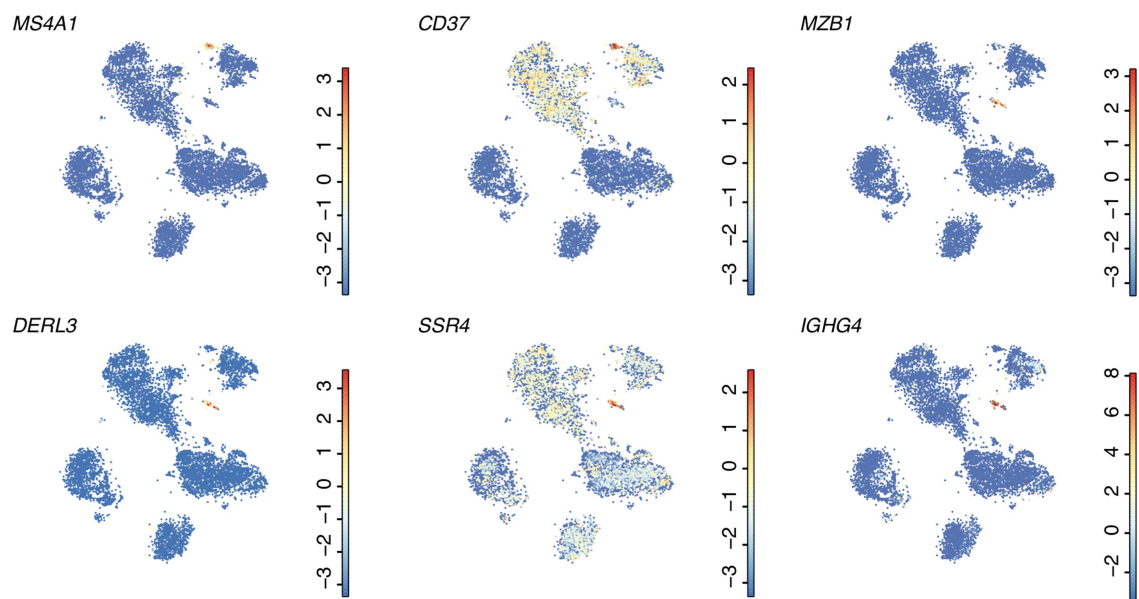




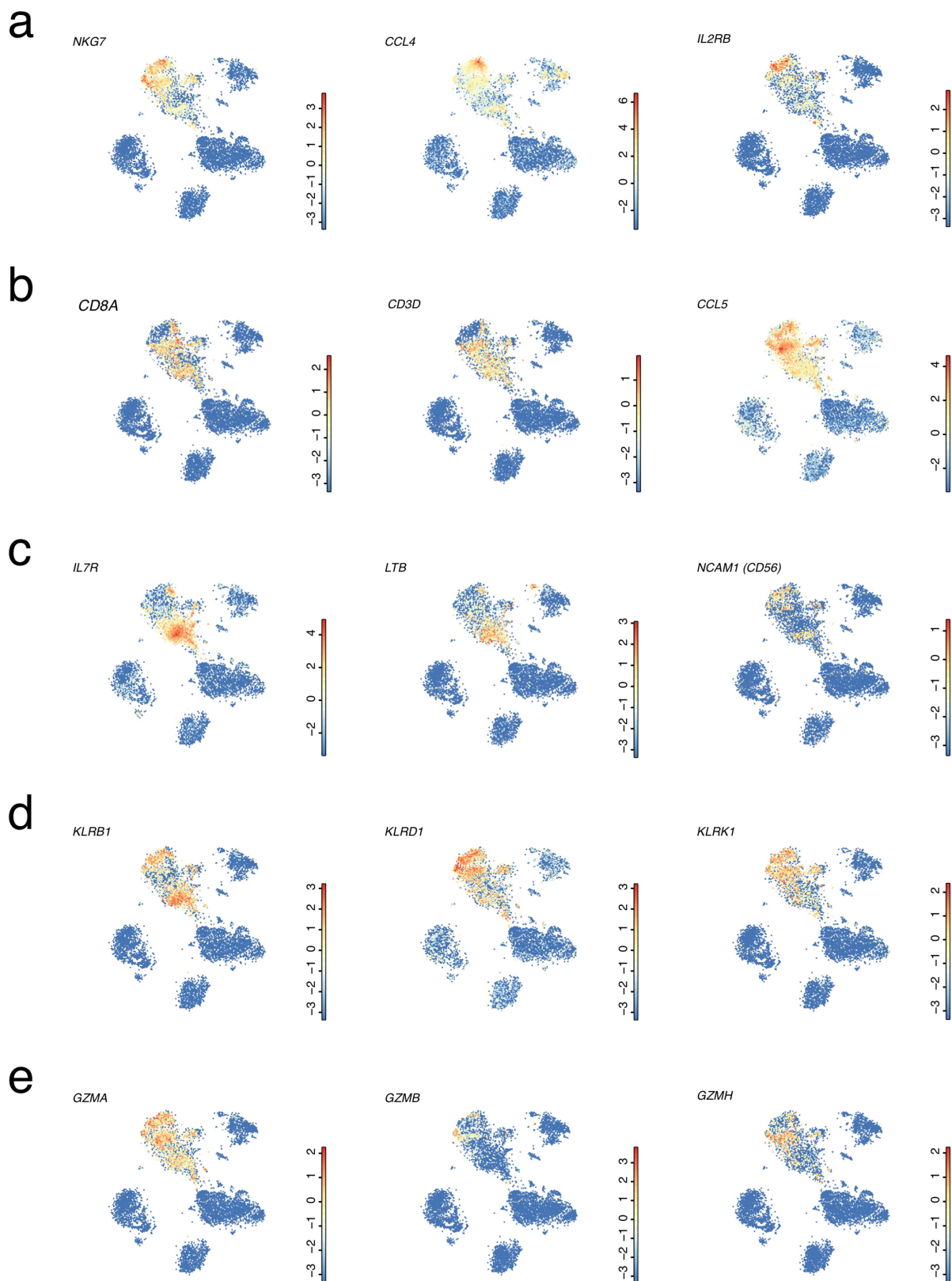
**Extended Data Fig. 4 | The human liver contains different Kupffer cell populations. a,** Expression *t*-SNE maps of marker genes for Kupffer cell subtypes. The colour bar indicates log<sub>2</sub> normalized expression ( $n = 10,372$  cells). **b,** Major pathways upregulated in the  $CD1C^{+}$  antigen-presenting ( $n = 12$  genes) and  $LILRB5^{+}$  metabolic/immunoregulatory ( $n = 35$  genes) Kupffer cell subsets as revealed by Reactome pathway analysis. The number of genes in each pathway is shown on the *x* axis.

*P* values were calculated using a hypergeometric model and adjusted using the Benjamini–Hochberg method. **c,** Expression heat map of genes upregulated in Kupffer cell clusters (Benjamini–Hochberg corrected  $P < 0.01$ , see Methods). For each cluster, the top ten upregulated genes were extracted and expression of the joint set is shown in the heat map across all Kupffer cell clusters. Genes were ordered by hierarchical clustering.



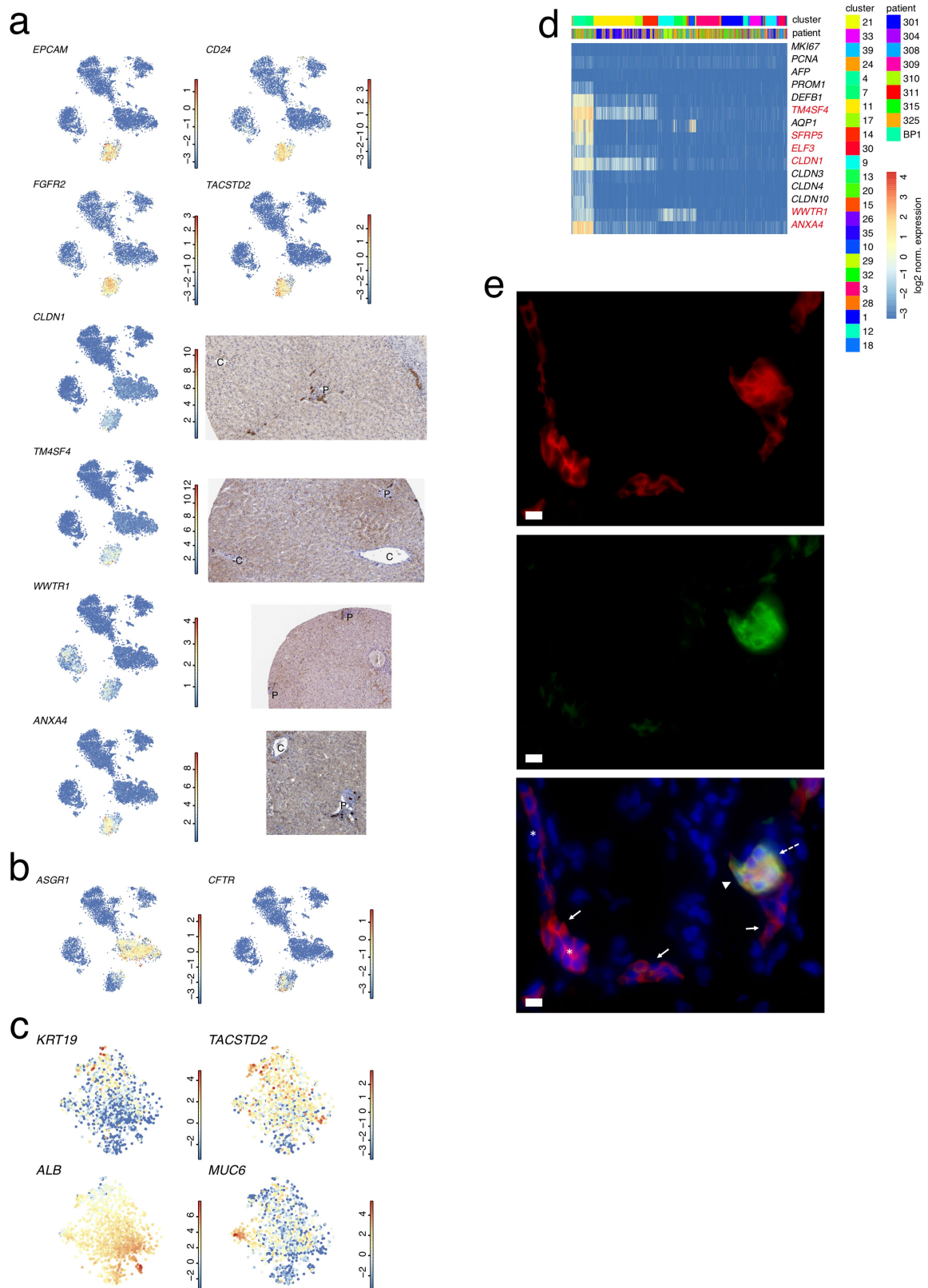


**Extended Data Fig. 5 | The human liver contains different B cell populations.** Expression *t*-SNE maps of the markers for the B cell subtypes. The colour bars indicate log<sub>2</sub> normalized expression ( $n = 10,372$  cells).



**Extended Data Fig. 6 | Heterogeneity of NK and NKT cells in the human liver.** **a–c**, Expression *t*-SNE maps of inferred markers of cluster 5 (**a**), cluster 1 (**b**) and cluster 3 (**c**). Cluster 5 comprises mainly  $CD56^+CD8A^-$  NK cells, some of which show upregulated *CCL4*. Cluster 1 comprises  $CD56^-CD8A^+$  NKT cells, which show upregulated *CCL5*. Cluster 3 consists of both  $CD56^+$  and  $CD56^-CD8A^+$  NKT cells. Clusters 1 and 3

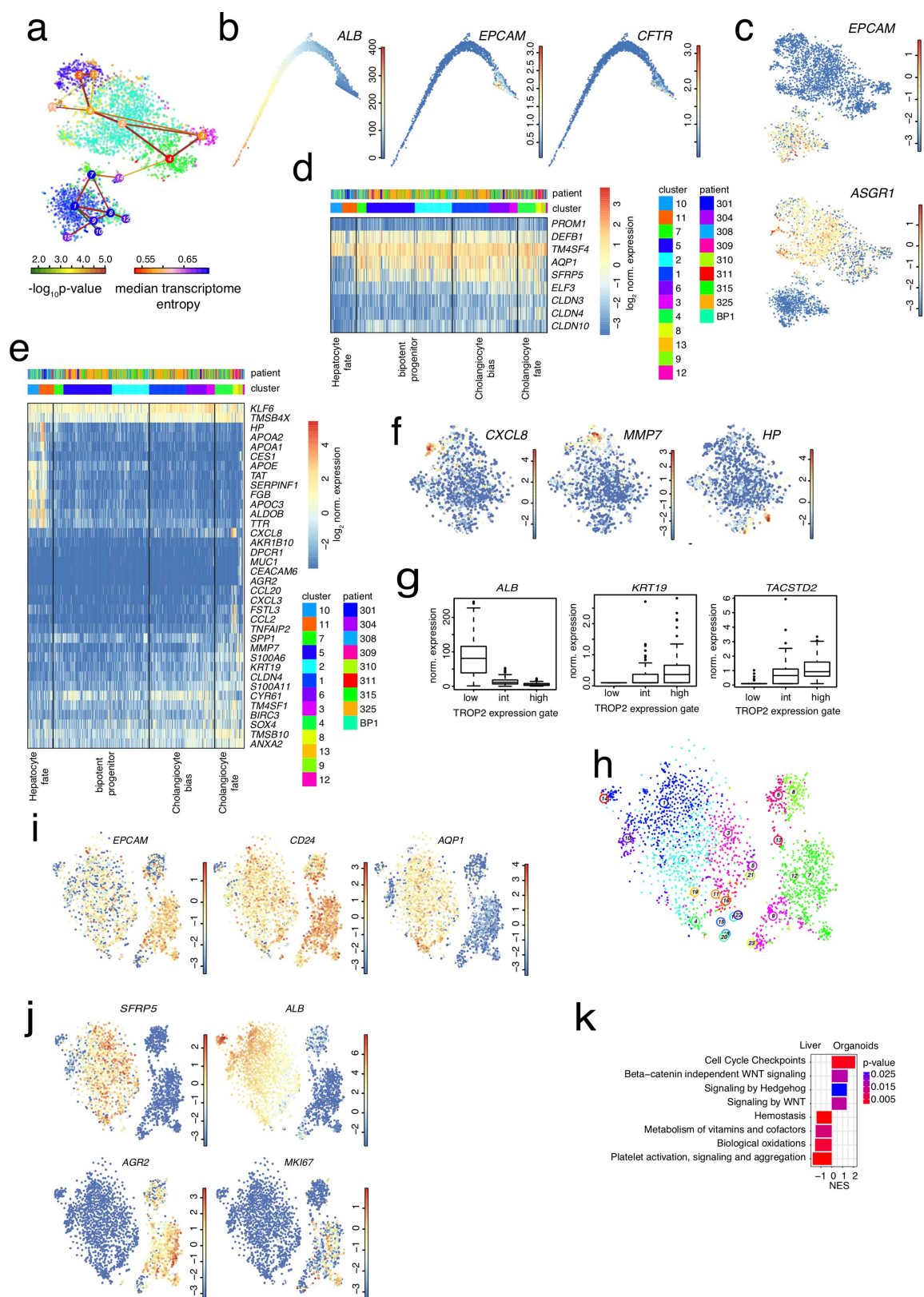
express T cell receptor components exemplified by *CD3D* co-receptor expression. **d**, Differential expression of killer cell lectin-like receptor genes across the different populations shown in **a–c**. **e**, Differential expression of granzyme genes across the different populations shown in **a–c**. Colour bars indicate log<sub>2</sub> normalized expression. **a–e**,  $n = 10,372$  cells.



Extended Data Fig. 7 | See next page for caption.

**Extended Data Fig. 7 | scRNA-seq identifies marker genes expressed by EPCAM<sup>+</sup> cells.** **a**, Expression *t*-SNE maps (left) for *EPCAM*, *CD24*, *FGFR2*, *TACSTD2*, *CLDN1*, *TM4SF4*, *WWTR1* and *ANXA4* ( $n = 10,372$  cells) and immunohistochemistry from the Human Protein Atlas (right) for *CLDN1*, *TM4SF4*, *WWTR1*, and *ANXA4*. **b**, Expression *t*-SNE maps for *ASGR1* and *CFTR* ( $n = 10,372$  cells). **c**, *t*-SNE maps showing expression of *KRT19*, *ALB*, *TACSTD2* and *MUC6* in the *EPCAM*<sup>+</sup> compartment ( $n = 1,087$  cells). **a–c**, Colour bars indicate log<sub>2</sub> normalized expression. **d**, Expression heat map of proliferation marker genes (*MKI67*, *PCNA*), *AFP*, and identified markers of the *EPCAM*<sup>+</sup> compartment.

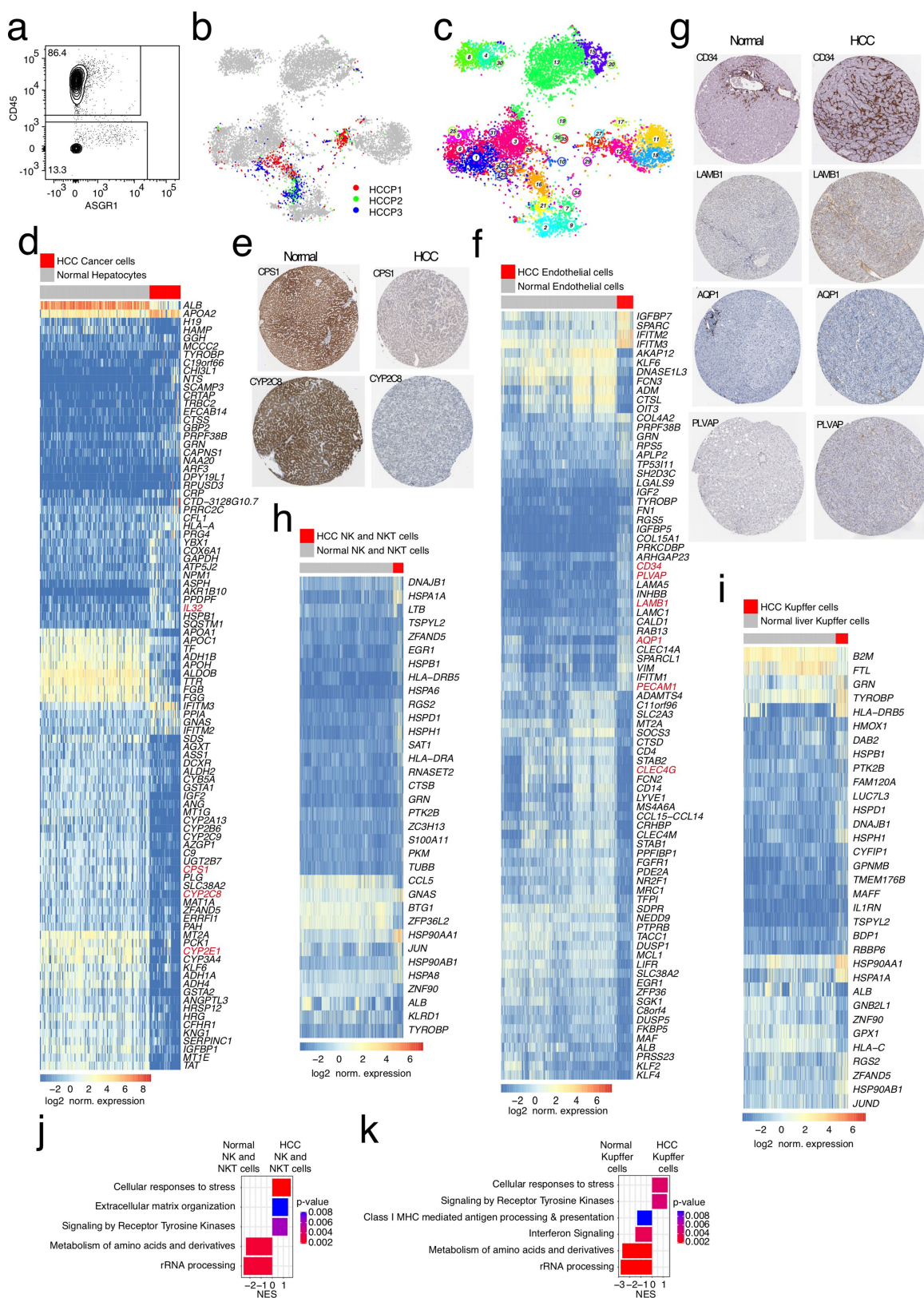
Genes highlighted in red correspond to newly identified markers of the *EPCAM*<sup>+</sup> compartment. The heat map comprises all clusters to show the specificity of the markers for the progenitor compartment. The expression analysis confirms the absence of proliferating and *AFP*<sup>+</sup> cells. **e**, Immunofluorescence labelling of *EPCAM* and *KRT19* on human liver tissue. *EPCAM*<sup>+</sup>*KRT19*<sup>low/-</sup> cells (solid arrow) in the canals of Hering (asterisk) and *EPCAM*<sup>+</sup>*KRT19*<sup>+</sup> cells (broken arrow) in the bile duct (arrowhead) are indicated. Nuclei are stained with DAPI. Scale bar, 10  $\mu$ m ( $n = 3$  independent experiments).





**Extended Data Fig. 8 | The EPCAM<sup>+</sup> compartment segregates into different major subpopulations.** **a**, Separate RaceID3 and StemID2 analyses of the EPCAM<sup>+</sup> and hepatocyte populations reveal a lineage tree connecting EPCAM<sup>+</sup> cells to mature hepatocytes via an EPCAM<sup>+</sup> hepatocyte progenitor cluster (part of the EPCAM<sup>+</sup> population in Fig. 3b). Shown are links with StemID2  $P < 0.05$ . The node colour highlights transcriptome entropy. **b**, Two-dimensional diffusion map representation of the population shown in **a**, highlighting expression of the hepatocyte marker *ALB* (left), *EPCAM* (centre), and the mature cholangiocyte marker *CFTR* (right). The maps suggest continuous transitions from the EPCAM<sup>+</sup> compartment towards hepatocytes and mature cholangiocytes. **c**, Expression *t*-SNE map of *EPCAM* (top) and the hepatocyte marker *ASGR1* (bottom) for the population shown in **a**. Colour bars indicate log<sub>2</sub> normalized expression. **b**, **c**,  $n = 3,877$  cells. **d**, Expression heat map of de novo identified markers of the EPCAM<sup>+</sup> compartment, highlighting the expression distribution within clusters of this population only (Fig. 3). **e**, Expression heat map of all genes that were differentially expressed in the

more mature clusters, belonging to the groups denoted as ‘hepatocyte fate’ and ‘cholangiocyte fate’. For each of these clusters, the top ten upregulated genes (Benjamini–Hochberg corrected  $P < 0.01$ ) were selected, and the joint set of these genes is shown in the figure. **f**, Expression *t*-SNE maps of *CXCL8*, *MMP7* and *HP*. Colour bars indicate log<sub>2</sub> normalized expression. **d–f**,  $n = 1,087$  cells. **g**, Normalized expression counts of *ALB*, *KRT19* and *TACSTD2* in cells sequenced from the gates in Fig. 4a ( $n = 293$  cells). Centre line, mean; boxes, interquartile range; whiskers, 5% and 95% quantiles; data points, outliers. **h**, *t*-SNE map of RaceID3 clusters for organoid cells and EPCAM<sup>+</sup> cells from patients (Fig. 3), including cells sorted from the gates in **a**. **i**, Expression *t*-SNE maps of *EPCAM*, *CD24* and *AQP1* in organoid cells and EPCAM<sup>+</sup> cells from patients. **j**, Expression *t*-SNE maps of *SFRP5*, *ALB*, *AGR2* and *MKI67*. Colour bars indicate log<sub>2</sub> normalized expression. **h–j**,  $n = 2,870$  cells. **k**, GSEA of genes that are differentially expressed between organoid and EPCAM<sup>+</sup> liver cells from patients (Benjamini–Hochberg corrected  $P < 0.01$ ,  $n = 11,610$  genes; see Methods).

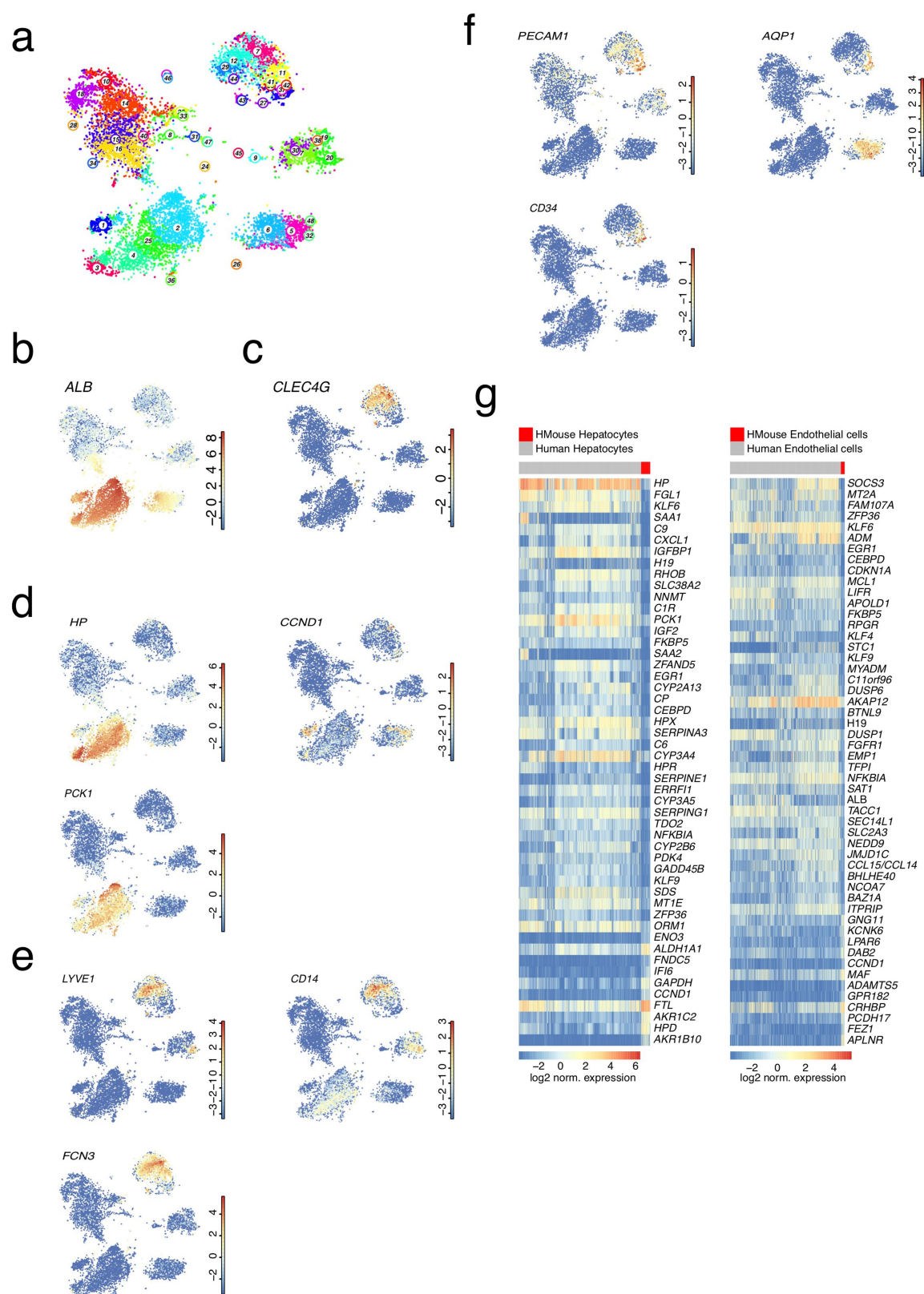


Extended Data Fig. 9 | See next page for caption.

**Extended Data Fig. 9 | Cell types from patient-derived HCC exhibit perturbed gene expression signatures.** **a**, FACS plot of CD45 and ASGR1 staining on cells from HCC samples ( $n = 3$  independent experiments). **b**, Symbol  $t$ -SNE map showing the IDs of HCC patients ( $n = 11,654$  cells). **c**,  $t$ -SNE map showing RaceID3 clusters for normal liver cells co-analysed with cells from HCC tissues ( $n = 3$  patients). **d**, Expression heat map of genes that are differentially expressed between cancer cells from HCC and normal hepatocytes (Benjamini–Hochberg corrected  $P < 0.05$  and  $\log_2$  fold change  $> 1.6$ ;  $n = 256$  cells; see Methods). Genes highlighted in red correspond to differentially expressed genes validated by immunohistochemistry. **e**, Immunostaining of CPS1 and CYP2C8 in normal liver and HCC tissues from the Human Protein Atlas. **f**, Expression heat map of genes that are differentially expressed between endothelial cells from HCC and normal endothelial cells from MaVEC and LSEC

clusters. Benjamini–Hochberg corrected  $P < 0.05$ ;  $\log_2$  fold change  $> 1.5$ ;  $n = 1,936$  cells; see Methods). Genes highlighted in red correspond to differentially expressed genes validated by immunohistochemistry. **g**, Immunostaining of CD34, LAMB1, AQP1 and PLVAP in normal liver and HCC tissues from the Human Protein Atlas. **h**, Heat map of genes that are differentially expressed between normal and HCC-resident NK and NKT cells (Benjamini–Hochberg corrected  $P < 0.05$ ;  $n = 2,754$  cells; see Methods). **i**, Heat map of genes that are differentially expressed between normal and HCC-resident Kupffer cells (Benjamini–Hochberg corrected  $P < 0.05$ ;  $n = 991$  cells; see Methods). **j**, GSEA of genes that are differentially expressed between normal and HCC-resident NK and NKT cells ( $n = 15,442$  genes). **k**, GSEA of genes that are differentially expressed between normal and HCC-resident Kupffer cells ( $n = 15,442$  genes).





**Extended Data Fig. 10 | Transplanted human liver cells in a humanized mouse model exhibit a distinct gene signature compared to cells within the human liver.** **a**, *t*-SNE map of RaceID3 clusters of liver cells from patients co-analysed with cells from the humanized mouse liver model. **b**, Expression *t*-SNE maps of the hepatocyte marker gene *ALB*. **c**, Expression *t*-SNE maps of the endothelial marker *CLEC4G*. **d**, Expression *t*-SNE maps of *HP*, *PCK1* and *CCND1*. **e**, Expression *t*-SNE maps of the liver endothelial cell zonated genes *LYVE1*, *FCN3* and *CD14*. **f**, Expression

*t*-SNE maps of *PECAM1*, *CD34* and *AQP1*. **a–f**, Colour bars indicate  $\log_2$  normalized expression.  $n = 10,683$  cells. **g**, Heat maps of genes that are differentially expressed between hepatocytes ( $n = 3,175$  cells) and endothelial cells ( $n = 1,710$  cells) from patients (human hepatocytes and human endothelial cells) and from the humanized mouse model (HMouse hepatocytes and HMouse endothelial cells). Benjamini-Hochberg corrected  $P < 0.05$ ; see Methods.

## Supplementary Note 1

### Single-cell RNA-seq reveals heterogeneity within the endothelial cell compartment

Differential gene expression analysis of the RaceID3 clusters within the endothelial cell compartment (Fig. 1b-d) uncovered two major populations of endothelial cells in the human liver, the macrovascular endothelial cells (MaVECs) and the liver sinusoidal endothelial cells (LSECs). In the normal liver, LSECs line the sinusoids of the liver lobule and are *CLEC4G<sup>+</sup>PECAMI<sup>low</sup>*, while MaVECs line the hepatic arteries and veins and are *CD34<sup>+</sup>PECAMI<sup>high</sup>* cells<sup>45,46</sup> (Extended Data Fig. 2). Interestingly, we find that CD34 also stains some LSECs in the periportal zone. LSECs have different functions and morphology from venous and arterial endothelial cells, i.e. they are fenestrated and have filtering and scavenging roles<sup>47-49</sup>. Within the *CD34<sup>+</sup>* compartment we identified an *AQP1<sup>high</sup>* population as well as a *CPE<sup>+</sup>* (clusters 29 and 32) and a *CPE<sup>-</sup>* (cluster 10) sub-population (Extended Data Fig. 2c, d). Antibody stainings extracted from the Human Protein Atlas revealed strong AQP1 expression in endothelial cells of the portal tract as well as in bile duct cells (Extended Data Fig. 2f). In addition, we identified several other novel sub-types, including a *CCL21<sup>+</sup>* population, which expresses angiogenesis-associated genes like the netrin receptor *UNC5B<sup>50,51</sup>* and *TFF3<sup>52</sup>* (cluster 35) (Extended Data Fig. 2c). Another population highly expresses *H19*, a non-coding RNA involved in early development<sup>53</sup>.

## Supplementary Note 2

### Co-zonation of gene expression indicates functional cooperation of hepatocytes and endothelial cells

Consistent with the observation of shared pathways expressed in hepatocytes and endothelial cells residing within the same zones (Fig. 2), we observed correlated expression of many genes across zones. For instance, we found *CD14* to be co-expressed by midzonal hepatocytes and LSECs (Extended Data Fig. 4a). We observed an expression peak of the aminopeptidase *ANPEP* in LSECs of the periportal zone (module 20). Immunostaining of *ANPEP* indicated strong expression in the bile canaliculi of hepatocytes and specific expression in LSECs closer to the periportal zone (Fig. 2b). Interestingly, pathway enrichment analysis of module 20 revealed *ANPEP*-containing pathways for peptide hormone metabolism and for the synthesis and secretion of incretins, which are metabolic hormones involved in stimulating a decrease in blood glucose levels by various mechanisms including

the augmentation of insulin secretion<sup>54</sup>; this results in the stimulation of glycogen synthesis, which peaks in periportal hepatocytes. These observations together with the pathways shared between LSECs and hepatocytes, support the idea that LSECs and hepatocytes, as a result of their zonation, are transcriptomically heterogeneous and may co-operate to regulate and carry out particular functions in a zone-specific manner.

### Supplementary Note 3

#### Comparison between mouse and human reveals limited evolutionary conservation of gene expression zonation

Since genome-wide zonation of hepatocytes and endothelial cells has been characterized recently in mouse, we investigated conservation of zonation profiles in human hepatocytes and endothelial cells (Methods). For endothelial cells we included both human MaVECs and LSECs to be consistent with the published mouse data. A correlation analysis of zonation patterns revealed only a limited degree of evolutionary conservation, comparable between endothelial cells and hepatocytes: only 68% (60%) of genes with significant zonation patterns exhibited a positive correlation of zonation profiles for hepatocytes (endothelial cells) (Extended Data Fig. 3k,l and Supplementary Data Tables 6,7), suggesting widespread evolutionary changes in zonation patterns. Changes were observed across all zones without enrichment of particular pathways. Reassuringly, profiles of known hepatocyte-zonated genes such as *ALB*, *HP*, *PCK1*, and *CYP2A1* were conserved (Supplementary Data Tables 6,7).

### Supplementary Note 4

#### The human liver hosts distinct subpopulations of Kupffer cells

Analysis of the *CD163*<sup>+</sup>*VSIG4*<sup>+</sup> Kupffer cell compartment<sup>55</sup> revealed two main subsets: a *LILRB5*<sup>+</sup>*CD5L*<sup>+</sup>*MARCO*<sup>+</sup>*HMOX1*<sup>high</sup> subset (cluster 6) and a *CD1C*<sup>+</sup>*FCERIA*<sup>+</sup> subset (cluster 2) (Extended Data Fig. 4a, c). Differential gene expression and pathway enrichment analysis indicated an immunoregulatory and metabolic gene signature for *LILRB5*<sup>+</sup> Kupffer cells while *CD1C*<sup>+</sup> Kupffer cells exhibit a signature with higher expression of genes involved in MHC Class II antigen presentation such as *HLA-DRA*. These observations are in agreement with recent single-cell based analysis of the human liver<sup>10</sup>. Furthermore, *LILRB5* belongs to the LILR subfamily B receptors, which bind to MHC class I molecules on antigen presenting cells and inhibits stimulation of an immune response<sup>56</sup>. Moreover, we detected higher *CD163*

expression in the *LILRB5*<sup>+</sup> subset compared to the *CD1C*<sup>+</sup> subset, which is consistent with the high expression of *CD163* in M2 macrophages<sup>57</sup>. Interestingly, we observed that *CD14*, *LYVE1* and *MRC1* are co-expressed by Kupffer cell and midzonal LSEC subsets (Extended Data Fig. 4a). Furthermore, pathway enrichment analysis indicated that the *LILRB5*<sup>+</sup>*HMOX*<sup>high</sup> Kupffer cell subset shares pathways such as binding and uptake of ligands by scavenger receptors with subsets of hepatocytes and endothelial cells (Extended Data Fig. 4b). Additional cell types co-clustering with these Kupffer cell subtypes comprise *CD163*<sup>+</sup>*VCAN*<sup>+</sup> cells (cluster 23), *CD163*<sup>+</sup>*AREG*<sup>+</sup> cells (cluster 25), and a small *CD163*<sup>+</sup> cluster expressing cytochrome P450 genes (cluster 31). The latter population could potentially be explained by doublets consisting of Kupffer cells or macrophages and hepatocytes, or hepatocytes being phagocytosed (Extended Data Fig. 4c).

## Supplementary Note 5

### Sinusoidal endothelial cells isolated from HCC tissue upregulate PLVAP

HCC sinusoidal endothelial cells upregulate PLVAP and LAMB1 compared to normal LSECs and MaVECs (Extended Data Fig. 9f,g). PLVAP is implicated in angiogenesis<sup>58</sup> and involved in the formation of the diaphragms of fenestrae on endothelial cells. Hence, it acts as a physical sieve or barrier to reduce vascular permeability and controls the entry of lymphocytes and soluble antigens. Related to this observation, we find that immune cell populations from the tumor, comprising Kupffer cells, NKT and NK cells upregulate stress response genes and pathways (Extended Data Fig. 9h-k).

These observations suggest that endothelial cells in the sinusoids of HCC tumors may undergo a transformation into a more macrovascular-like endothelial cell phenotype, promoting angiogenesis while becoming less permeable to immune cells as a result of PLVAP overexpression. It is conceivable, that, on the one hand, such an effect could result in lower immune cell infiltration into the tumor while, on the other, it may lead to the trapping of immune cells, which are activated by cancer cells (for example via IL32 expression). The latter could thus be capable of evading the immune system while promoting inflammation, ultimately leading to cytokine mediated feeding of the cancer by immune cells.

## References

45. Pusztaszeri, M. P., Seelentag, W. & Bosman, F. T. Immunohistochemical expression of endothelial markers CD31, CD34, von Willebrand factor, and Fli-1 in normal human tissues. *J Histochem Cytochem* **54**, 385-395, doi:10.1369/jhc.4A6514.2005 (2006).
46. Lalor, P. F., Lai, W. K., Curbishley, S. M., Shetty, S. & Adams, D. H. Human hepatic sinusoidal endothelial cells can be distinguished by expression of phenotypic markers related to their specialised functions in vivo. *World J Gastroenterol* **12**, 5429-5439 (2006).
47. Braet, F. & Wisse, E. Structural and functional aspects of liver sinusoidal endothelial cell fenestrae: a review. *Comp Hepatol* **1**, 1 (2002).
48. Fraser, R., Dobbs, B. R. & Rogers, G. W. Lipoproteins and the liver sieve: the role of the fenestrated sinusoidal endothelium in lipoprotein metabolism, atherosclerosis, and cirrhosis. *Hepatology* **21**, 863-874 (1995).
49. Wisse, E., De Zanger, R. B., Jacobs, R. & McCuskey, R. S. Scanning electron microscope observations on the structure of portal veins, sinusoids and central veins in rat liver. *Scan Electron Microsc*, 1441-1452 (1983).
50. Lu, X. *et al.* The netrin receptor UNC5B mediates guidance events controlling morphogenesis of the vascular system. *Nature* **432**, 179-186, doi:10.1038/nature03080 (2004).
51. Navankasattusas, S. *et al.* The netrin receptor UNC5B promotes angiogenesis in specific vascular beds. *Development* **135**, 659-667, doi:10.1242/dev.013623 (2008).
52. Ahmed, A. R. H., Griffiths, A. B., Tilby, M. T., Westley, B. R. & May, F. E. B. TFF3 is a normal breast epithelial protein and is associated with differentiated phenotype in early breast cancer but predisposes to invasion and metastasis in advanced disease. *Am J Pathol* **180**, 904-916 (2012).
53. Jinno, Y. *et al.* Establishment of functional imprinting of the H19 gene in human developing placentae. *Nat Genet* **10**, 318-324 (1995).
54. Kim, W. & Egan, J. M. The role of incretins in glucose homeostasis and diabetes treatment. *Pharmacol Rev* **60**, 470-512, doi:10.1124/pr.108.000604 (2008).
55. Mowat, A. M., Scott, C. L. & Bain, C. C. Barrier-tissue macrophages: functional adaptation to environmental challenges. *Nat Med* **23**, 1258-1270, doi:10.1038/nm.4430 (2017).
56. Borges, L., Hsu, M. L., Fanger, N., Kubin, M. & Cosman, D. A family of human lymphoid and myeloid Ig-like receptors, some of which bind to MHC class I molecules. *J Immunol* **159**, 5192-5196 (1997).
57. Gonzalez-Dominguez, E. *et al.* CD163L1 and CLEC5A discriminate subsets of human resident and inflammatory macrophages in vivo. *J Leukoc Biol* **98**, 453-466, doi:10.1189/jlb.3HI1114-531R (2015).
58. Carson-Walter, E. B. *et al.* Plasmalemmal vesicle associated protein-1 is a novel marker implicated in brain tumor angiogenesis. *Clin Cancer Res* **11**, 7643-7650, doi:10.1158/1078-0432.Ccr-05-1099 (2005).

## **3.2 VIRAL COMPARTMENTALIZATION, CANCER HETEROGENEITY IN HBV-INDUCED HEPATOCELLULAR CARCINOMA**

Chronic HBV infection is a major cause of HCC and molecular mechanisms of virus-host interactions and hepatocarcinogenesis are still partially understood. In this study, we analyzed at single-cell level, HCC cells from a patient with a low HBV load and investigated HBV-host cell interactions and intratumor heterogeneity. We applied the Smart-Seq2 protocol that allows deep full transcript sequencing including the analysis of putative viral integration sites of the host genome. Computational analyses of gene expression revealed a marked heterogeneity of the HCC and the tumor microenvironment. Analyses of virus-induced host responses identified previously undiscovered pathways mediating viral carcinogenesis, including a marked correlation of HBV load and the oncogene SERTAD2. Finally, mapping of fused HBV-host cell transcripts unraveled integration sites in individual cancer cells. Importantly, scRNA-Seq unraveled heterogeneity and compartmentalization of both virus and cancer. High HBV levels were associated with an induction of genes involved in bile acid- and fatty acid metabolism as in more differentiated cancer cells. Furthermore, the impact of HBV on gene expression in cancer cells harboring HBV transcripts was highly similar to the gene expression profile of HBV-infected primary human hepatocytes. The perturbation of gene expression mediating carcinogenesis in cells with low viral RNA levels highlights the importance of curing HBV chronic infection to eliminate HCC risk. The marked tumor heterogeneity suggests that combination therapies targeting multiple drivers are required for HCC chemotherapeutic approaches. The paper is in revision in *Journal of Clinical Investigation* (2019).

# **Viral compartmentalization and cancer heterogeneity in HBV-induced hepatocellular carcinoma**

Frank Jühling<sup>1,\$</sup>, Eloi R. Verrier<sup>1,\$</sup>, Antonio Saviano<sup>1,\$</sup>, Houssein El Saghire<sup>1</sup>, Laura Heydmann<sup>1</sup>,  
Patrick Pessaux<sup>1,2</sup>, Nathalie Pochet<sup>3</sup>, Catherine Schuster<sup>1,2</sup>, Thomas F. Baumert<sup>1,2,4,\*</sup>

<sup>\$</sup>Co-first authors

**Affiliations:** <sup>1</sup>Université de Strasbourg, Inserm, Institut de Recherche sur les Maladies Virales et Hépatiques UMR\_S1110, F-67000 Strasbourg, France; <sup>2</sup>Institut Hospitalo-Universitaire, Pôle Hépatodigestif, Nouvel Hôpital Civil, F-67000 Strasbourg, France; <sup>3</sup>Ann Romney Center for Neurologic Diseases, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA, Cell Circuits Program, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA; <sup>4</sup>Institut Universitaire de France (IUF), Paris, France.

**\*Corresponding author:** Prof. Thomas F. Baumert, MD; Inserm U1110, Institut de Recherche sur les Maladies Virales et Hépatiques, 3 Rue Koeberlé, 67000 Strasbourg, France; Phone: +33 3 68 85 37 03, Fax: +33 3 68 85 37 24, e-mail: [thomas.baumert@unistra.fr](mailto:thomas.baumert@unistra.fr)

**Conflict of interest statement:** The authors have declared that no conflict of interest exists.

**Total word count:** 3986

**ABSTRACT** (184/200 words)

Chronic hepatitis B virus (HBV) infection is a major cause of hepatocellular carcinoma (HCC) world-wide. Molecular mechanisms of virus-host interactions and hepatocarcinogenesis are still partially understood. Here, we report the case of a HCC despite low HBV load and investigated HBV-host cell interactions in this patient-derived HCC using single-cell sequencing. Applying the Smart-Seq2 protocol allowed for deep full transcript sequencing including the analysis of putative viral integration sites of the host genome. Computational analyses of gene expression revealed a marked heterogeneity of the HCC and the tumor microenvironment. Analyses of virus-induced host responses identified previously undiscovered pathways mediating viral carcinogenesis, including a marked correlation of HBV load and the oncogene *SERTAD2*. Finally, mapping of fused HBV-host cell transcripts unraveled integration sites in individual cancer cells. Collectively, single-cell RNA-Seq unravels a heterogeneity and compartmentalization of both, virus and cancer. The perturbation of gene expression mediating carcinogenesis in cells with low viral RNA levels highlights the importance of curing HBV chronic infection to eliminate HCC risk. The marked tumor heterogeneity suggests that combination therapies targeting multiple drivers are required for HCC chemotherapeutic approaches.



## INTRODUCTION

Hepatitis B virus (HBV) infection is a major cause of chronic liver disease and hepatocellular carcinoma (HCC)<sup>1</sup>. An estimated 2 billion people have evidence of exposure to HBV and more than 250 million people are chronically infected with HBV worldwide. HBV infected patients have an approximately 100-fold increased risk for HCC compared to uninfected patients<sup>2</sup>. HCC is the second leading and fastest rising cause of cancer death worldwide<sup>3</sup>. Each year, close to 600,000 people are newly diagnosed with HCC. The future significance and impact of the disease is not only illustrated by its rising incidence over the last two decades, but also by its unchanged high mortality<sup>3,4</sup>. Thus, the burden of established, incurable HBV-induced liver disease represents a major challenge to public health and efficient treatment strategies to cure chronic hepatitis B (CHB) are urgently needed<sup>4</sup>. The pathogenesis of HBV-induced HCC is multifactorial. CHB induces HCC through direct and indirect mechanisms (reviewed in<sup>5</sup>). First, the viral DNA has been shown to integrate into the host genome inducing both genomic instability and direct insertional mutagenesis of diverse cancer-related genes<sup>5</sup>. Compared with tumors associated with other risk factors, HBV-related tumors have a higher rate of chromosomal alterations including p53 inactivation by mutations. Moreover, epigenetic changes targeting the expression of tumor suppressor genes have been shown to occur early in the development of HCC<sup>5</sup>. Second, HBV proteins such as the viral regulatory protein HBx or altered versions of the preS/S envelope proteins have been shown to modulate cell transcription, resulting in alteration in host cell proliferation and sensitizing the hepatocytes to carcinogenic factors<sup>5</sup>. HBV-related HCCs can also arise in non-cirrhotic livers, supporting the notion that HBV plays a direct role in liver transformation by triggering both common and etiology-specific oncogenic pathways in addition to stimulating the host immune response and driving liver chronic necro-inflammation<sup>5</sup>.

Single-cell RNA-Seq is a high-resolution technique allowing transcriptome-wide analyses of individual cells and represents a precious tool to study heterogeneous tissues including cancer<sup>6</sup>. Tumors are characterized by multiple neoplastic sub-clones as well as non-neoplastic cells constituting the tumor microenvironment. Moreover, single-cell RNA-Seq enables to distinguish cells with different levels of HBV infection, cells exposed to the virus but not infected as well as noninfected cells. Aiming to

investigate HBV-host interactions and viral carcinogenesis in individual cells of HCC, we performed a single-cell RNA-Seq analysis of a CHB-related HCC. Applying Smart-Seq2 instead of another single cell protocol optimized for high throughput of cells to be sequenced, we were able to deeply sequence full RNA transcripts in a sufficient number of cells. Thereby, only Smart-Seq2 allowed us to study integration sites with genetic material originating from both, human and virus, on the same RNA transcript.

## COMBINED RESULTS AND DISCUSSION

**Clinical case.** Here, we study the case of a patient who developed HCC despite of a low HBV load (Table S1). The patient was an overweight 61-year-old man with a history of type-2 diabetes treated by sulfonylureas and of streptococcal endocarditis of the aortic valve that occurred 2 years earlier. His family history was positive for HCC and hemochromatosis. He presented to the emergency department with acute abdominal epigastric pain. Blood tests revealed mild anemia without leukocytosis and normal liver/kidney chemistry. A computed tomography (CT) scan showed an intrahepatic lesion of 5 cm associated with hemoperitoneum without active bleeding. The patient was therefore admitted to the liver surgery unit. Subsequent magnetic resonance imaging (MRI) revealed a liver tumor in the liver segment III displaying typical radiological features of HCC (Figure 1A-C). Virological analyses revealed HBeAg-negative chronic HBV infection<sup>7</sup>. Serological tests for hepatitis C virus and HIV were negative. Levels of alpha-fetoprotein (AFP), CA 19-9, CEA and PSA were within the normal range (Table S1). Gastroscopy and colonoscopy were negative for neoplastic lesions and portal hypertension. A segmental liver resection was performed, the histopathological examination of the resected tissue revealed a well to moderately differentiated HCC displaying both a trabecular and a pseudoglandular pattern (Figure 1D-F). The surrounding liver tissue showed portal fibrosis with some incomplete septa, iron overload and macrovesicular steatosis without lobular inflammation nor hepatocyte ballooning. Hereditary hemochromatosis was ruled out by the detection of normal ferritin levels and negative HFE test. Following surgical recovery, an antiviral therapy with entecavir was started. 18 months following treatment start HBV-DNA were undetectable and HBsAg was lost. No HCC recurrence was detected during a 2 years follow-up.

**Heterogeneity of HCC and infiltrating nonparenchymal cells (NPCs).** Single cells were isolated from the resected tissue and gene expression was quantified by RNA-seq. Primary human hepatocytes isolated (PHH) from a healthy donor were used as a reference. The intratumoral heterogeneity was assessed using clustering and marker gene expression analyses. Gene expressions were highlighted

on T-distributed Stochastic Neighbor Embedding (t-SNE) maps (Figure 2A) enabling the separation and deep characterization of parenchymal and non-parenchymal cells using cell-specific markers. Cancer cells were identified by expression of *GPC3*, a well-documented HCC marker<sup>8</sup>, as well as osteopontin (*SPP1*) expression, not detectable in healthy hepatocytes<sup>9</sup>. In addition to hepatocyte-derived cancer cells, the following non-parenchymal cells were identified: Macrophages (expressing *CD14* and *ITGAM*), endothelial cells (expressing *PECAM* and *KDR*, but not *ITGAM*), and one antigen-presenting Kupffer cell (expressing *CLEC4F*, and *CD1-A/B/C/D*) (Figure 2B-C), as determined according the Human Protein Atlas<sup>10</sup> ([www.proteinatlas.org](http://www.proteinatlas.org)).

Single cell-specific clustering algorithms<sup>11</sup> revealed a marked heterogeneity of HCC cells as visualized on a diffusion map highlighting the clusters of different cancer cell populations and their branching (Figure 3A). Main branches of cells comprising clusters 1 and 2, and with cluster 4 connecting them, show very distinct marker gene expressions (Figure 3B, Table S2). In contrast, gene set enrichment analysis (GSEA) calculated cell-wise vs. control PHHs (Figure 3C) revealed that nearly all cells express a *CTNNB1*-positive HCC subclass-like profile which corresponds well with the patient's clinical data (Figure S1). Tumors overexpressing *CTNNB1* are grouped in one specific molecular class<sup>12,13</sup>. While the intratumoral and intertumoral heterogeneity of *CTNNB1* mutations in HCC were already reported<sup>14,15</sup>, no study unraveled the transcriptomic intratumoral heterogeneity of *CTNNB1*+ HCC at the single cell level. In contrast, we found cell populations in both branches differ in the expression of cancer stem cells genes as well as in genes involved in the DNA repair and in the metastatic process (Figure 3C).

Interestingly, cells of clusters 1 and 2 show similar features such as up-regulated genes involved in metastasis (Figure S1). However, they exhibit very distinct features overall. Cells in cluster 1 express a different profile, i.e., the “unannotated” subclass in Chiang et al.<sup>16</sup> and simultaneously features of high proliferating HCCs of high proliferation, clusters 2 and 4 show up as differentiated HCC cells (subclass S3 according to<sup>13</sup>) (Figure 3C). Cluster 3 shows cells of both subclasses and seems to be in a transition state between both. Cluster 1 is consistently enriched for cancer stem cell genes to be up-regulated, and DNA repair down-regulated, which is consistent with previous findings of higher chromosomal

instability in subclasses with high proliferation<sup>16</sup> (Figure 3C). This suggests a cluster of high heterogeneity and fast evolving cells in cluster 1, while converse gene profiles are expressed in cluster 2. Moreover, growth factor activity is higher and HCC-specific growth factors *VEGFA* and *EGFR* (Figure S2) are more expressed in cluster 1 compared to the others. A similar heterogeneity was observed when analyzing the expression of gene sets associated with prognosis of HCC. Intersecting differentially expressed cluster marker genes (compared to control PHHs) with prognostic genes (identified in TCGA patient's data and listed in the Human Protein Atlas) resulted only in 10 hits listed as “prognostic, unfavorable” (corresponding to poor prognosis), and were identified only in cluster 1. Examples of such poor prognosis marker genes in cluster 1 are shown in Figure 3D (full dataset: Table S2). Some of them were already deeply analyzed for their prognostic capabilities in HCC<sup>16</sup>. *GGA3* and *ACACA*, implicated in regulating intracellular trafficking and in the fatty acid synthesis, respectively, are known to be involved in HCC pathogenesis and associated with a poor prognosis in HCC<sup>17,18</sup>. On the other hand, *CPSF7* and *SRRM2* are involved in splicing processes and were previously reported to be associated with a poor prognosis in non-hepatocellular cancers<sup>19-21</sup>. We found that *CPS7* and *SRRM2* were heterogeneously expressed among the HCC clusters and their overexpression was almost confined in the clusters displaying a metastatic gene expression profile suggesting a role in HCC prognosis and progression. Overall, this supports our approach for the identification of poor prognostic markers associated with liver disease within specific cell compartments of a heterogeneous primary HCC tumor.

**Compartmentalization of HBV RNA and virus-host interactions in single cells of the tumor.** To unravel the relationship between HBV expression and host gene expression in HCC cells, we analyzed HBV RNA levels in tumor single cells. While all HCC cells were HBV RNA positive, an intra-tumor cells comparison revealed marked differences in HBV RNA levels in individual cells (Figure 4A). Interestingly, HBV RNA loads were significantly modulated between the four cell clusters that we identified in the HCC analyses (Figure 4B), with the minimum HBV load in cluster 1. Collectively, these findings suggest a compartmentalization of HBV similarly as it has recently been suggested for HCV<sup>22</sup>. Finally, two NPCs expressing macrophage markers were positive for HBV RNA (Figure 4C). Further investigating the

characteristics of these macrophages, we detected liver-specific sequences in one macrophage (*ALB*, data not shown), suggesting that the presence of HBV RNA in the macrophage is most likely due a phagocytosis of HBV-expressing cancer cell by this macrophage and not the consequence of a productive infection.

We then calculated the correlation between HBV RNA levels and transcripts of recently identified HBV host factors at the single cell level. Host factors included HBV entry factor sodium taurocholate co-transporting polypeptide NTCP, transcription factor HNF4A, HBx-binding protein DDB1 and cccDNA DNA repair enzyme TDP2. An absent correlation between the gene expression of previously described HBV host factors (*SLC10A1*, *HNF4A*, *DDB1*, *TDP2*) (Figure S3) suggests that HBV RNA level differences were not due to different expression of these virus-dependency hepatocyte factors.

To understand the functional impact of HBV on HCC gene expression, we performed a GSEA pathway analysis using the MSigDB Hallmark collection and identified cellular pathways associated with HBV infection (Figure 4D). Notably, high HBV levels were associated with an induction of genes involved in bile acid- and fatty acid metabolism. Furthermore, the impact of HBV on gene expression in cancer cells harboring HBV transcripts was highly similar to the gene expression profile of HBV-infected PHH<sup>23</sup>. A down-regulation of E2F- and MYC targets was observed, as well as a strong induction of xenobiotic metabolism and adipogenesis (Figure 4D). This observation suggests that HBV infection induces long-term modifications of the hepatocyte transcriptomic profile, which in turn are likely to be involved in the development of liver disease and carcinogenesis. Interestingly, several cancer-related key pathways were only up-regulated in HCC-derived single cells, highlighting the specificity of the virus's impact on cancer cells. Notably, the up-regulation of genes involved in hypoxia was only associated with HBV RNA levels in HCC cells (Figure 4B). Hypoxia plays a key role in hepatocarcinogenesis and liver tumor progression, through the ability of the hypoxia inducible factor 1 alpha (HIF-1 $\alpha$ ) to target the expression of oncogenic genes such as the proliferation-specific transcription factor Forkhead box M1<sup>24</sup>. HIF-1 $\alpha$  overexpression in HCC has been correlated with worse clinical outcomes and is considered as a poor prognosis factor and molecular target for liver disease therapy<sup>24,25</sup>. To identify functionally relevant virus-host interactions for hepatocarcinogenesis, we analyzed the correlation between HBV RNA levels and

expression of individual genes, using the set of oncogenes and tumor suppressor genes from Tumor Associated Gene (TAG) database (<http://www.binfo.ncku.edu.tw/TAG/GeneFinder.php>). As shown in Figure 4E, HBV loads in single cells correlated positively with the expression of *SERTAD2*, which encodes TRIP-Br2, a regulator of fat lipolysis playing a key role in obesity and insulin resistance<sup>26</sup>. Interestingly, TRIP-Br2 is overexpressed in HepG2 cells and HCC, and a high expression of *SERTAD2* is associated with poor prognosis in HCC patients<sup>27</sup>. On the other side, a negative correlation was observed between HBV loads and the expression of *RB1* (Figure 4E), encoding the tumor suppressor retinoblastoma 1 (RB1), known to play a key role in HCC development<sup>28</sup>.

As the integration of the HBV genome into host cell has been suggested as one mechanism of HBV-induced liver carcinogenesis<sup>5</sup>, we investigated whether HBV integration was detectable at the single cell level through RNA-Seq. To address this question, we analyzed discordant pairs of reads where the read mates were mapping to both the HBV genome and to human genes. While many reads link HBV with unannotated regions of the human genome, we detected also affected gene bodies or their promoter regions, e.g., as shown in Figure 4F for *SFXN1*. In addition, we identified HBV reads in sequences corresponding to cancer genes like *SYPL1* (see Table S3 for the full list of genes). Its gene expression has been recently shown to predict HCC poor prognosis<sup>29</sup>. Even if we could not detect expression changes, the mutational impact at the single cell level might have consequences and is likely to play a role in HBV-induced cancer development. Taken together, our results suggest a profound impact of HBV RNA on the transcriptomic profile of single cells, including modulation of key cancer-related genes and pathways, likely contributing to hepatocarcinogenesis.

Collectively, the clinical implications of this study are two-fold: despite low levels of HBV replication with negative HBeAg, the presence of HBV RNA was associated with perturbation of expression of genes involved in carcinogenesis. These data suggest that even at low levels, HBV can trigger the expression of cancer-related genes, underlining the importance of curing HBV chronic infection in infected patients. Second, the marked heterogeneity of the tumor highlights the challenges of effective therapies: combination therapy targeting different drivers are most likely required for

treatment of HCC in this patient in case of non-resectable recurrence or advanced disease.

In conclusion, we show that the virus is compartmentalized within the tumor and liver cancer cells are characterized by marked heterogeneity within several clusters. Analysis of HBV-host cell interactions at the single cell levels revealed previously undiscovered pathways and driver candidates for carcinogenesis. These mechanisms represent previously undiscovered challenges but also provide opportunities for novel approaches for prevention and treatment of virus-induced HCC.



## **METHODS**

**Protocols.** Single cell RNA isolation, sequencing and data analyses are presented in Supplemental data.

**Data availability.** The SRA study accession number for the data reported in this study is SRP165160.

**Study approval.** Human liver tissue was obtained from patients followed at the Strasbourg University Hospitals, Strasbourg, France with informed consent. PHH were obtained from liver tissue from patients undergoing liver resection for liver metastasis at the Strasbourg University Hospitals with informed consent. Protocols were approved by the local Ethics Committee of the Strasbourg University Hospitals (CPP) and the Ministry of Higher Education and Research of France (DC 2016 2616).

**AUTHOR CONTRIBUTIONS:** TFB initiated the study and supervised research. PP provided human liver tissue. LH isolated and processed single cells. FJ performed computational analyses. FJ, ERV, AS, HES, CS, NP and TFB analyzed data and wrote the MS.

## **ACKNOWLEDGMENTS**

We thank E. Heuillard (Inserm UMR1110) for excellent technical support. This work was supported by Inserm, the University of Strasbourg, the European Union (Infect-ERA hepBccc, ERC-2014-AdG-671231-HEPCIR and Horizon 2020 research and innovation programme under grant agreement 667273 - HEPCAR), the Agence Nationale de Recherches sur le Sida et les Hépatites Virales (ANRS ECTZ50121), the French Cancer Agency (ARC IHU201301187) and the US National Institutes of Health (NIH/NIAID U19 AI123862-01, NIH/NIAID R03 AI131066, NIH/NCI R21 CA209940). This work has been published under the framework of the LabEx ANR-10-LAB-28 and benefits from a funding from the state managed by the French National Research Agency as part of the Investments for the Future (Investissements d'Avenir) program. A.S. is the recipient of a fellowship co-funded by the Région Alsace, France, the LabEx HepSys and IHU Strasbourg.

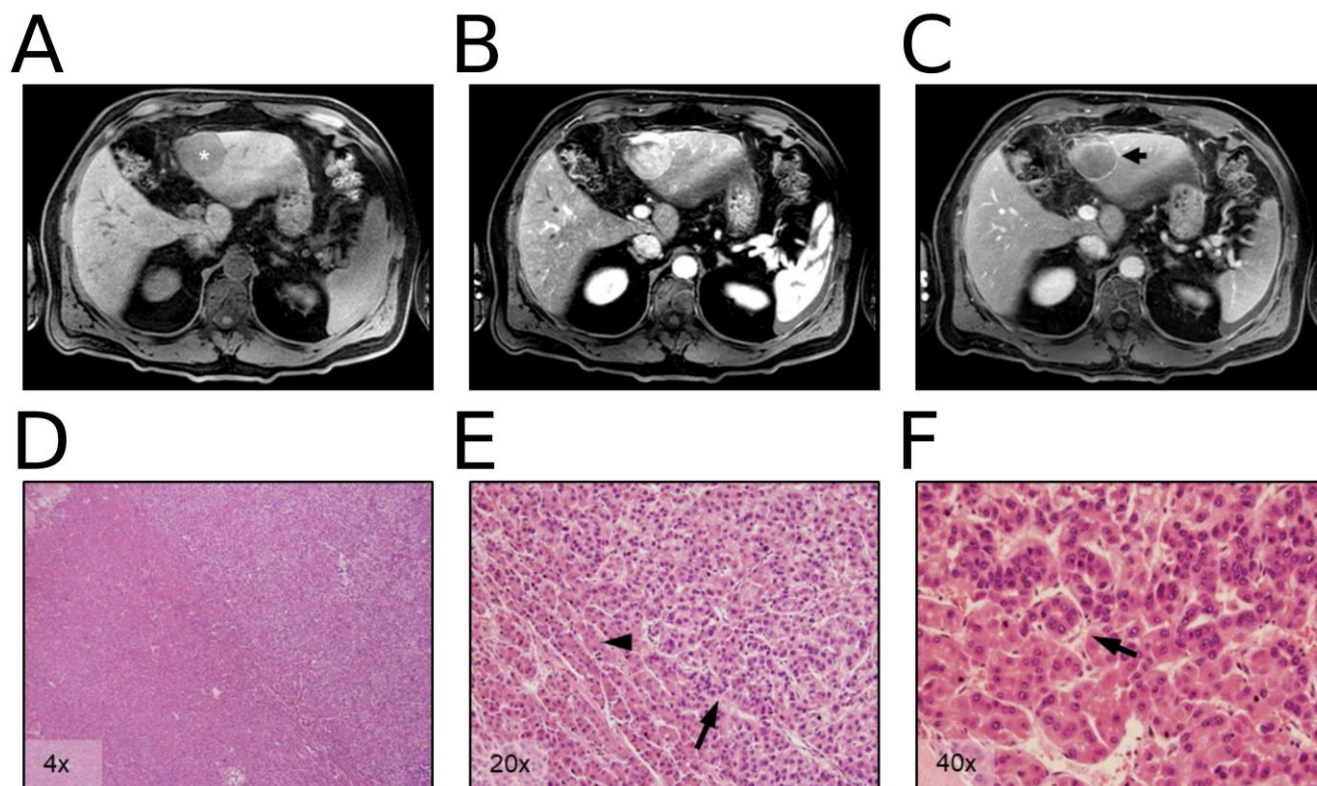
## REFERENCES

1. El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma. *Gastroenterology*. 2012;142(6):1264-1273 e1261.
2. Perz JF, Armstrong GL, Farrington LA, Hutin YJ, Bell BP. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J Hepatol*. 2006;45(4):529-538.
3. American Cancer Society. Cancer Facts & Figures 2014. *Atlanta: American Cancer Society*. 2014.
4. Flores A, Marrero JA. Emerging trends in hepatocellular carcinoma: focus on diagnosis and therapeutics. *Clin Med Insights Oncol*. 2014;8:71-76.
5. Levrero M, Zucman-Rossi J. Mechanisms of HBV-induced hepatocellular carcinoma. *J Hepatol*. 2016;64(1 Suppl):S84-S101.
6. Sandberg R. Entering the era of single-cell transcriptomics in biology and medicine. *Nat Methods*. 2014;11(1):22-24.
7. EASL. EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. *J Hepatol*. 2017;67(2):370-398.
8. Filmus J, Capurro M. Glypican-3: a marker and a therapeutic target in hepatocellular carcinoma. *FEBS J*. 2013;280(10):2471-2476.
9. Luo JH, et al. Transcriptomic and genomic analysis of human hepatocellular carcinomas and hepatoblastomas. *Hepatology*. 2006;44(4):1012-1024.
10. Uhlen M, et al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015;347(6220):1260419.
11. Kiselev VY, et al. SC3: consensus clustering of single-cell RNA-seq data. *Nat Methods*. 2017;14(5):483-486.
12. Boyault S, et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology*. 2007;45(1):42-52.

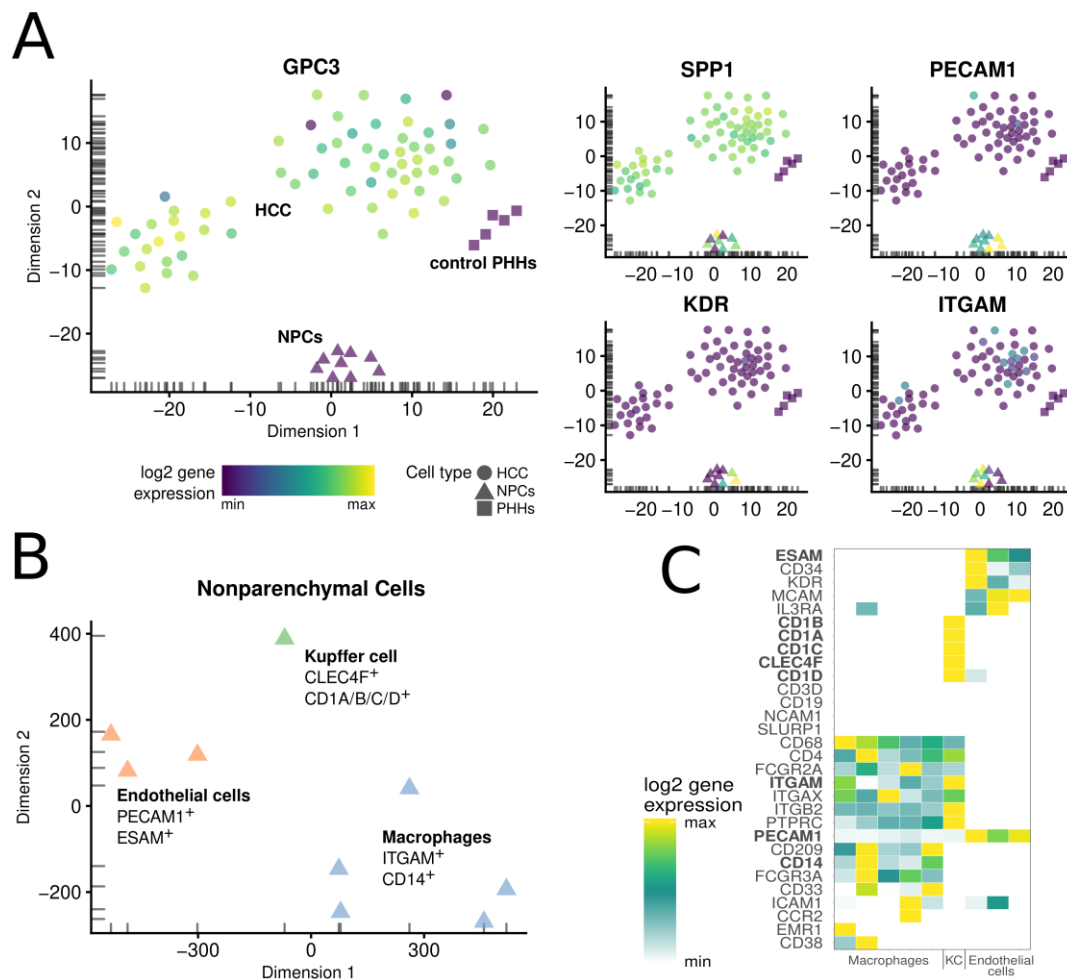
13. Hoshida Y, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res.* 2009;69(18):7385-7392.
14. Torrecilla S, et al. Trunk mutational events present minimal intra- and inter-tumoral heterogeneity in hepatocellular carcinoma. *J Hepatol.* 2017;67(6):1222-1231.
15. Friemel J, et al. Intratumor heterogeneity in hepatocellular carcinoma. *Clin Cancer Res.* 2015;21(8):1951-1961.
16. Chiang DY, et al. Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. *Cancer Res.* 2008;68(16):6779-6788.
17. Hu CT, Cheng CC, Wu JR, Pan SM, Wu WS. PKCepsilon-mediated c-Met endosomal processing directs fluctuant c-Met-JNK-paxillin signaling for tumor progression of HepG2. *Cell Signal.* 2015;27(7):1544-1555.
18. Jiang H, et al. Genetic variants in de novo lipogenic pathway genes predict the prognosis of surgically-treated hepatocellular carcinoma. *Sci Rep.* 2015;5:9536.
19. Tomsic J, et al. A germline mutation in SRRM2, a splicing factor gene, is implicated in papillary thyroid carcinoma predisposition. *Sci Rep.* 2015;5:10566.
20. Sveen A, Kilpinen S, Ruusulehto A, Lothe RA, Skotheim RI. Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes. *Oncogene.* 2016;35(19):2413-2427.
21. Verghese ET, et al. MiR-26b is down-regulated in carcinoma-associated fibroblasts from ER-positive breast cancers leading to enhanced cell migration and invasion. *J Pathol.* 2013;231(3):388-399.
22. Hedegaard DL, et al. High resolution sequencing of hepatitis C virus reveals limited intra-hepatic compartmentalization in end-stage liver disease. *J Hepatol.* 2017;66(1):28-38.
23. Yoneda M, et al. Hepatitis B Virus and DNA Stimulation Trigger a Rapid Innate Immune Response through NF-kappaB. *J Immunol.* 2016;197(2):630-643.
24. Lin D, Wu J. Hypoxia inducible factor in hepatocellular carcinoma: A therapeutic target. *World J Gastroenterol.* 2015;21(42):12171-12178.

25. Ju C, Colgan SP, Eltzschig HK. Hypoxia-inducible factors as molecular targets for liver diseases. *J Mol Med (Berl)*. 2016;94(6):613-627.
26. Liew CW, et al. Ablation of TRIP-Br2, a regulator of fat lipolysis, thermogenesis and oxidative metabolism, prevents diet-induced obesity and insulin resistance. *Nat Med*. 2013;19(2):217-226.
27. Cheong JK, et al. TRIP-Br2 promotes oncogenesis in nude mice and is frequently overexpressed in multiple human tumors. *J Transl Med*. 2009;7:8.
28. Kent LN, et al. Dosage-dependent copy number gains in E2f1 and E2f3 drive hepatocellular carcinoma. *J Clin Invest*. 2017;127(3):830-842.
29. Chen DH, Wu QW, Li XD, Wang SJ, Zhang ZM. SYPL1 overexpression predicts poor prognosis of hepatocellular carcinoma and associates with epithelial-mesenchymal transition. *Oncol Rep*. 2017;38(3):1533-1542.

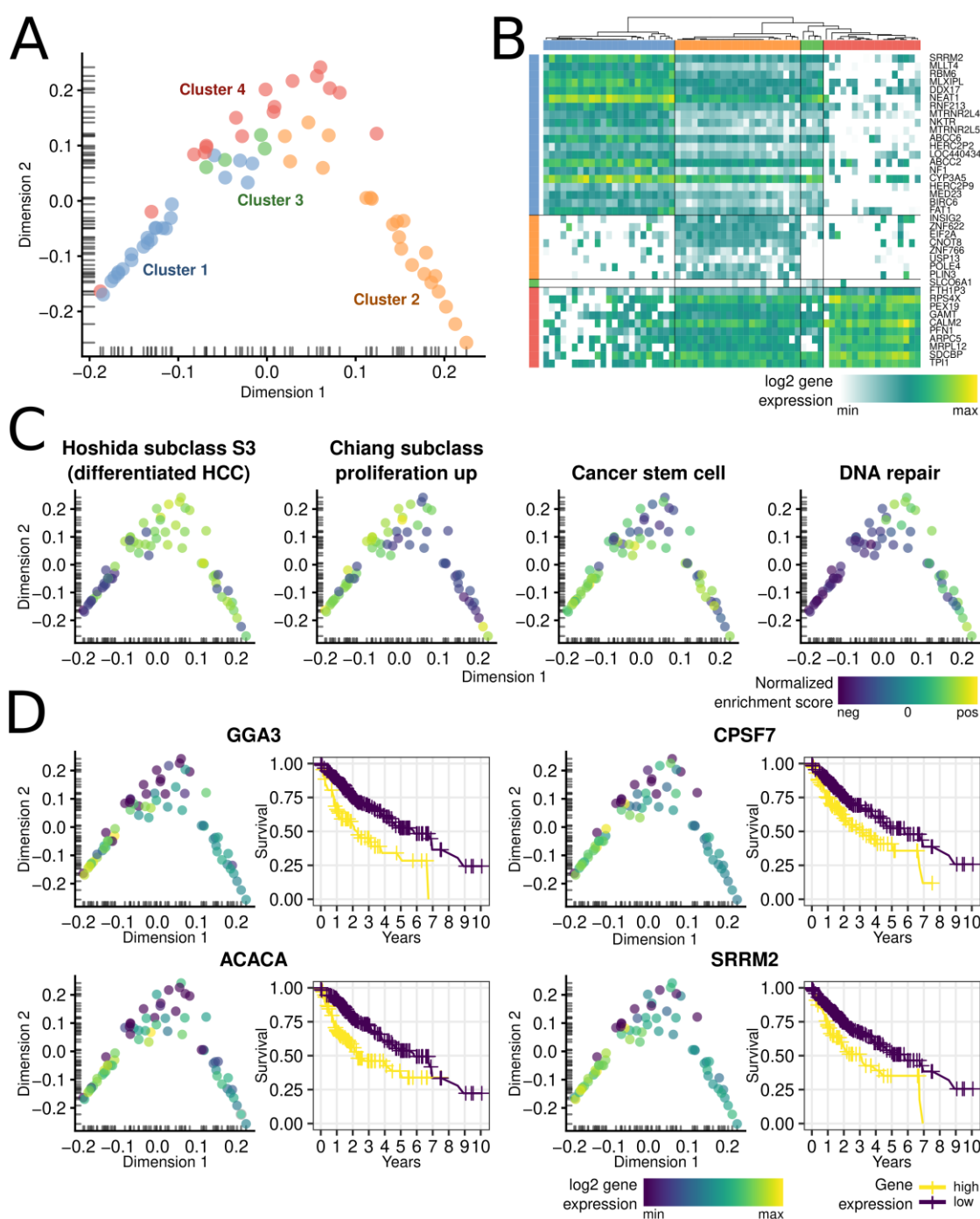
## FIGURES



**Figure 1. Radiological and histopathological features of the resected HCC.** At MRI, the nodule was hypointense in T2w sequences (A, white \*) and exhibited the typical radiological features of HCC consisting in arterial hyperenhancement (B) followed by wash-out in the portal phase (C). A peripheral delayed enhancement in the portal phase suggesting a tumor pseudocapsule was also noted (C, black arrow). The histopathological analyses showed a well to moderately differentiated HCC (D) displaying both a trabecular (E, black arrowheads) and pseudoglandular pattern (E, F black arrows).

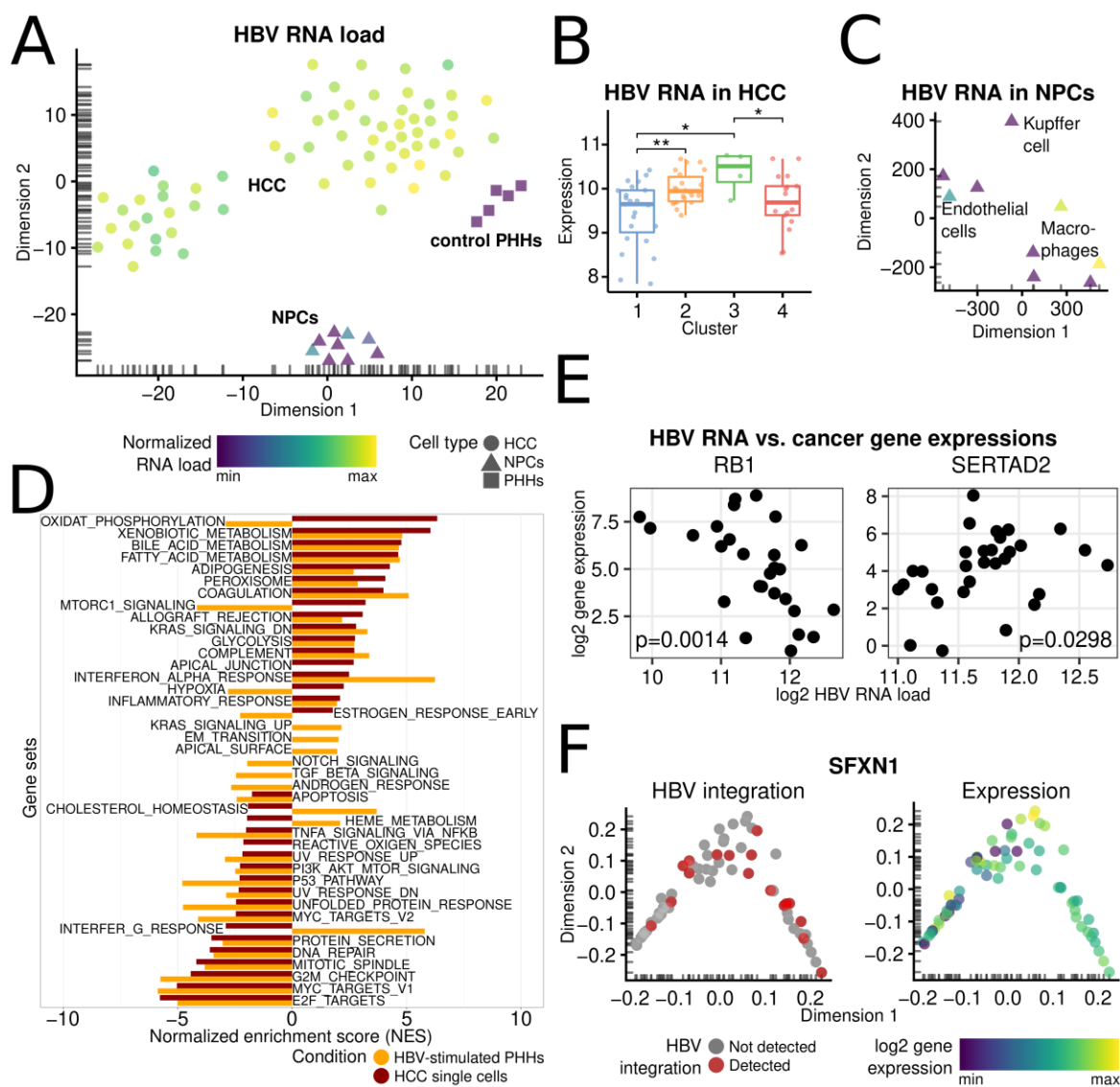


**Figure 2. Microenvironment of HCC.** A) T-SNE map and specific gene expressions in HCC cells and control PHHs. Control PHHs cluster together close to HCC cells, and also NPCs span their own cluster. Specific marker genes are expressed only in HCC (GPC3 and SPP1), or in NPCs (PECAM1, KDR, and ITGAM). Thereby, KDR and ITGAM are specifically expressed in macrophages or endothelial cells, respectively. B) A t-SNE map of NPCs indicates different cell types, i.e., macrophages, endothelial cells, and a Kupffer cell (KC) which were defined by specific marker genes expressed cell type-specific as shown in C).



**Figure 3. Intratumoral heterogeneity of HCC.** A) K-means clustering and branching of HCC-derived single cells illustrated on a diffusion map and B) Top specific marker genes for the 4 clusters (a maximum of 20 genes are shown, see Table S2 for full list). C) Cell-specific pathway analysis revealed different subclasses of HCC cells (differentiated phenotype in cluster 1 vs. high proliferation in cluster 2) with distinct features in comparison to control PHHs. D) Marker gene expression profiles as listed in A) with corresponding survival curves from TCGA patient data. A compelling proportion of marker genes upregulated specifically in cluster 1 are linked with poor survival (see Table S2 for full list).





**Figure 4. Heterogeneity of HBV RNA load in HCC.** A) T-SNE map of all HCC-derived cells including control PHHs, and HBV RNA load indicated. B) Compartmentalization with significant differences (Mann-Whitney test. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ) of HBV load in HCC cell clusters. C) HBV RNA levels in NPCs. D) Pathway analysis comparing enrichments in HBV-stimulated PHHs (GEO ID GSE69590), and enrichments of correlations of genes with HBV RNA load. E) HBV RNA load in comparison to cancer cell gene expression for *RB1* (tumor suppressor) and *SERTAD2* (oncogene) with  $p$ -values for spearman correlations shown in single cells with detected gene expression. F) Detected HBV integrations in *SFXN1* (left) and corresponding gene expression of single cells (right).

# Viral compartmentalization and cancer heterogeneity in HBV-induced hepatocellular carcinoma

Frank Jühling<sup>\$</sup>, Eloi R. Verrier<sup>\$</sup>, Antonio Saviano<sup>\$</sup>, Houssein El Saghire, Laura Heydmann, Patrick Pessaux, Nathalie Pochet, Catherine Schuster, Thomas F. Baumert

<sup>\$</sup>co-first authors

## SUPPLEMENTAL DATA

### Supplemental Methods

**Single cell RNA isolation and sequencing.** Single cells were isolated from fresh HCC tissue using collagenase digestion using a modified protocol as previously describe<sup>1</sup>. Human single hepatocytes were isolated from healthy liver tissue as described<sup>1</sup>. Tumor dissociation was performed using gentleMACS dissociator (Milteny) and the Tumor Dissociation kit for human biopsies (Milteny, 130-095-929) following the manufacturer's procedure. Single cells were isolated and sorted into 96-well plates using a MoFlo sorter (Beckman Coulters). Single cells were lysed in 15 µl TCL buffer (Qiagen, Hilden, Germany) supplemented with 1% 2-mercaptoethanol. Cellular mRNA was isolated and analyzed as described<sup>2,3</sup>. Paired-end 25bp reads were sequenced for control PHHs and HCC cells (one plate of 96 single cells for each) using the Smart-Seq2 protocol<sup>2,4,5</sup>. Reads were aligned to the human hg19 UCSC reference as well as the HBV genome (included as an additional chromosome) using hisat2 and suppressing discordant alignments for paired reads. Reads were counted using htseq-count<sup>6</sup>, and only 75 HCC-derived cells and 5 control PHHs with at least 200.000 reads mapping to genes were kept for further analysis. This resulted in a set of single cells with ~2.8 million reads and ~5.5 thousand genes covered (both median, see Figure S4 for details). Gene expression levels were quantified using DESeq2, including the calculation of normalized expression levels.

**Single cell clustering.** Single cell clustering and marker gene analysis was performed using SC3<sup>7</sup> (URL: <http://dx.doi.org/10.1038/nmeth.4236>) on log2 normalized expression levels. Clustering was performed for all cells, and for NPCs as well as cancer cells separately to provide deep distinctions of cell types and cell sub-types. The overall k-means clustering was performed with k=4 due to highest silhouette with (=0.73) for k>2. T-SNE maps and diffusion maps were calculated using the scater R package<sup>8,9</sup>.

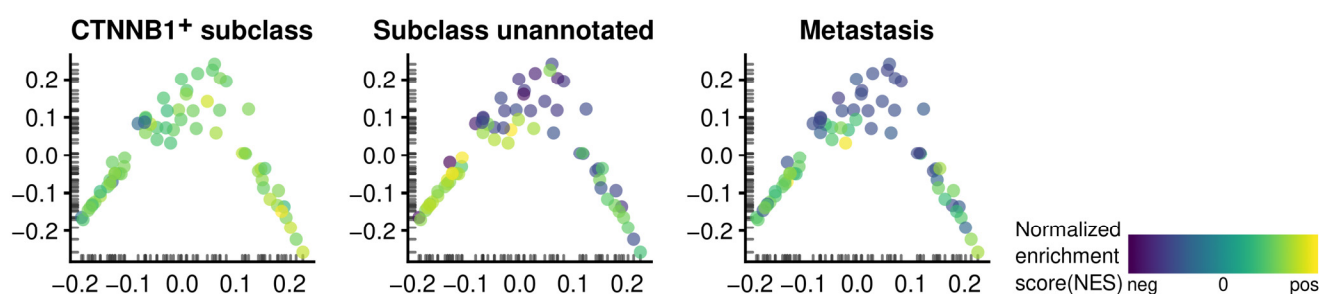
**HBV integration sites.** Reads were mapped again as described before, but here we allowed paired reads to be mapped to different chromosomes. Multiple discordant pairs of reads mapping uniquely to both, a human gene and the HBV chromosome, were considered as integration events in a single cell.

**Gene set enrichment analysis of single cells in association with HBV loads.** HBV RNA loads were processed identical to human gene expressions and normalized together. Correlations between individual gene expression and HBV loads were determined through the Spearman's rank correlation significance (p value), and genes were then ranked according to the Spearman's p value. Pathway enrichment analysis was assessed using the Pre-ranked Gene Set Enrichment Analysis (GSEA)<sup>10</sup> according to the obtained ranking. False discovery rate (FDR) below 0.05 was considered as statistically significant. Another experimental dataset (GSE69590) was used in this study, where gene expression of HBV-stimulated PHHs was compared to naïve PHHs. Pathway analysis was also assessed through

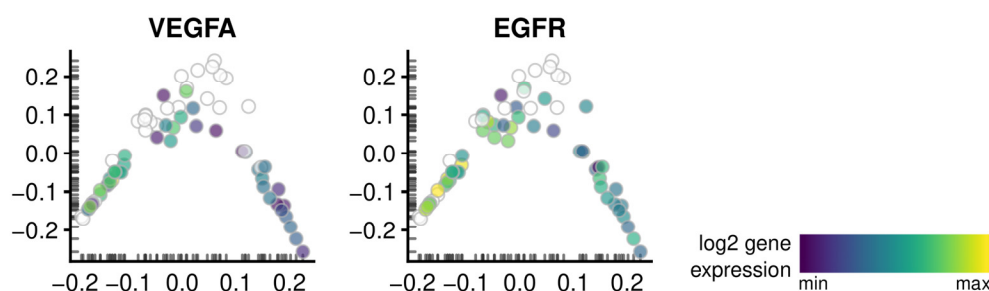
the pre-ranked GSEA based on p values of DESeq2 differential expression analysis. The normalized enrichment scores (NES) of pathways significantly differentially expressed between HBV-stimulated PHHs and naïve PHHs were then compared to the NES of pathways significantly modulated in association with HBV load in single cells.

**Statistics.** Clinical data was imported from The Human Protein Atlas<sup>11</sup>. The full data for cluster-specific marker genes identified as significant (un)favorable predictive genes was retrieved from the TCGA website<sup>12</sup>, and survival curves were calculated and drawn using corresponding R packages (survminer, survival, and ggplot2)<sup>13-15</sup>.

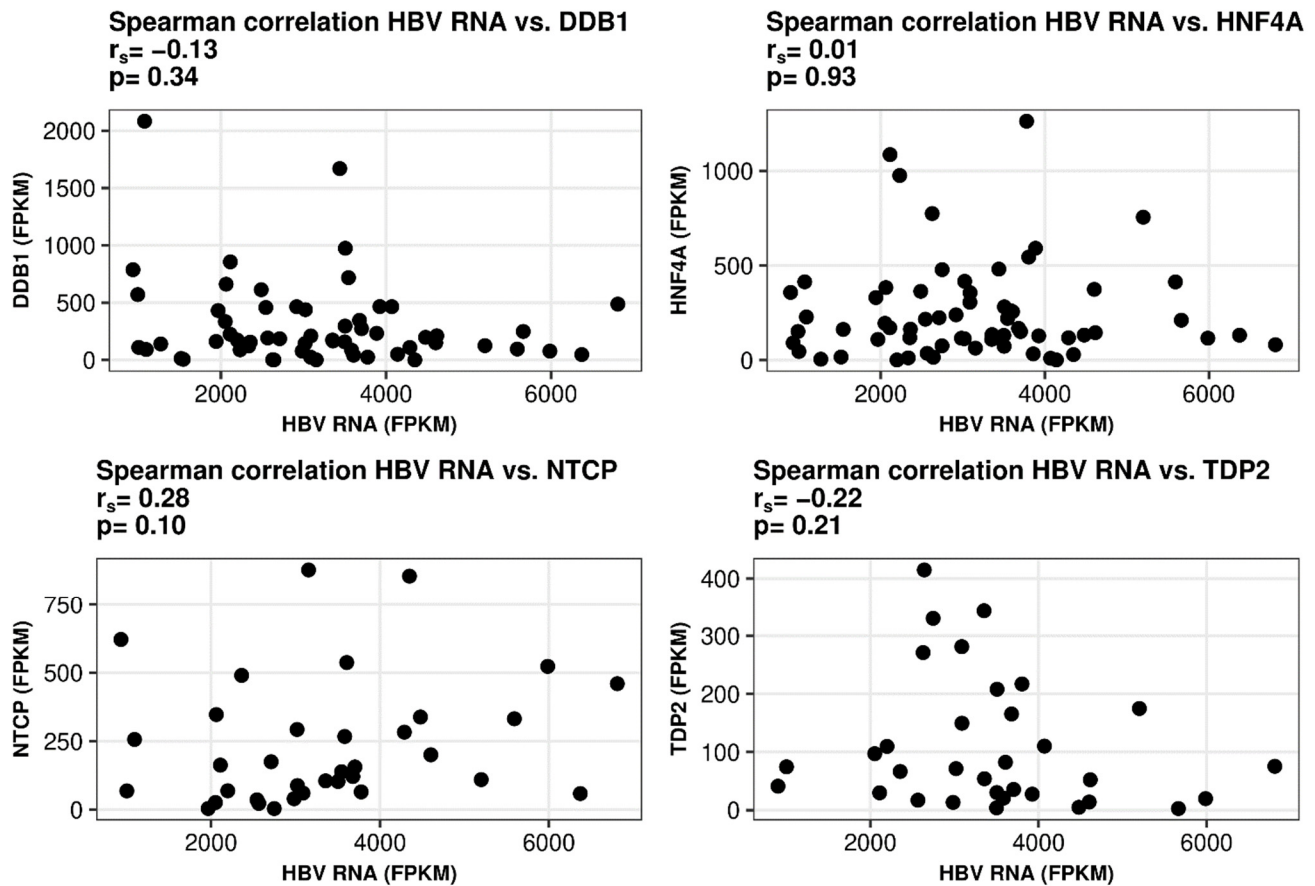
## Supplemental Figures



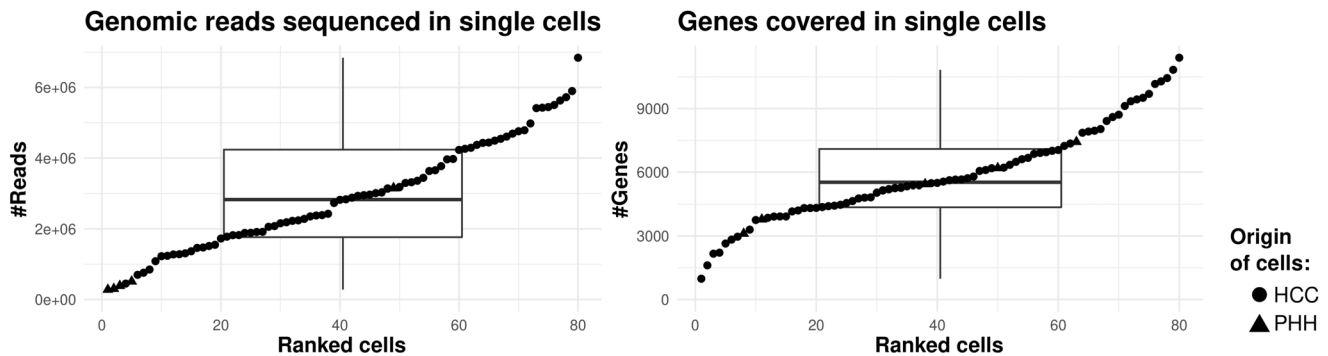
**Fig. S1: Intratumoral heterogeneity in HCC: cell-specific pathway analysis revealed subclasses of HCC cells with common and specific features.** While all HCC cells exhibit transcriptomic profiles common for the CTNNB1-positive subclass of HCC, the “unannotated” subclass defined in Chiang et al. (24) is specific for cells in cluster 1 (left bottom). In contrast, metastasis promoting genes are induced in both branches representing clusters 1 and 2 (left and right bottom) on a nearly equal level.



**Fig. S2: Intratumor heterogeneity in HCC: cell-specific expression of growth factors *VEGFA* and *EGFR*.** Both genes are predominantly expressed in clusters 1 and 2, and are higher expressed in cluster 1 (bottom left) compared to cluster 2 (bottom right).



**Fig. S3: Expression of HBV entry factors does not correlate with HBV RNA load in cancer cells.** Spearman correlations are shown for *DDB1*, *HNF4A*, *NTCP* (*SLC10A1*), and *TDP2* expressions compared to HBV RNA load.



**Fig. S4: Sequenced reads and covered genes in single cells.** Cells were sorted for number of reads mapping to genomic regions (left) or for covered genes (right), and a box plot was drawn indicating the distributions. In median, cells were sequenced with 2,831,670 reads and thereby covered 5,522.5 genes.

## Supplemental Tables

**Table S1.** Patient's laboratory data

|                                       | Reference range* | Patient         |
|---------------------------------------|------------------|-----------------|
| Glucose (mg/dl)                       | 74-100           | <b>170</b>      |
| Urea (mmol/l)                         | 2.5-7.0          | 7               |
| Creatinin (μmol/l)                    | 53.0-97.0        | 57.3            |
| Albumin (g/dl)                        | 3.5-5.0          | 4.4             |
| Total bilirubin (μmol/l)              | 1.7-21.0         | 8.6             |
| ALT (UI/l)                            | 15-33            | 27              |
| AST (UI/l)                            | 16-32            | 21              |
| ALP (UI/l)                            | 41-117           | 103             |
| Gamma-gt (UI/l)                       | 11-69            | 43              |
| Total Cholesterol (mg/dl)             | < 200            | 135             |
| Triglycerides (mg/dl)                 | < 150            | 101             |
| HbA1c (%)                             | < 6.0            | <b>8.1</b>      |
| Ferritin (μg/l)                       | 58-319           | 284             |
| Hemoglobin (g/dl)                     | 13.0-18.0        | 12.6            |
| Platelets count (10 <sup>9</sup> /l)  | 150-400          | 202             |
| White-cell count (10 <sup>9</sup> /l) | 4.10-10.50       | 5.37            |
| INR                                   | < 1.3            | 1.12            |
| CEA (kU/l)                            | < 5.0            | < 1.0           |
| Ca 19-9 (μg/l)                        | < 37.0           | 9.5             |
| AFP (μg/l)                            | < 13.0           | 2.1             |
| Total PSA (μg/l)                      | < 3.50           | 0.15            |
| HBsAg (UI/ml)                         | Negative         | <b>150.53</b>   |
| anti-HBc                              | Negative         | <b>Positive</b> |
| HBV-DNA (UI/ml)                       | Negative         | <b>71</b>       |
| anti-HCV                              | Negative         | Negative        |

AFP: alpha-fetoprotein, ALT: alanine aminotransferase, AST: aspartate aminotransferase, CEA: carcinoembryonic antigen, HBV: hepatitis B virus, HBsAg: hepatitis B surface antigen HCV: hepatitis C virus, INR: international normalized ratio, PSA: prostate specific antigen.

All the blood tests, except ferritin, were performed before liver resection. In bold are reported the values not included in the corresponding reference range.

\*Reference values are affected by many variables. The ranges used at the Nouvel Hôpital Civil of Strasbourg are for non-pregnant adults who do not have medical conditions affecting the results.

**Table S2:** Full list of marker genes with prognostic data (*favorable* corresponds with good prognosis and *unfavorable* corresponds with poor prognosis; *p*-values adjusted according to Holm<sup>1</sup>) for clusters 1-4. Top marker genes from this table are shown in Fig. 3B.

<sup>1</sup>Holm, Sture. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*. 6. 65-70. 10.2307/4615733.

\*AUC: Area Under the ROC Curve

\*\*Prognostic data imported from <https://www.proteinatlas.org/>

| Cluster | AUC*  | Adjusted<br><i>p</i> -value | Symbol    | Prognostic<br>favorable** | Prognostic<br>unfavorable** |
|---------|-------|-----------------------------|-----------|---------------------------|-----------------------------|
| 1       | 0.996 | 4.68E-07                    | SRRM2     |                           | 5.09E-04                    |
| 1       | 0.992 | 6.04E-07                    | MLLT4     |                           |                             |
| 1       | 0.984 | 1.04E-06                    | RBM6      |                           |                             |
| 1       | 0.980 | 1.95E-06                    | MLXIPL    |                           |                             |
| 1       | 0.979 | 2.11E-06                    | DDX17     |                           |                             |
| 1       | 0.979 | 1.99E-06                    | NEAT1     |                           |                             |
| 1       | 0.977 | 2.54E-06                    | RNF213    |                           |                             |
| 1       | 0.974 | 3.30E-06                    | MTRNR2L4  |                           |                             |
| 1       | 0.970 | 3.58E-06                    | NKTR      |                           |                             |
| 1       | 0.968 | 5.03E-06                    | MTRNR2L5  |                           |                             |
| 1       | 0.967 | 5.77E-06                    | ABCC6     | 3.82E-04                  |                             |
| 1       | 0.967 | 4.65E-06                    | HERC2P2   |                           |                             |
| 1       | 0.966 | 6.40E-06                    | LOC440434 |                           |                             |
| 1       | 0.965 | 6.97E-06                    | ABCC2     |                           |                             |
| 1       | 0.963 | 6.90E-06                    | NF1       |                           |                             |
| 1       | 0.961 | 9.94E-06                    | CYP3A5    | 3.72E-05                  |                             |

|   |       |          |          |          |          |
|---|-------|----------|----------|----------|----------|
| 1 | 0.959 | 9.71E-06 | HERC2P9  |          |          |
| 1 | 0.959 | 8.21E-06 | MED23    |          |          |
| 1 | 0.958 | 1.01E-05 | BIRC6    |          |          |
| 1 | 0.957 | 1.35E-05 | FAT1     |          |          |
| 1 | 0.952 | 1.84E-05 | PCSK6    |          |          |
| 1 | 0.947 | 2.79E-05 | COL18A1  | 6.16E-04 |          |
| 1 | 0.943 | 2.91E-05 | CPS1-IT1 |          |          |
| 1 | 0.942 | 3.85E-05 | CLMN     |          |          |
| 1 | 0.941 | 3.84E-05 | TNRC6A   |          |          |
| 1 | 0.941 | 4.33E-05 | UBR4     |          |          |
| 1 | 0.938 | 5.68E-05 | CPSF7    |          | 9.83E-04 |
| 1 | 0.938 | 2.93E-05 | KAT2A    |          |          |
| 1 | 0.938 | 6.12E-05 | LAMB2    |          |          |
| 1 | 0.938 | 4.13E-05 | DNHD1    |          |          |
| 1 | 0.937 | 5.33E-05 | PILRB    |          |          |
| 1 | 0.937 | 6.70E-05 | ABCA1    |          |          |
| 1 | 0.936 | 6.04E-05 | CCNL2    |          |          |
| 1 | 0.936 | 4.86E-05 | TRIM66   |          |          |
| 1 | 0.934 | 7.63E-05 | PRKDC    |          |          |
| 1 | 0.931 | 1.08E-04 | CHD1L    |          | 3.07E-04 |
| 1 | 0.929 | 1.26E-04 | SLC38A10 |          |          |

|   |       |          |              |          |          |
|---|-------|----------|--------------|----------|----------|
| 1 | 0.928 | 1.37E-04 | ATP2A2       |          |          |
| 1 | 0.927 | 1.24E-04 | NRG1         |          |          |
| 1 | 0.927 | 1.40E-04 | SPATA13      |          |          |
| 1 | 0.926 | 1.59E-04 | MST1P2       |          |          |
| 1 | 0.924 | 1.69E-04 | SRCAP        |          |          |
| 1 | 0.924 | 1.74E-04 | CDK12        |          |          |
| 1 | 0.924 | 1.58E-04 | RALGAPA2     |          |          |
| 1 | 0.922 | 1.31E-04 | MTRNR2L7     |          |          |
| 1 | 0.921 | 2.35E-04 | C5           | 9.43E-04 |          |
| 1 | 0.921 | 2.24E-04 | ABCC6P1      |          |          |
| 1 | 0.920 | 1.77E-04 | LOC100132247 |          |          |
| 1 | 0.919 | 2.62E-04 | TEAD1        |          |          |
| 1 | 0.919 | 2.24E-04 | LOC100131564 |          |          |
| 1 | 0.918 | 2.62E-04 | AASS         | 9.81E-04 |          |
| 1 | 0.918 | 2.52E-04 | ACACA        |          | 1.24E-05 |
| 1 | 0.918 | 2.31E-04 | STAG3L2      |          |          |
| 1 | 0.917 | 1.79E-04 | SLFNL1-AS1   |          |          |
| 1 | 0.915 | 2.10E-04 | FAM20A       |          |          |
| 1 | 0.915 | 3.19E-04 | GANC         |          |          |
| 1 | 0.915 | 3.30E-04 | TBC1D8       |          |          |
| 1 | 0.914 | 3.56E-04 | INTS3        |          |          |



|   |       |          |         |  |          |
|---|-------|----------|---------|--|----------|
| 1 | 0.914 | 3.56E-04 | PRRC2B  |  |          |
| 1 | 0.914 | 3.00E-04 | LENG8   |  |          |
| 1 | 0.911 | 4.93E-04 | IL6R    |  |          |
| 1 | 0.910 | 5.40E-04 | ZKSCAN1 |  |          |
| 1 | 0.910 | 4.30E-04 | SPDYE6  |  |          |
| 1 | 0.908 | 4.59E-04 | DYNC1H1 |  | 3.03E-07 |
| 1 | 0.908 | 5.45E-04 | SNRNP70 |  |          |
| 1 | 0.906 | 7.30E-04 | H6PD    |  |          |
| 1 | 0.904 | 7.38E-04 | PLXNB1  |  |          |
| 1 | 0.902 | 9.35E-04 | KANSL1  |  | 3.74E-04 |
| 1 | 0.900 | 1.04E-03 | POGZ    |  |          |
| 1 | 0.900 | 1.08E-03 | HTT     |  |          |
| 1 | 0.900 | 5.68E-04 | LAMA5   |  |          |
| 1 | 0.899 | 1.07E-03 | CLTCL1  |  |          |
| 1 | 0.899 | 9.44E-04 | STAG3L1 |  |          |
| 1 | 0.899 | 5.54E-04 | TLK2    |  | 9.40E-04 |
| 1 | 0.898 | 1.15E-03 | MDM4    |  |          |
| 1 | 0.898 | 7.59E-04 | RICTOR  |  |          |
| 1 | 0.898 | 1.32E-03 | IDO2    |  |          |
| 1 | 0.896 | 1.24E-03 | IL17RB  |  |          |
| 1 | 0.896 | 1.42E-03 | ACACB   |  |          |

|   |       |          |              |          |          |
|---|-------|----------|--------------|----------|----------|
| 1 | 0.895 | 1.65E-03 | CELF1        |          | 7.56E-04 |
| 1 | 0.895 | 1.44E-03 | PRPF3        |          | 6.59E-05 |
| 1 | 0.893 | 9.71E-04 | LOC100133331 |          |          |
| 1 | 0.892 | 2.07E-03 | PPARGC1A     | 3.35E-06 |          |
| 1 | 0.891 | 1.48E-03 | PLEC         |          |          |
| 1 | 0.891 | 2.22E-03 | CPT1A        |          |          |
| 1 | 0.890 | 6.81E-04 | DNAJB3       |          |          |
| 1 | 0.890 | 1.65E-03 | TRIM25       |          | 9.37E-04 |
| 1 | 0.890 | 2.37E-03 | CNOT1        |          |          |
| 1 | 0.888 | 1.84E-03 | SS18L1       |          |          |
| 1 | 0.888 | 2.70E-03 | MYO18A       |          |          |
| 1 | 0.888 | 2.13E-03 | MAN2C1       |          |          |
| 1 | 0.886 | 1.87E-03 | MDN1         |          |          |
| 1 | 0.885 | 2.97E-03 | GLG1         |          |          |
| 1 | 0.885 | 3.16E-03 | RBM25        |          | 1.42E-04 |
| 1 | 0.885 | 2.77E-03 | USP24        |          | 1.34E-07 |
| 1 | 0.884 | 3.19E-03 | PPIE         |          | 1.76E-04 |
| 1 | 0.884 | 3.21E-03 | SLC23A2      | 3.35E-04 |          |
| 1 | 0.883 | 3.20E-03 | MYOM1        |          |          |
| 1 | 0.882 | 3.13E-03 | FLJ45340     |          |          |
| 1 | 0.882 | 3.70E-03 | GGA3         |          | 2.77E-06 |

|   |       |          |          |  |          |
|---|-------|----------|----------|--|----------|
| 1 | 0.882 | 3.13E-03 | GOLGA8B  |  |          |
| 1 | 0.882 | 3.44E-03 | DST      |  |          |
| 1 | 0.881 | 4.58E-03 | RREB1    |  |          |
| 1 | 0.881 | 4.58E-03 | SF1      |  | 4.85E-05 |
| 1 | 0.880 | 3.98E-03 | ERN1     |  |          |
| 1 | 0.879 | 4.55E-03 | POLR2J3  |  |          |
| 1 | 0.879 | 1.75E-03 | ATP13A1  |  |          |
| 1 | 0.878 | 5.65E-03 | NPLOC4   |  | 1.16E-05 |
| 1 | 0.876 | 4.30E-03 | SCMH1    |  | 3.82E-04 |
| 1 | 0.875 | 4.62E-03 | RGPD1    |  |          |
| 1 | 0.870 | 5.75E-03 | RNF217   |  |          |
| 1 | 0.869 | 1.02E-02 | ABCA5    |  |          |
| 1 | 0.869 | 9.71E-03 | N4BP2L2  |  |          |
| 1 | 0.869 | 5.24E-03 | LAMA3    |  |          |
| 1 | 0.868 | 1.06E-02 | HELZ     |  |          |
| 1 | 0.867 | 1.13E-02 | SIPA1L2  |  |          |
| 1 | 0.867 | 1.03E-02 | CCNL1    |  |          |
| 1 | 0.867 | 6.05E-03 | CCDC144B |  |          |
| 1 | 0.866 | 1.21E-02 | ACSF2    |  |          |
| 1 | 0.866 | 5.71E-03 | ZMIZ2    |  | 6.82E-05 |
| 1 | 0.866 | 1.10E-02 | ARGLU1   |  |          |

|   |       |          |              |          |          |
|---|-------|----------|--------------|----------|----------|
| 1 | 0.866 | 1.25E-02 | SYVN1        |          |          |
| 1 | 0.865 | 1.37E-02 | MTMR4        |          |          |
| 1 | 0.865 | 1.28E-02 | FOXK2        |          | 1.24E-06 |
| 1 | 0.863 | 1.37E-02 | CLK1         |          |          |
| 1 | 0.863 | 7.49E-03 | DENND4B      |          | 4.14E-04 |
| 1 | 0.863 | 1.22E-02 | GCN1L1       |          |          |
| 1 | 0.861 | 1.01E-02 | AFG3L1P      |          |          |
| 1 | 0.861 | 1.56E-02 | KIAA1731     |          |          |
| 1 | 0.861 | 1.34E-02 | UBR5         |          | 1.26E-04 |
| 1 | 0.860 | 1.63E-02 | PNISR        |          |          |
| 1 | 0.860 | 1.49E-02 | HECTD4       |          |          |
| 1 | 0.859 | 1.92E-02 | ABCB11       |          |          |
| 1 | 0.859 | 8.23E-03 | LOC100286922 |          |          |
| 1 | 0.859 | 1.47E-02 | MLL3         |          |          |
| 1 | 0.859 | 1.70E-02 | MTOR         |          |          |
| 1 | 0.858 | 1.64E-02 | PTPRM        |          |          |
| 1 | 0.857 | 2.17E-02 | IL6ST        |          |          |
| 1 | 0.857 | 2.04E-02 | ABCG5        | 2.60E-05 |          |
| 1 | 0.857 | 1.43E-02 | EP400        |          | 2.11E-04 |
| 1 | 0.857 | 2.00E-02 | NBEAL1       |          |          |
| 1 | 0.857 | 2.00E-02 | WDR59        |          |          |

|   |       |          |          |  |          |
|---|-------|----------|----------|--|----------|
| 1 | 0.857 | 1.82E-02 | NADSYN1  |  | 5.17E-06 |
| 1 | 0.856 | 1.80E-02 | ATG4B    |  |          |
| 1 | 0.856 | 1.80E-02 | TPTE2P5  |  |          |
| 1 | 0.855 | 1.53E-02 | AGRN     |  | 4.62E-04 |
| 1 | 0.855 | 1.31E-02 | CCDC18   |  |          |
| 1 | 0.855 | 2.08E-02 | MON2     |  |          |
| 1 | 0.855 | 2.08E-02 | SLC17A9  |  |          |
| 1 | 0.854 | 2.58E-02 | DYNC1LI2 |  |          |
| 1 | 0.853 | 2.87E-02 | LEPR     |  |          |
| 1 | 0.853 | 2.38E-02 | UBR2     |  |          |
| 1 | 0.853 | 2.87E-02 | PTPRF    |  | 3.47E-04 |
| 1 | 0.853 | 1.17E-02 | TRPC4AP  |  | 6.25E-08 |
| 1 | 0.853 | 1.17E-02 | SIN3B    |  |          |
| 1 | 0.852 | 2.45E-02 | OGT      |  |          |
| 1 | 0.852 | 1.88E-02 | MLLT6    |  |          |
| 1 | 0.852 | 3.01E-02 | SORL1    |  |          |
| 1 | 0.851 | 2.97E-02 | NRBP2    |  |          |
| 1 | 0.851 | 2.62E-02 | RECQL5   |  |          |
| 1 | 0.851 | 3.07E-02 | CARD8    |  |          |
| 1 | 0.850 | 3.35E-02 | TRIP12   |  | 3.60E-04 |
| 2 | 0.881 | 1.12E-03 | INSIG2   |  |          |

|   |       |          |         |  |          |
|---|-------|----------|---------|--|----------|
| 2 | 0.876 | 1.19E-03 | ZNF622  |  |          |
| 2 | 0.875 | 3.60E-03 | EIF2A   |  | 2.36E-05 |
| 2 | 0.869 | 2.00E-03 | CNOT8   |  | 5.33E-04 |
| 2 | 0.865 | 1.93E-04 | ZNF766  |  |          |
| 2 | 0.856 | 8.58E-04 | USP13   |  | 2.40E-05 |
| 2 | 0.854 | 1.19E-04 | POLE4   |  | 6.49E-04 |
| 2 | 0.852 | 1.21E-03 | PLIN3   |  | 3.28E-05 |
| 3 | 0.948 | 1.80E-03 | SLCO6A1 |  |          |
| 4 | 0.909 | 6.62E-03 | FTH1P3  |  |          |
| 4 | 0.908 | 6.80E-03 | RPS4X   |  |          |
| 4 | 0.904 | 8.19E-03 | PEX19   |  |          |
| 4 | 0.899 | 1.23E-02 | GAMT    |  |          |
| 4 | 0.898 | 1.31E-02 | CALM2   |  | 8.35E-05 |
| 4 | 0.891 | 2.01E-02 | PFN1    |  |          |
| 4 | 0.885 | 2.82E-02 | ARPC5   |  | 3.63E-04 |
| 4 | 0.884 | 2.76E-02 | MRPL12  |  |          |
| 4 | 0.882 | 3.33E-02 | SDCBP   |  |          |
| 4 | 0.878 | 4.47E-02 | TPI1    |  | 1.04E-05 |

**Table S3** List of genes with detected HBV integration sites with the number of affected single cells.

| #Cells | Gene   |
|--------|--|
| 17     | SFXN1  |
| 4      | KIRREL   |
| 3      | CAMTA1, DLG2, LDLRAD3, RP11-138M12.1, TCERG1L  |
| 2      | MEF2A, RNF114  |
| 1      | AC007563.5, ADH4, ANKRD30BL, APOH, ARHGEF38, BAZ2B, CASC15, CCM2, CNTRL, CTD-2043I16.1, CTD-2547E10.2, CUTC, DCC, DIS3L2, DOK7, DPP10, DYSF, DZIP1, EFHC2, EGLN1, EXOSC8, FMN2, GPATCH2, GPHN, KCNIP4, LEKR1, LHFPL2, LHFPL3, LINC01029, MAP4, MELK, MICU2, MYBPC1, NCOA4, PDE1C, PDE3A, PIGU, PLOD1, PROS1, RAB11FIP5, RP11-22P4.1, RP11-383H13.1, RP11-3J1.1, RP11-696N14.1, RP4-536B24.3, RP4-694A7.4, SAR1B, SETD3, SPATS2L, SRD5A1, SYPL1, SYT14, TBL1XR1, THSD7B, TPRX1, ZFH4-AS1, ZKSCAN8, ZNF765 |

### **Supplemental references**

1. Krieger SE, et al. Inhibition of hepatitis C virus infection by anti-claudin-1 antibodies is mediated by neutralization of E2-CD81-claudin-1 associations. *Hepatology*. 2010;51(4):1144-1157.
2. Trombetta JJ, et al. Preparation of Single-Cell RNA-Seq Libraries for Next Generation Sequencing. *Curr Protoc Mol Biol*. 2014;107:4 22 21-24 22 17.
3. Shalek AK, et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature*. 2014;510(7505):363-369.
4. Picelli S, et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*. 2014;9(1):171-181.
5. Picelli S, et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods*. 2013;10(11):1096-1098.
6. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-169.
7. Kiselev VY, et al. SC3: consensus clustering of single-cell RNA-seq data. *Nat Methods*. 2017;14(5):483-486.
8. Team RC. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. 2018(06 Feb 2018).
9. McCarthy DJ, Campbell KR, Lun AT, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics*. 2017;33(8):1179-1186.
10. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
11. Uhlen M, et al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015;347(6220):1260419.
12. Weinstein JN, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013;45(10):1113-1120.
13. Wickham H. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*. 2009.
14. Kassambara A, Kosinski M. survminer: Drawing Survival Curves using 'ggplot2'. R package version 0.4.2. . 2018.
15. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. Springer, New York. ed2000.

## 4. CONCLUSIONS AND PERSPECTIVES

We have established a robust pipeline to perform scRNA-seq on fresh and cryopreserved liver tissues and demonstrated to be able to characterize rare cell types in normal liver such as bipotent progenitor cells. In single cell studies, tissue dissociation is a crucial step. Soft dissociation procedures can yield a low number of single cells especially in complex and fibrotic tissues. Hard and long dissociation steps can induce transcriptomic changes and compromise the analysis<sup>207</sup>.

The liver is the central hub of human metabolism, a major immunoregulatory organ and it is capable of impressive regeneration. My work demonstrated that normal liver is highly organized and composed by heterogeneous cell populations. Using a high-resolution scRNA-seq technique, we identified 39 cell clusters with significant transcriptomic differences. As an example, endothelial cells can be divided in at least 11 subtypes. Importantly, we showed that in the normal liver both immunosuppressive KCs with metabolic/scavenger functions and KCs expressing more inflammatory genes coexist, suggesting that in normal conditions the different types of macrophages are finely regulated and balanced.

Recent studies carried out in mouse showed that liver zonation is not monotonic and is extended also to non-parenchymal cells (e.g. endothelial cells) playing an important role in determining hepatocyte function and organization<sup>181,182</sup>. Little was known about zonation in human liver. We were able to demonstrate that, also in human liver, the zonation is not monotonic and is extended to non-parenchymal cells. We accurately described zonation patterns gene by gene and found



that 41% of hepatocytes genes and 67% of endothelial cells genes are significantly zoned. A comparison of zonation patterns between mouse and human showed only partial common gene patterns. These findings highlight, once again, the complexity of the human liver and the low grade of similarity between mouse and human.

Regeneration is a hallmark of liver physiopathology. However, this process is still poorly understood and most of the data available so far were generated in mice. We explored the heterogeneity of EPCAM<sup>+</sup> population and found hepatocyte-biased and cholangiocyte-biased cells. A fate analysis corroborated by functional experiments in liver organoids allowed us to discover and describe a TROP2<sup>int</sup> cell population that resides in the small bile ducts, displays lower level of mature cholangiocyte markers (e.g. CK19) and has the potential to differentiate into hepatocytes. This finding opens the door to a better understanding of mechanisms and players involved in liver regeneration and repair.

Our human liver cell atlas is a resource for the entire liver community and could serve as a reference to study human diseases and models at single-cell level. As a proof-of-concept we performed scRNA-seq of human hepatocytes and endothelial cells engrafted in FRG-NOD mice and demonstrated that human cells change their transcriptomic in the mouse microenvironment upregulating genes involved in cell cycle, WNT, Hedgehog and VEGF signaling.

We also performed scRNA-seq on three HCCs and used the human liver cell atlas to study the perturbations associated with cancer transformation. HCC cells showed a downregulation in pathways associated with metabolism and oxidation while upregulated cell cycle checkpoints,

WNT and Hedgehog signaling. Interestingly, HCC endothelial cells were enriched of genes involved in extracellular matrix organization, receptor tyrosine kinase signaling and downregulated in genes belonging to the toll-like receptor cascade, suggesting that these cells play also an important role in modulating innate immunity in liver cancer.

Chronic HBV infection is a major cause of liver disease mortality and HCC worldwide. HBV can integrate to the human genome and induce liver cancer via direct and/or indirect mechanisms. Not all HBV-infected patients will receive antiviral treatment. Patients with HBeAg-negative chronic infection (previously called “inactive carriers”) that have no sign of hepatitis or advanced liver fibrosis or any virus-associated complications are not treated with antiviral drugs since the infection is considered indolent. Using Smart-Seq2 scRNA-seq, allowing for deep sequencing of full RNA transcripts, we analyzed cells from an HBV-related HCC and showed a marked tumor heterogeneity and HBV-RNA compartmentalization. HBV-RNA levels were different among HCC clusters and such a difference was not only related to viral integration which we mapped at the single cell level. HBV-RNA is higher in more differentiated HCC cells, suggesting that HBV plays a major role in the first phases of carcinogenesis and tumor progression. HBV-RNA load was associated with higher levels of oncogenes and lower levels of tumor suppressor (e.g. *RB1*). We reported for the first time the correlation of HBV-RNA load with the level of the oncogene *SERTAD2*. Finally, pathways enriched in HCC cells with higher HBV load were comparable with the ones enriched in primary human hepatocytes stimulated with HBV. These findings suggest that even at low levels, HBV has an important role in HCC pathogenesis. In this context, our

analysis of HBV-host interaction at single cell level gave new insights and revealed new pathways, drivers and mechanism of HBV-related carcinogenesis.

These works open important perspectives for future translational research in liver diseases. ScRNA-seq can be used to characterize the physiopathology of NASH and HCC heterogeneity to ultimately identify new therapeutic targets for these diseases.

Tumor resistance to systemic drugs can also be analyzed at single-cell level. Indeed, tumor composition changes upon treatment, with selection of resistant difficult-to-treat clones. Several studies reported cancer stem cell-like enrichment in HCC after sorafenib<sup>43,45</sup>. These data, deriving exclusively from cell lines or patient-derived models, would need confirmation in primary human tumors. ScRNA-seq of treatment-resistant HCC will give remarkable new insights of HCC biology and changes following systemic treatment.

Finally, the human liver cell atlas could also be used to study the physiopathology and discover therapeutic strategies for rare liver cancers such as ICC and rare liver diseases as primary sclerosing cholangitis.

In my next research project, I will use our human liver cell atlas as reference and applying scRNA-seq to dissect NASH physiopathology and hepatocarcinogenesis. Using our established pipeline, I will perform paired scRNA-seq on HCC and the adjacent non-tumor tissue from NASH patients at different disease stages. In particular, I will focus on the interactions between hepatocytes and the liver microenvironment and its role in triggering and maintain chronic inflammation and inducing HCC. Our human liver cell atlas data integrated with a paired sequencing at single cell

level of HCC and adjacent non-tumor liver tissue as well as early and high grade dysplastic nodules will help to (i) identify chronic liver disease specific changes associated with HCC development, which could be extremely relevant for NASH carcinogenesis that is peculiar and still poorly understood, (ii) discover drivers of tumor development and modulators of cancer immune surveillance to establish effective HCC chemopreventive strategies as well as adjuvant therapies for HCC, (iii) provide new insights on the role of microenvironment in sustaining inflammation in NASH.

In conclusion, my research work has lead the way to the study of liver physiopathology at single cell level and provided the scientific community with valuable data and tools to identify novel therapeutic targets for liver fibrosis treatment and HCC chemoprevention. In future translational research projects, I will focus on the mechanisms of liver carcinogenesis in NASH and advanced fibrosis using the human liver cell atlas as reference.

## 5. RESUME DE LA RECHERCHE

### 5.1 INTRODUCTION

La cirrhose décompensée est la quatrième cause de décès chez les adultes en Europe centrale et le carcinome hépatocellulaire (CHC) est la deuxième cause de décès par cancer dans le monde (<http://globocan.iarc.fr>). Plusieurs études suggèrent qu'il existe des voies communes conduisant à la pathogenèse des maladies hépatiques avancées et du CHC, quelle que soit leur étiologie. Toutes les maladies chroniques du foie sont caractérisées par une inflammation chronique, une fibrose progressive et un risque élevé de CHC<sup>208</sup>. Le CHC survient presque toujours dans le contexte de la fibrose hépatique, démontrant le rôle critique de l'environnement dans la carcinogénèse hépatique<sup>208</sup>. En effet, la sévérité de la fibrose s'est avérée être le plus important des facteurs prédictifs de survie et de risque de CHC<sup>209</sup>. Il n'existe actuellement aucun traitement antifibrotique validé et autorisé<sup>210,211</sup>.

Dans le cas du CHC avancé, les options thérapeutiques ne sont que peu satisfaisantes. Seuls quelques traitements systémiques, d'efficacité et de sécurité très limitées, sont actuellement disponibles<sup>173,174</sup>. En effet, le bénéfice en termes de survie de ces thérapies par rapport au placebo est d'environ 3 mois et les patients présentent fréquemment des effets secondaires majeurs<sup>173,174</sup>.

Étant donné le rôle critique de la fibrose dans la progression de la maladie hépatique ainsi que le risque et la mortalité liés au CHC, pallier ce besoin médical non satisfait de nouvelles approches pour traiter la fibrose hépatique avancée et l'hépatocarcinogenèse reste de première importance. Le but de ces nouvelles thérapies étant l'amélioration et des résultats en termes de survie et de qualité de vie des patients.

L'absence de traitements préventifs pour la fibrose hépatique et le CHC tient au fait que les circuits cellulaires à l'origine de la fibrose et de l'hépatocarcinogenèse, ne sont encore que partiellement compris. La pathogenèse de la fibrose est considérée comme un processus multifactoriel impliquant une interaction complexe entre les hépatocytes et les cellules non parenchymateuses, notamment les cellules étoilées, les macrophages et les myofibroblastes présents dans le foie<sup>212,213</sup>. La même complexité est retrouvée dans le cas du CHC. Le microenvironnement du CHC est une structure dynamique de cellules tumorales au sein d'une matrice extracellulaire constituée d'un mélange complexe de cellules stromales et des protéines qu'elles sécrètent. Les interactions des cellules tumorales avec les cellules hépatiques sinusoïdales et extrasinusoïdales contribuent au développement de la tumeur. En effet, les cellules cancéreuses ne conduisent pas seules à la maladie mais le processus se fait en interaction étroite avec le stroma qui est généralement activé de manière inappropriée en réponse à une inflammation chronique. Il existe des preuves solides que les cellules stromales contribuent au développement et à la progression du cancer en délivrant des signaux de maintien de la prolifération cellulaire, d'inhibition l'apoptose cellulaire, d'induction de la transformation cellulaire

et de l'angiogenèse, de promotion de l'invasion cellulaire et des métastases, de reprogrammation du métabolisme énergétique et d'échappement à la destruction immunitaire<sup>44</sup>. Cet écosystème hétérogène complexe, renforcé par la variation clonale des cellules cancéreuses, façonne et détermine la réponse au traitement des tumeurs<sup>214,215</sup>. À l'ère de l'immunothérapie, comprendre la biologie du microenvironnement tumoral et cibler les cellules stromales constitue une stratégie thérapeutique rationnelle. Malgré les preuves solides soutenant que le CHC est étroitement lié à un microenvironnement perturbé, on en sait peu sur les profils transcriptomiques à l'échelle de la cellule unique des composants du microenvironnement de la tumeur hépatique.

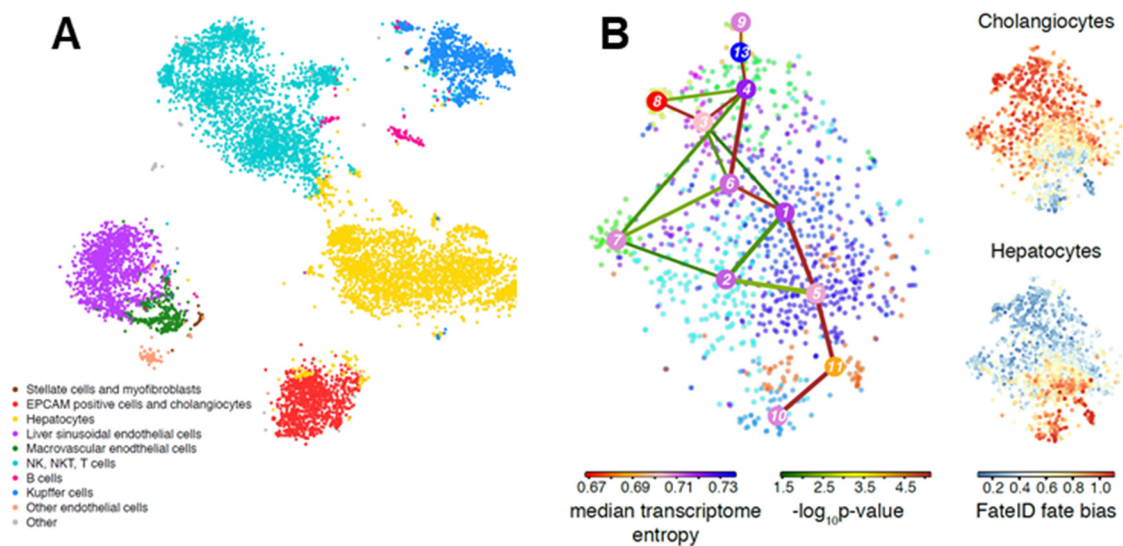
Par conséquent, les recherches portant sur la résolution cellulaire des mécanismes moléculaires sous-jacents à la progression de la maladie du foie et au développement du CHC auront un impact majeur sur la découverte de nouvelles cibles pour les stratégies de prévention du CHC impatientement attendues pour le traitement de la fibrose avancée et du CHC.

## 5.2 RESULTATS

### 5.2.1 UN ATLAS DES CELLULES DU FOIE HUMAIN REVELE SON HETEROGENEITE CELLULAIRE ET IDENTIFIE LES PROGENITEURS HEPATIQUES

Pour étudier précisément les processus et le rôle du microenvironnement dans la fibrogenèse et carcinogénèse hépatique, nous avons développé un pipeline de séquençage ARN sur cellule unique (scRNA-seq) à partir de tissus primaires de foie humain<sup>216</sup>. De cette manière nous avons assemblé le premier atlas de cellules du foie humain à partir de 10372 cellules provenant de six patients. Nous avons ainsi pu caractériser les principaux types de cellules hépatiques y compris des cellules progénitrices bipotentes capables de se différencier à la fois en cholangiocytes et en hépatocytes (**Fig. 27**).



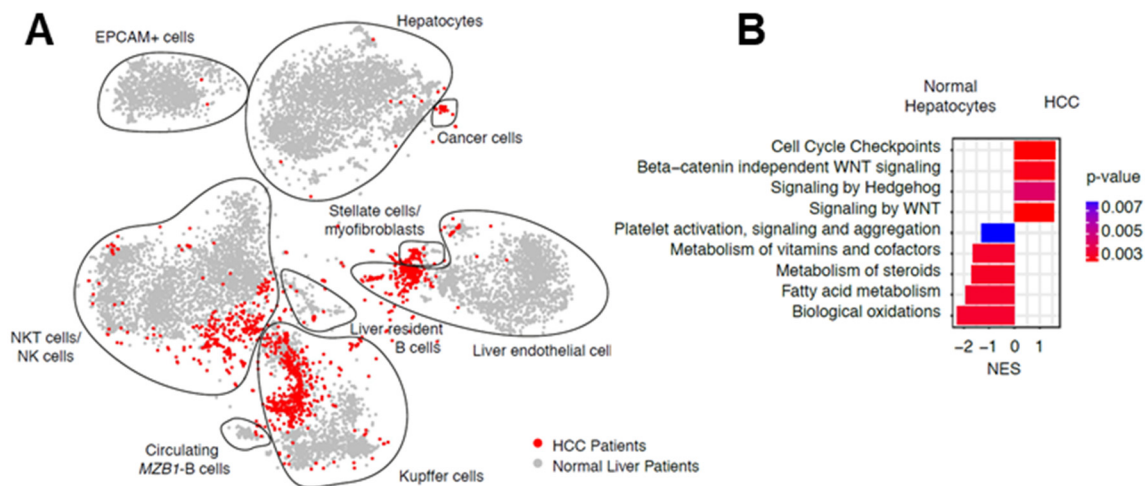


**Figure 27. L'analyse scRNA-seq identifie les différents types de cellules présents dans le foie humain adulte. A) Carte t-SNE de transcriptomes de cellule unique identifiant les principaux compartiments cellulaires hépatiques à partir de tissu hépatique humain normal provenant de six donneurs différents. B) L'algorithme FATE-ID révèle et récapitule le profil de différenciation d'une nouvelle cellule progénitrice bipotente.**

En utilisant cet atlas, nous avons comparé le profil scRNA-seq du foie normal au profil de 6319 cellules de CHC de 3 patients (**Fig. 28**). L'analyse des différences d'expression génique et l'analyse GSEA (Gene Set Enrichment Analysis) ont révélé que les hépatocytes transformés provenant du CHC régulent négativement les gènes du métabolisme cellulaire comme ALB et PCK1 et perdent la signature métabolique des hépatocytes normaux.

Nous avons découvert que les hépatocytes transformés régulaient positivement les voies de signalisation WNT et Hedgehog, qui sont en revanche fortement exprimées dans les cellules EPCAM+, soulignant les similitudes entre les progéniteurs hépatiques normaux EPCAM+ et la population de cellules cancéreuses observée (**Fig. 28**). De manière inattendue, l'expression de

marqueurs retenus pan-exprimé dans le CHC, tels que S100A9 et NTS, est limitée à une sous-population de cellules cancéreuses. De même, le marqueur pan-CHC GPC3 n'était pas exprimé dans tous les hépatocytes transformés, mais régulé positivement dans les cellules étoilées tumorales.



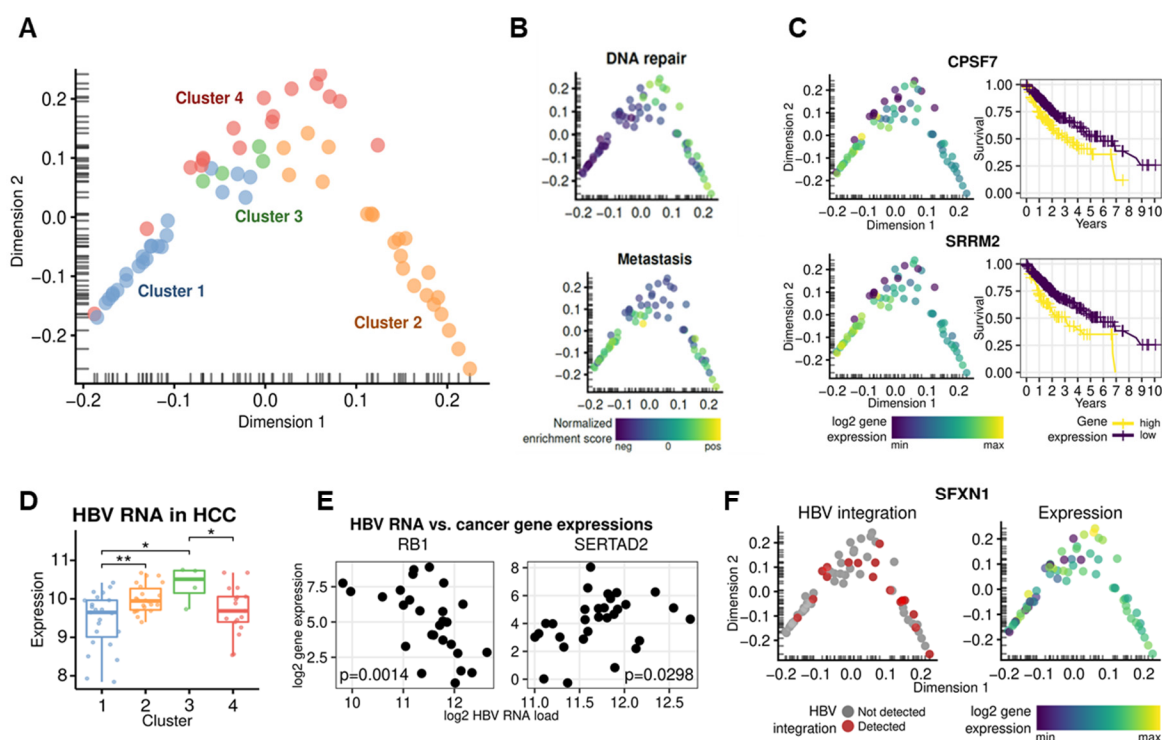
**Figure 28. L'analyse scRNA-seq de cellules provenant de CHC révèle les ou des signatures de gènes spécifiques du cancer et des phénotypes cellulaires perturbés.** A) Carte t-SNE montrant les clusters détectés par l'algorithme RaceID3 pour les cellules hépatiques normales et les cellules CHC provenant de trois patients. B) Analyse GSEA (Gene Set Enrichment Analysis) pour les groupes de gènes exprimés différemment entre les hépatocytes normaux et les cellules cancéreuses du CHC. Le bar-graphe montre le score d'enrichissement normalisé (NES) et met en évidence la valeur p.

Les cellules endothéliales de la tumeur, comparées aux cellules endothéliales normales, régulent positivement l'expression de gènes d'organisation de la matrice extracellulaire tels que COL4A1, COL4A2 et SPARC, ainsi que des gènes liés à des facteurs de croissance comme IGFBP7. Les populations de cellules immunitaires de la tumeur, comprenant les cellules de Küpffer, les cellules

NKT et NK, régulent positivement les gènes de réponse au stress et perdent leurs fonctions immunorégulatrices présentes chez leurs homologues hépatiques normaux. Dans le cadre de ces travaux, accepté pour publication en *Nature*, j'ai contribué en particulier en gérant l'approvisionnement des tissus hépatiques, en isolant les cellules à l'échelle de la cellule unique et en réalisant de nombreuses expériences de validation des hypothèses élaborées par l'analyse bio-informatique (expériences animales et immunofluorescence sur coupes de foie humain).

### 5.2.3 COMPARTIMENTATION VIRALE ET HETEROGENEITE CANCEREUSE DANS LE CHC INDUITS PAR LE VIRUS DE L'HEPATITE B

Dans le but d'étudier les interactions VHB-hôte et la carcinogenèse virale du CHC, nous avons effectué du scRNA-seq à partir de cellules d'un CHC associé à une infection par le virus de l'hépatite B (VHB). En utilisant la méthode Smart-Seq2 (Smart - Switching Mechanism At the end of the 5'-end of the RNA Transcript), nous avons pu séquencer en détail l'expression génique complète dans un nombre suffisant de cellules provenant de la tumeur. Nous avons démontré que l'approche scRNA-seq permet de mettre en évidence l'hétérogénéité cellulaire des tumeurs du foie et de caractériser les interactions VHB-hôte. Il est important de noter que différents clusters de cellules tumorales expriment des gènes associés à différents pronostics à long terme. De même, les sites d'intégration du VHB peuvent être identifiés au niveau de la cellule unique et permettent de caractériser l'évolution de la clonalité tumorale et être corrélés à l'expression génique cellulaire (**Fig. 29, Juehling F\*, Verrier E\*, Saviano A\* et al. JCI 2019 en révision, \*co-premier auteur**).



**Figure 29. Hétérogénéité clonale et compartimentation virale d'un CHC associé au VHB.** A) Regroupement de cellules CHC par SC3 montrant 4 clusters différents. B) Analyse GSEA montrant un enrichissement en gènes différent dans différents clusters. C) Nouveaux marqueurs pronostiques exprimés de manière hétérogène dans le CHC et validés par The Cancer Genome Atlas (TCGA). D) Compartimentation avec des différences significatives de charge d'ARN VHB dans les clusters de CHC. E) Charge d'ARN VHB par rapport à l'expression génique de RB1 (suppresseur de tumeur) et de SERTAD2 (oncogène) avec les valeurs  $p$  pour les corrélations de Spearman. F) Intégrations détectées du VHB dans SFXN1 et expression du gène correspondant.

## 5.3 DISCUSSION

L'atlas de cellules hépatiques humaines, établi en collaboration avec le groupe de D Grün à Freiburg, a révélé une hétérogénéité au sein des principales populations de cellules hépatiques, l'existence d'un progéniteur épithélial dans le foie humain adulte et la zonation transcriptomique des hépatocytes et des cellules endothéliales. L'atlas constitue une référence essentielle pour l'étude des maladies du foie et contribue à la progression de connaissances concernant les modèles et la pathogenèse des maladies du foie<sup>216</sup>.

L'analyse du patient VHB-CHC a démontré que malgré la faible réplication du VHB, la présence d'ARN du VHB était associée à une perturbation de l'expression des gènes impliqués dans la carcinogenèse soulignant l'importance de guérir cette infection chronique chez les patients infectés. L'hétérogénéité marquée de la tumeur met en évidence les défis posés par cette maladie : une thérapie combinée ciblant l'expression de différents gènes promoteurs de cancer sera probablement nécessaire pour le traitement du CHC chez ce patient en cas de récurrence non opérable ou de maladie avancée. L'analyse des interactions entre le VHB et les cellules hôtes au niveau de la cellule unique a révélé des voies jusque-là inconnues et des facteurs moteurs de la cancérogenèse (**Juehling F\*, Verrier E\*, Saviano A\* et al. JCI 2019 in révision, \*co-premier auteur**).

## 6. REFERENCES

- 1 Mokdad, A. H. A. *et al.* Liver cirrhosis mortality in 187 countries between 1980 and 2010: a systematic analysis. *BMC medicine* **12**, 145-145, doi:10.1186/s12916-014-0145-y (2014).
- 2 Hamdane, N. *et al.* HCV-induced Epigenetic Changes Associated With Liver Cancer Risk Persist After Sustained Virologic Response. *Gastroenterology*, doi:10.1053/j.gastro.2019.02.038 (2019).
- 3 Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1736-1788, doi:10.1016/S0140-6736(18)32203-7 (2018).
- 4 International Agency for Research on Cancer -World Health, O.
- 5 Asrani, S. K., Devarbavi, H., Eaton, J. & Kamath, P. S. Burden of liver diseases in the world. *J Hepatol* **70**, 151-171, doi:10.1016/j.jhep.2018.09.014 (2019).
- 6 Stepanova, M. *et al.* Direct and Indirect Economic Burden of Chronic Liver Disease in the United States. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association* **15**, 759-766 e755, doi:10.1016/j.cgh.2016.07.020 (2017).
- 7 Lozano, R. *et al.* Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2095-2128, doi:10.1016/S0140-6736(12)61728-0 (2012).
- 8 Lim, Y.-S. *et al.* Mortality, liver transplantation, and hepatocellular carcinoma among patients with chronic hepatitis B treated with entecavir vs lamivudine. *Gastroenterology* **147**, 152-161, doi:10.1053/j.gastro.2014.02.033 (2014).
- 9 Goldberg, D. *et al.* Changes in the Prevalence of Hepatitis C Virus Infection, Nonalcoholic Steatohepatitis, and Alcoholic Liver Disease Among Patients With Cirrhosis or Liver Failure on the Waitlist for Liver Transplantation. *Gastroenterology* **152**, 1090-1099.e1091, doi:10.1053/j.gastro.2017.01.003 (2017).
- 10 Akuta, N. *et al.* Liver Fibrosis and Body Mass Index Predict Hepatocarcinogenesis following Eradication of Hepatitis C Virus RNA by Direct-Acting Antivirals. *Oncology* **91**, 341-347, doi:10.1159/000450551 (2016).
- 11 Calleja, J. L. *et al.* Effectiveness, safety and clinical outcomes of direct-acting antiviral therapy in HCV genotype 1 infection: Results from a Spanish real-world cohort. *J Hepatol* **66**, 1138-1148, doi:10.1016/j.jhep.2017.01.028 (2017).
- 12 Kanwal, F. *et al.* Risk of Hepatocellular Cancer in HCV Patients Treated With Direct-Acting Antiviral Agents. *Gastroenterology* **153**, 996-1005.e1001, doi:10.1053/j.gastro.2017.06.012 (2017).
- 13 Romano, A. *et al.* Newly diagnosed hepatocellular carcinoma in patients with advanced hepatitis C treated with DAAs: A prospective population study. *J Hepatol* **69**, 345-352, doi:10.1016/j.jhep.2018.03.009 (2018).

- 14 van der Meer, A. J. *et al.* Risk of cirrhosis-related complications in patients with advanced fibrosis following hepatitis C virus eradication. *J Hepatol* **66**, 485-493, doi:10.1016/j.jhep.2016.10.017 (2017).
- 15 Llovet, J. M. & Villanueva, A. Liver cancer: Effect of HCV clearance with direct-acting antiviral agents on HCC. *Nature reviews. Gastroenterology & hepatology* **13**, 561-562, doi:10.1038/nrgastro.2016.140 (2016).
- 16 Conti, F. *et al.* Early occurrence and recurrence of hepatocellular carcinoma in HCV-related cirrhosis treated with direct-acting antivirals. *J Hepatol* **65**, 727-733, doi:10.1016/j.jhep.2016.06.015 (2016).
- 17 Reig, M. *et al.* Unexpected high rate of early tumor recurrence in patients with HCV-related HCC undergoing interferon-free therapy. *J Hepatol* **65**, 719-726, doi:10.1016/j.jhep.2016.04.008 (2016).
- 18 Ioannou, G. N., Green, P., Lowy, E., Mun, E. J. & Berry, K. Differences in hepatocellular carcinoma risk, predictors and trends over time according to etiology of cirrhosis. *PloS one* **13**, e0204412-e0204412, doi:10.1371/journal.pone.0204412 (2018).
- 19 Kim, D. *et al.* Changing Trends in Etiology- and Ethnicity-Based Annual Mortality Rates of Cirrhosis and Hepatocellular Carcinoma in the United States. *Hepatology (Baltimore, Md.)*, doi:10.1002/hep.30161 (2018).
- 20 Saviano, A. & Baumert, T. F. Mortality from liver cirrhosis and HCC in the DAA era: success in viral control is darkened by raise of metabolic disease. *Hepatobiliary Surg Nutr* **8**, 307-310, doi:10.21037/hbsn.2019.01.21 (2019).
- 21 Jenne, C. N. & Kubes, P. Immune surveillance by the liver. *Nat Immunol* **14**, 996-1006, doi:10.1038/ni.2691 (2013).
- 22 Marrone, G., Shah, V. H. & Gracia-Sancho, J. Sinusoidal communication in liver fibrosis and regeneration. *J Hepatol* **65**, 608-617, doi:10.1016/j.jhep.2016.04.018 (2016).
- 23 Pellicoro, A., Ramachandran, P., Iredale, J. P. & Fallowfield, J. A. Liver fibrosis and repair: immune regulation of wound healing in a solid organ. *Nat Rev Immunol* **14**, 181-194, doi:10.1038/nri3623 (2014).
- 24 Mossanen, J. C. *et al.* Chemokine (C-C motif) receptor 2-positive monocytes aggravate the early phase of acetaminophen-induced acute liver injury. *Hepatology* **64**, 1667-1682, doi:10.1002/hep.28682 (2016).
- 25 Lemmers, A. *et al.* The interleukin-17 pathway is involved in human alcoholic liver disease. *Hepatology* **49**, 646-657, doi:10.1002/hep.22680 (2009).
- 26 Meng, F. *et al.* Interleukin-17 signaling in inflammatory, Kupffer cells, and hepatic stellate cells exacerbates liver fibrosis in mice. *Gastroenterology* **143**, 765-776 e763, doi:10.1053/j.gastro.2012.05.049 (2012).
- 27 Ringelhan, M., Pfister, D., O'Connor, T., Pikarsky, E. & Heikenwalder, M. The immunology of hepatocellular carcinoma. *Nat Immunol* **19**, 222-232, doi:10.1038/s41590-018-0044-z (2018).
- 28 Tu, T. *et al.* Hepatocytes in liver injury: Victim, bystander, or accomplice in progressive fibrosis? *Journal of gastroenterology and hepatology* **30**, 1696-1704, doi:10.1111/jgh.13065 (2015).



- 29 Mederacke, I. *et al.* Fate tracing reveals hepatic stellate cells as dominant contributors to liver fibrosis independent of its aetiology. *Nature communications* **4**, 2823, doi:10.1038/ncomms3823 (2013).
- 30 Maeda, S., Kamata, H., Luo, J. L., Leffert, H. & Karin, M. IKKbeta couples hepatocyte death to cytokine-driven compensatory proliferation that promotes chemical hepatocarcinogenesis. *Cell* **121**, 977-990, doi:10.1016/j.cell.2005.04.014 (2005).
- 31 McCartney, E. M. *et al.* Signal transducer and activator of transcription 3 is a proviral host factor for hepatitis C virus. *Hepatology* **58**, 1558-1568, doi:10.1002/hep.26496 (2013).
- 32 Kang, T. W. *et al.* Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. *Nature* **479**, 547-551, doi:10.1038/nature10599 (2011).
- 33 Ma, C. *et al.* NAFLD causes selective CD4(+) T lymphocyte loss and promotes hepatocarcinogenesis. *Nature* **531**, 253-257, doi:10.1038/nature16969 (2016).
- 34 Li, K., Liu, H. & Guo, T. Th17/Treg imbalance is an indicator of liver cirrhosis process and a risk factor for HCC occurrence in HBV patients. *Clinics and research in hepatology and gastroenterology* **41**, 399-407, doi:10.1016/j.clinre.2016.12.004 (2017).
- 35 Read, S., Malmstrom, V. & Powrie, F. Cytotoxic T lymphocyte-associated antigen 4 plays an essential role in the function of CD25(+)CD4(+) regulatory cells that control intestinal inflammation. *The Journal of experimental medicine* **192**, 295-302, doi:10.1084/jem.192.2.295 (2000).
- 36 Carter, L. *et al.* PD-1:PD-L inhibitory pathway affects both CD4(+) and CD8(+) T cells and is overcome by IL-2. *European journal of immunology* **32**, 634-643, doi:10.1002/1521-4141(200203)32:3<634::AID-IMMU634>3.0.CO;2-9 (2002).
- 37 Kehrl, J. H. *et al.* Production of transforming growth factor beta by human T lymphocytes and its potential role in the regulation of T cell growth. *The Journal of experimental medicine* **163**, 1037-1050, doi:10.1084/jem.163.5.1037 (1986).
- 38 Espevik, T., Waage, A., Faxvaag, A. & Shalaby, M. R. Regulation of interleukin-2 and interleukin-6 production from T-cells: involvement of interleukin-1 beta and transforming growth factor-beta. *Cellular immunology* **126**, 47-56, doi:10.1016/0008-8749(90)90299-7 (1990).
- 39 Smyth, M. J., Strobl, S. L., Young, H. A., Ortaldo, J. R. & Ochoa, A. C. Regulation of lymphokine-activated killer activity and pore-forming protein gene expression in human peripheral blood CD8+ T lymphocytes. Inhibition by transforming growth factor-beta. *J Immunol* **146**, 3289-3297 (1991).
- 40 Knolle, P. *et al.* Human Kupffer cells secrete IL-10 in response to lipopolysaccharide (LPS) challenge. *J Hepatol* **22**, 226-229, doi:10.1016/0168-8278(95)80433-1 (1995).
- 41 Hattori, E. *et al.* Possible contribution of circulating interleukin-10 (IL-10) to anti-tumor immunity and prognosis in patients with unresectable hepatocellular carcinoma. *Hepatol Res* **27**, 309-314 (2003).
- 42 El-Emshaty, H. M., Nasif, W. A. & Mohamed, I. E. Serum Cytokine of IL-10 and IL-12 in Chronic Liver Disease: The Immune and Inflammatory Response. *Disease markers* **2015**, 707254, doi:10.1155/2015/707254 (2015).

- 43 Nishida, N., Kitano, M., Sakurai, T. & Kudo, M. Molecular Mechanism and Prediction of Sorafenib Chemoresistance in Human Hepatocellular Carcinoma. *Dig Dis* **33**, 771-779, doi:10.1159/000439102 (2015).
- 44 Hanahan, D. & Coussens, L. M. Accessories to the Crime: Functions of Cells Recruited to the Tumor Microenvironment. *Cancer Cell* **21**, 309-322, doi:10.1016/j.ccr.2012.02.022 (2012).
- 45 Tovar, V. *et al.* Tumour initiating cells and IGF/FGF signalling contribute to sorafenib resistance in hepatocellular carcinoma. *Gut* **66**, 530-540, doi:10.1136/gutjnl-2015-309501 (2017).
- 46 Saviano, A. *et al.* in *Hepatocellular carcinoma* (ed Y. Hoshida) Ch. 15, (Springer, 2019).
- 47 White, D. L., Kanwal, F. & El-Serag, H. B. Association between nonalcoholic fatty liver disease and risk for hepatocellular cancer, based on systematic review. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association* **10**, 1342-1359 e1342, doi:10.1016/j.cgh.2012.10.001 (2012).
- 48 Piscaglia, F. *et al.* Clinical patterns of hepatocellular carcinoma in nonalcoholic fatty liver disease: A multicenter prospective study. *Hepatology* **63**, 827-838, doi:10.1002/hep.28368 (2016).
- 49 Tavakoli, H. *et al.* Cirrhosis Patients with Nonalcoholic Steatohepatitis Are Significantly Less Likely to Receive Surveillance for Hepatocellular Carcinoma. *Digestive diseases and sciences* **62**, 2174-2181, doi:10.1007/s10620-017-4595-x (2017).
- 50 Anstee, Q. M., Reeves, H. L., Kotsiliti, E., Govaere, O. & Heikenwalder, M. From NASH to HCC: current concepts and future challenges. *Nature reviews. Gastroenterology & hepatology* **16**, 411-428, doi:10.1038/s41575-019-0145-7 (2019).
- 51 Yuan, D. *et al.* Kupffer Cell-Derived Tnf Triggers Cholangiocellular Tumorigenesis through JNK due to Chronic Mitochondrial Dysfunction and ROS. *Cancer Cell* **31**, 771-789 e776, doi:10.1016/j.ccell.2017.05.006 (2017).
- 52 Singh, R. *et al.* Autophagy regulates lipid metabolism. *Nature* **458**, 1131-1135, doi:10.1038/nature07976 (2009).
- 53 Tanaka, S. *et al.* Rubicon inhibits autophagy and accelerates hepatocyte apoptosis and lipid accumulation in nonalcoholic fatty liver disease in mice. *Hepatology* **64**, 1994-2014, doi:10.1002/hep.28820 (2016).
- 54 Gabele, E. *et al.* DSS induced colitis increases portal LPS levels and enhances hepatic inflammation and fibrogenesis in experimental NASH. *J Hepatol* **55**, 1391-1399, doi:10.1016/j.jhep.2011.02.035 (2011).
- 55 Wolf, M. J. *et al.* Metabolic activation of intrahepatic CD8+ T cells and NKT cells causes nonalcoholic steatohepatitis and liver cancer via cross-talk with hepatocytes. *Cancer Cell* **26**, 549-564, doi:10.1016/j.ccell.2014.09.003 (2014).
- 56 Laurent-Puig, P. *et al.* Genetic alterations associated with hepatocellular carcinomas define distinct pathways of hepatocarcinogenesis. *Gastroenterology* **120**, 1763-1773, doi:10.1053/gast.2001.24798 (2001).

- 57 Guichard, C. *et al.* Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet* **44**, 694-698, doi:10.1038/ng.2256 (2012).
- 58 Shen, S. *et al.* Biomarker MicroRNAs for Diagnosis, Prognosis and Treatment of Hepatocellular Carcinoma: A Functional Survey and Comparison. *Scientific reports* **6**, 38311, doi:10.1038/srep38311 (2016).
- 59 Seeger, C. & Mason, W. S. Molecular biology of hepatitis B virus infection. *Virology* **479-480**, 672-686, doi:10.1016/j.virol.2015.02.031 (2015).
- 60 Bonilla Guerrero, R. & Roberts, L. R. The role of hepatitis B virus integrations in the pathogenesis of human hepatocellular carcinoma. *J Hepatol* **42**, 760-777, doi:10.1016/j.jhep.2005.02.005 (2005).
- 61 Bouchard, M. J. & Navas-Martin, S. Hepatitis B and C virus hepatocarcinogenesis: lessons learned and future challenges. *Cancer letters* **305**, 123-143, doi:10.1016/j.canlet.2010.11.014 (2011).
- 62 Tu, T., Budzinska, M. A., Vondran, F. W. R., Shackel, N. A. & Urban, S. Hepatitis B Virus DNA Integration Occurs Early in the Viral Life Cycle in an In Vitro Infection Model via Sodium Taurocholate Cotransporting Polypeptide-Dependent Uptake of Enveloped Virus Particles. *J Virol* **92**, doi:10.1128/JVI.02007-17 (2018).
- 63 Feitelson, M. A. & Lee, J. Hepatitis B virus integration, fragile sites, and hepatocarcinogenesis. *Cancer letters* **252**, 157-170, doi:10.1016/j.canlet.2006.11.010 (2007).
- 64 Ferber, M. J. *et al.* Integrations of the hepatitis B virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers. *Oncogene* **22**, 3813-3820, doi:10.1038/sj.onc.1206528 (2003).
- 65 Gozuacik, D. *et al.* Identification of human cancer-related genes by naturally occurring Hepatitis B Virus DNA tagging. *Oncogene* **20**, 6233-6240, doi:10.1038/sj.onc.1204835 (2001).
- 66 Murakami, Y. *et al.* Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas. *Gut* **54**, 1162-1168, doi:10.1136/gut.2004.054452 (2005).
- 67 Paterlini-Brechot, P. *et al.* Hepatitis B virus-related insertional mutagenesis occurs frequently in human liver cancers and recurrently targets human telomerase gene. *Oncogene* **22**, 3911-3916, doi:10.1038/sj.onc.1206492 (2003).
- 68 Sung, W. K. *et al.* Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nat Genet* **44**, 765-769, doi:10.1038/ng.2295 (2012).
- 69 Lau, C. C. *et al.* Viral-human chimeric transcript predisposes risk to liver cancer development and progression. *Cancer Cell* **25**, 335-349, doi:10.1016/j.ccr.2014.01.030 (2014).
- 70 Liang, H. W. *et al.* Hepatitis B virus-human chimeric transcript HBx-LINE1 promotes hepatic injury via sequestering cellular microRNA-122. *J Hepatol* **64**, 278-291, doi:10.1016/j.jhep.2015.09.013 (2016).

- 71 Amaddeo, G. *et al.* Integration of tumour and viral genomic characterizations in HBV-related hepatocellular carcinomas. *Gut* **64**, 820-829, doi:10.1136/gutjnl-2013-306228 (2015).
- 72 Zhao, L. H. *et al.* Genomic and oncogenic preference of HBV integration in hepatocellular carcinoma. *Nature communications* **7**, 12992, doi:10.1038/ncomms12992 (2016).
- 73 Kekule, A. S. *et al.* The preS2/S region of integrated hepatitis B virus DNA encodes a transcriptional transactivator. *Nature* **343**, 457-461, doi:10.1038/343457a0 (1990).
- 74 Ma, N. F. *et al.* COOH-terminal truncated HBV X protein plays key role in hepatocarcinogenesis. *Clin Cancer Res* **14**, 5061-5068, doi:10.1158/1078-0432.CCR-07-5082 (2008).
- 75 Chisari, F. V. *et al.* A transgenic mouse model of the chronic hepatitis B surface antigen carrier state. *Science* **230**, 1157-1160, doi:10.1126/science.3865369 (1985).
- 76 Locarnini, S. & Zoulim, F. Molecular genetics of HBV infection. *Antiviral therapy* **15 Suppl 3**, 3-14, doi:10.3851/IMP1619 (2010).
- 77 Chen, C. H. *et al.* Pre-S deletion and complex mutations of hepatitis B virus related to advanced liver disease in HBeAg-negative patients. *Gastroenterology* **133**, 1466-1474, doi:10.1053/j.gastro.2007.09.002 (2007).
- 78 Keasler, V. V., Hodgson, A. J., Madden, C. R. & Slagle, B. L. Enhancement of hepatitis B virus replication by the regulatory X protein in vitro and in vivo. *J Virol* **81**, 2656-2662, doi:10.1128/JVI.02020-06 (2007).
- 79 Tateno, C. *et al.* Near completely humanized liver in mice shows human-type metabolic responses to drugs. *Am J Pathol* **165**, 901-912, doi:10.1016/S0002-9440(10)63352-4 (2004).
- 80 Bouchard, M. J. & Schneider, R. J. The enigmatic X gene of hepatitis B virus. *J Virol* **78**, 12725-12734, doi:10.1128/JVI.78.23.12725-12734.2004 (2004).
- 81 Tsuge, M. *et al.* Infection of human hepatocyte chimeric mouse with genetically engineered hepatitis B virus. *Hepatology* **42**, 1046-1054, doi:10.1002/hep.20892 (2005).
- 82 Leupin, O., Bontron, S., Schaeffer, C. & Strubin, M. Hepatitis B virus X protein stimulates viral genome replication via a DDB1-dependent pathway distinct from that leading to cell death. *J Virol* **79**, 4238-4245, doi:10.1128/JVI.79.7.4238-4245.2005 (2005).
- 83 Sitterlin, D., Bergametti, F., Tiollais, P., Tennant, B. C. & Transy, C. Correct binding of viral X protein to UVDDDB-p127 cellular protein is critical for efficient infection by hepatitis B viruses. *Oncogene* **19**, 4427-4431, doi:10.1038/sj.onc.1203770 (2000).
- 84 Tarn, C., Zou, L., Hullinger, R. L. & Andrisani, O. M. Hepatitis B virus X protein activates the p38 mitogen-activated protein kinase pathway in dedifferentiated hepatocytes. *J Virol* **76**, 9763-9772, doi:10.1128/jvi.76.19.9763-9772.2002 (2002).
- 85 Chami, M., Ferrari, D., Nicotera, P., Paterlini-Brechot, P. & Rizzuto, R. Caspase-dependent alterations of Ca<sup>2+</sup> signaling in the induction of apoptosis by hepatitis B virus X protein. *The Journal of biological chemistry* **278**, 31745-31755, doi:10.1074/jbc.M304202200 (2003).

- 86 Oh, J. C., Jeong, D. L., Kim, I. K. & Oh, S. H. Activation of calcium signaling by hepatitis B virus-X protein in liver cells. *Experimental & molecular medicine* **35**, 301-309, doi:10.1038/emm.2003.41 (2003).
- 87 Yoo, Y. D. *et al.* Regulation of transforming growth factor-beta 1 expression by the hepatitis B virus (HBV) X transactivator. Role in HBV pathogenesis. *The Journal of clinical investigation* **97**, 388-395, doi:10.1172/JCI118427 (1996).
- 88 Xu, J., Lamouille, S. & Derynck, R. TGF-beta-induced epithelial to mesenchymal transition. *Cell research* **19**, 156-172, doi:10.1038/cr.2009.5 (2009).
- 89 Clippinger, A. J., Gearhart, T. L. & Bouchard, M. J. Hepatitis B virus X protein modulates apoptosis in primary rat hepatocytes by regulating both NF-kappaB and the mitochondrial permeability transition pore. *J Virol* **83**, 4718-4731, doi:10.1128/JVI.02590-08 (2009).
- 90 Karin, M., Lawrence, T. & Nizet, V. Innate immunity gone awry: linking microbial infections to chronic inflammation and cancer. *Cell* **124**, 823-835, doi:10.1016/j.cell.2006.02.016 (2006).
- 91 Lee, S. *et al.* Hepatitis B virus X protein differentially regulates cell cycle progression in X-transforming versus nontransforming hepatocyte (AML12) cell lines. *The Journal of biological chemistry* **277**, 8730-8740, doi:10.1074/jbc.M108025200 (2002).
- 92 Yoo, Y. G. *et al.* Hepatitis B virus X protein enhances transcriptional activity of hypoxia-inducible factor-1alpha through activation of mitogen-activated protein kinase pathway. *The Journal of biological chemistry* **278**, 39076-39084, doi:10.1074/jbc.M305101200 (2003).
- 93 Sanz-Cameno, P. *et al.* Hepatitis B virus promotes angiopoietin-2 expression in liver tissue: role of HBV x protein. *Am J Pathol* **169**, 1215-1222, doi:10.2353/ajpath.2006.051246 (2006).
- 94 Yamashita, T., Budhu, A., Forgues, M. & Wang, X. W. Activation of hepatic stem cell marker EpCAM by Wnt-beta-catenin signaling in hepatocellular carcinoma. *Cancer Res* **67**, 10831-10839, doi:10.1158/0008-5472.CAN-07-0908 (2007).
- 95 Madden, C. R., Finegold, M. J. & Slagle, B. L. Hepatitis B virus X protein acts as a tumor promoter in development of diethylnitrosamine-induced preneoplastic lesions. *J Virol* **75**, 3851-3858, doi:10.1128/JVI.75.8.3851-3858.2001 (2001).
- 96 Liu, Z. M. *et al.* Hepatitis B virus infection contributes to oxidative stress in a population exposed to aflatoxin B1 and high-risk for hepatocellular carcinoma. *Cancer letters* **263**, 212-222, doi:10.1016/j.canlet.2008.01.006 (2008).
- 97 Boyault, S. *et al.* Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology* **45**, 42-52, doi:10.1002/hep.21467 (2007).
- 98 Hoshida, Y. *et al.* Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res* **69**, 7385-7392, doi:10.1158/0008-5472.CAN-09-1089 (2009).
- 99 Goossens, N., Sun, X. & Hoshida, Y. Molecular classification of hepatocellular carcinoma: potential therapeutic implications. *Hepatic oncology* **2**, 371-379, doi:10.2217/hep.15.26 (2015).

- 100 Zhou, L. *et al.* Edmondson-Steiner grade: A crucial predictor of recurrence and survival in hepatocellular carcinoma without microvascular invasion. *Pathology, research and practice* **213**, 824-830, doi:10.1016/j.prp.2017.03.002 (2017).
- 101 Calderaro, J., Ziol, M., Paradis, V. & Zucman-Rossi, J. Molecular and histological correlations in liver cancer. *J Hepatol* **71**, 616-630, doi:10.1016/j.jhep.2019.06.001 (2019).
- 102 Calderaro, J. *et al.* Histological subtypes of hepatocellular carcinoma are related to gene mutations and molecular tumour classification. *J Hepatol* **67**, 727-738, doi:10.1016/j.jhep.2017.05.014 (2017).
- 103 Tan, P. S. *et al.* Clinicopathological indices to predict hepatocellular carcinoma molecular classification. *Liver international : official journal of the International Association for the Study of the Liver* **36**, 108-118, doi:10.1111/liv.12889 (2016).
- 104 Audard, V. *et al.* Cholestasis is a marker for hepatocellular carcinomas displaying beta-catenin mutations. *The Journal of pathology* **212**, 345-352, doi:10.1002/path.2169 (2007).
- 105 Ziol, M. *et al.* Macrotrabecular-massive hepatocellular carcinoma: A distinctive histological subtype with clinical relevance. *Hepatology* **68**, 103-112, doi:10.1002/hep.29762 (2018).
- 106 Salomao, M., Yu, W. M., Brown, R. S., Jr., Emond, J. C. & Lefkowitz, J. H. Steatohepatic hepatocellular carcinoma (SH-HCC): a distinctive histological variant of HCC in hepatitis C virus-related cirrhosis with associated NAFLD/NASH. *The American journal of surgical pathology* **34**, 1630-1636, doi:10.1097/PAS.0b013e3181f31caa (2010).
- 107 Lee, J. S. *et al.* Tumor stroma with senescence-associated secretory phenotype in steatohepatic hepatocellular carcinoma. *PloS one* **12**, e0171922, doi:10.1371/journal.pone.0171922 (2017).
- 108 Seok, J. Y. *et al.* A fibrous stromal component in hepatocellular carcinoma reveals a cholangiocarcinoma-like gene expression trait and epithelial-mesenchymal transition. *Hepatology* **55**, 1776-1786, doi:10.1002/hep.25570 (2012).
- 109 Durnez, A. *et al.* The clinicopathological and prognostic relevance of cytokeratin 7 and 19 expression in hepatocellular carcinoma. A possible progenitor cell origin. *Histopathology* **49**, 138-151, doi:10.1111/j.1365-2559.2006.02468.x (2006).
- 110 Lee, J. S. *et al.* A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nature medicine* **12**, 410-416, doi:10.1038/nm1377 (2006).
- 111 Engelholm, L. H. *et al.* CRISPR/Cas9 Engineering of Adult Mouse Liver Demonstrates That the Dnajb1-Prkaca Gene Fusion Is Sufficient to Induce Tumors Resembling Fibrolamellar Hepatocellular Carcinoma. *Gastroenterology* **153**, 1662-1673 e1610, doi:10.1053/j.gastro.2017.09.008 (2017).
- 112 Kasthuber, E. R. *et al.* DNAJB1-PRKACA fusion kinase interacts with beta-catenin and the liver regenerative response to drive fibrolamellar hepatocellular carcinoma. *Proc Natl Acad Sci U S A* **114**, 13076-13084, doi:10.1073/pnas.1716483114 (2017).
- 113 Malouf, G. G. *et al.* Methylome sequencing for fibrolamellar hepatocellular carcinoma depicts distinctive features. *Epigenetics* **10**, 872-881, doi:10.1080/15592294.2015.1076955 (2015).

- 114 Malouf, G. G. *et al.* Transcriptional profiling of pure fibrolamellar hepatocellular carcinoma reveals an endocrine signature. *Hepatology* **59**, 2228-2237, doi:10.1002/hep.27018 (2014).
- 115 Malouf, G. G. *et al.* Pure and mixed fibrolamellar hepatocellular carcinomas differ in natural history and prognosis after complete surgical resection. *Cancer* **118**, 4981-4990, doi:10.1002/cncr.27520 (2012).
- 116 Oikawa, T. *et al.* Model of fibrolamellar hepatocellular carcinomas reveals striking enrichment in cancer stem cells. *Nature communications* **6**, 8070, doi:10.1038/ncomms9070 (2015).
- 117 Chan, A. W. *et al.* Lymphoepithelioma-like hepatocellular carcinoma: an uncommon variant of hepatocellular carcinoma with favorable outcome. *The American journal of surgical pathology* **39**, 304-312, doi:10.1097/PAS.0000000000000376 (2015).
- 118 Labgaa, I., Stueck, A. & Ward, S. C. Lymphoepithelioma-Like Carcinoma in Liver. *Am J Pathol* **187**, 1438-1444, doi:10.1016/j.ajpath.2017.02.022 (2017).
- 119 Calderaro, J. *et al.* Programmed death ligand 1 expression in hepatocellular carcinoma: Relationship With clinical and pathological features. *Hepatology* **64**, 2038-2046, doi:10.1002/hep.28710 (2016).
- 120 Moeini, A. *et al.* Mixed hepatocellular cholangiocarcinoma tumors: Cholangiolocellular carcinoma is a distinct molecular entity. *J Hepatol* **66**, 952-961, doi:10.1016/j.jhep.2017.01.010 (2017).
- 121 Wood, L. D. *et al.* Chromophobe hepatocellular carcinoma with abrupt anaplasia: a proposal for a new subtype of hepatocellular carcinoma with unique morphological and molecular features. *Mod Pathol* **26**, 1586-1593, doi:10.1038/modpathol.2013.68 (2013).
- 122 Heaphy, C. M. *et al.* Altered telomeres in tumors with ATRX and DAXX mutations. *Science* **333**, 425, doi:10.1126/science.1207313 (2011).
- 123 Kurebayashi, Y. *et al.* Landscape of immune microenvironment in hepatocellular carcinoma and its additional impact on histological and molecular classification. *Hepatology* **68**, 1025-1041, doi:10.1002/hep.29904 (2018).
- 124 Calderaro, J. *et al.* Intra-tumoral tertiary lymphoid structures are associated with a low risk of early recurrence of hepatocellular carcinoma. *J Hepatol* **70**, 58-65, doi:10.1016/j.jhep.2018.09.003 (2019).
- 125 Liu, C. Q. *et al.* Expression patterns of programmed death ligand 1 correlate with different microenvironments and patient prognosis in hepatocellular carcinoma. *Br J Cancer* **119**, 80-88, doi:10.1038/s41416-018-0144-4 (2018).
- 126 Nault, J. C. *et al.* A hepatocellular carcinoma 5-gene score associated with survival of patients after liver resection. *Gastroenterology* **145**, 176-187, doi:10.1053/j.gastro.2013.03.051 (2013).
- 127 Nault, J. C. *et al.* Clinical Impact of Genomic Diversity From Early to Advanced Hepatocellular Carcinoma. *Hepatology*, doi:10.1002/hep.30811 (2019).
- 128 Grau-Bove, X., Ruiz-Trillo, I. & Rodriguez-Pascual, F. Origin and evolution of lysyl oxidases. *Scientific reports* **5**, 10568, doi:10.1038/srep10568 (2015).

- 129 Loomba, R. *et al.* The ASK1 inhibitor selonsertib in patients with nonalcoholic steatohepatitis: A randomized, phase 2 trial. *Hepatology* **67**, 549-559, doi:10.1002/hep.29514 (2018).
- 130 Gilead Science, I. *Gilead Announces Topline Data From Phase 3 STELLAR-4 Study of Selonsertib in Compensated Cirrhosis (F4) Due to Nonalcoholic Steatohepatitis (NASH)*, <<https://www.gilead.com/news-and-press/press-room/press-releases/2019/2/gilead-announces-topline-data-from-phase-3-stellar4-study-of-selonsertib-in-compensated-cirrhosis-f4-due-to-nonalcoholic-steatohepatitis-nash>> (2019).
- 131 Harrison, S. A. *et al.* Simtuzumab Is Ineffective for Patients With Bridging Fibrosis or Compensated Cirrhosis Caused by Nonalcoholic Steatohepatitis. *Gastroenterology* **155**, 1140-1153, doi:10.1053/j.gastro.2018.07.006 (2018).
- 132 Muir, A. J. *et al.* Simtuzumab for Primary Sclerosing Cholangitis: Phase 2 Study Results With Insights on the Natural History of the Disease. *Hepatology* **69**, 684-698, doi:10.1002/hep.30237 (2019).
- 133 Friedman, S. L. *et al.* A randomized, placebo-controlled trial of cenicriviroc for treatment of nonalcoholic steatohepatitis with fibrosis. *Hepatology* **67**, 1754-1767, doi:10.1002/hep.29477 (2018).
- 134 Gross, B., Pawlak, M., Lefebvre, P. & Staels, B. PPARs in obesity-induced T2DM, dyslipidaemia and NAFLD. *Nature reviews. Endocrinology* **13**, 36-49, doi:10.1038/nrendo.2016.135 (2017).
- 135 Ratziu, V. *et al.* Elafibranor, an Agonist of the Peroxisome Proliferator-Activated Receptor- $\alpha$  and - $\delta$ , Induces Resolution of Nonalcoholic Steatohepatitis Without Fibrosis Worsening. *Gastroenterology* **150**, 1147-1159 e1145, doi:10.1053/j.gastro.2016.01.038 (2016).
- 136 Luketic, V. in *International Liver Congress, ILC 2019* (ed liver European Association for the study of the) (Vienna, 2019).
- 137 Sinal, C. J. *et al.* Targeted disruption of the nuclear receptor FXR/BAR impairs bile acid and lipid homeostasis. *Cell* **102**, 731-744, doi:10.1016/s0092-8674(00)00062-3 (2000).
- 138 Seyer, P. *et al.* Hepatic glucose sensing is required to preserve beta cell glucose competence. *The Journal of clinical investigation* **123**, 1662-1676, doi:10.1172/JCI65538 (2013).
- 139 Neuschwander-Tetri, B. A. *et al.* Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. *Lancet* **385**, 956-965, doi:10.1016/S0140-6736(14)61933-4 (2015).
- 140 Younossi, I. in *International Liver Congress, ILC 2019* (ed liver European Association for the study of the) (Vienna, 2019).
- 141 Su, T. H. *et al.* Four-year entecavir therapy reduces hepatocellular carcinoma, cirrhotic events and mortality in chronic hepatitis B patients. *Liver international : official journal of the International Association for the Study of the Liver* **36**, 1755-1764, doi:10.1111/liv.13253 (2016).
- 142 Ni, Y. H. *et al.* Minimization of hepatitis B infection by a 25-year universal vaccination program. *J Hepatol* **57**, 730-735, doi:10.1016/j.jhep.2012.05.021 (2012).



- 143 Fujiwara, N., Friedman, S. L., Goossens, N. & Hoshida, Y. Risk factors and prevention of hepatocellular carcinoma in the era of precision medicine. *J Hepatol* **68**, 526-549, doi:10.1016/j.jhep.2017.09.016 (2018).
- 144 Simon, T. G., King, L. Y., Zheng, H. & Chung, R. T. Statin use is associated with a reduced risk of fibrosis progression in chronic hepatitis C. *J Hepatol* **62**, 18-23, doi:10.1016/j.jhep.2014.08.013 (2015).
- 145 Chang, F. M. *et al.* Statins decrease the risk of decompensation in hepatitis B virus- and hepatitis C virus-related cirrhosis: A population-based study. *Hepatology* **66**, 896-907, doi:10.1002/hep.29172 (2017).
- 146 Dyson, J. *et al.* Hepatocellular cancer: the impact of obesity, type 2 diabetes and a multidisciplinary team. *J Hepatol* **60**, 110-117, doi:10.1016/j.jhep.2013.08.011 (2014).
- 147 Tseng, C. H. Metformin and risk of hepatocellular carcinoma in patients with type 2 diabetes. *Liver international : official journal of the International Association for the Study of the Liver* **38**, 2018-2027, doi:10.1111/liv.13872 (2018).
- 148 Zhou, Y. Y. *et al.* Systematic Review with Network Meta-Analysis: Antidiabetic Medication and Risk of Hepatocellular Carcinoma. *Scientific reports* **6**, 33743, doi:10.1038/srep33743 (2016).
- 149 Chen, H. *et al.* Hepatic cyclooxygenase-2 overexpression induced spontaneous hepatocellular carcinoma formation in mice. *Oncogene* **36**, 4415-4426, doi:10.1038/onc.2017.73 (2017).
- 150 Simon, T. G. *et al.* Association Between Aspirin Use and Risk of Hepatocellular Carcinoma. *JAMA oncology* **4**, 1683-1690, doi:10.1001/jamaoncol.2018.4154 (2018).
- 151 Funaki, M. *et al.* Peretinoin, an acyclic retinoid, inhibits hepatocarcinogenesis by suppressing sphingosine kinase 1 expression in vitro and in vivo. *Scientific reports* **7**, 16978, doi:10.1038/s41598-017-17285-2 (2017).
- 152 Malehmir, M. *et al.* Platelet GPIIb/IIIa is a mediator and potential interventional target for NASH and subsequent liver cancer. *Nature medicine* **25**, 641-655, doi:10.1038/s41591-019-0379-5 (2019).
- 153 Nakagawa, S. *et al.* Molecular Liver Cancer Prevention in Cirrhosis by Organ Transcriptome Analysis and Lysophosphatidic Acid Pathway Inhibition. *Cancer Cell* **30**, 879-890, doi:10.1016/j.ccell.2016.11.004 (2016).
- 154 Geissler, E. K. *et al.* Sirolimus Use in Liver Transplant Recipients With Hepatocellular Carcinoma: A Randomized, Multicenter, Open-Label Phase 3 Trial. *Transplantation* **100**, 116-125, doi:10.1097/TP.0000000000000965 (2016).
- 155 Erstad, D. J., Tager, A. M., Hoshida, Y. & Fuchs, B. C. The autotaxin-lysophosphatidic acid pathway emerges as a therapeutic target to prevent liver cancer. *Molecular & cellular oncology* **4**, e1311827, doi:10.1080/23723556.2017.1311827 (2017).
- 156 Oakley, F. *et al.* Angiotensin II activates I kappaB kinase phosphorylation of RelA at Ser 536 to promote myofibroblast survival and liver fibrosis. *Gastroenterology* **136**, 2334-2344 e2331, doi:10.1053/j.gastro.2009.02.081 (2009).

- 157 Moreno, M. *et al.* Reduction of advanced liver fibrosis by short-term targeted delivery of an angiotensin receptor blocker to hepatic stellate cells in rats. *Hepatology* **51**, 942-952, doi:10.1002/hep.23419 (2010).
- 158 Pelusi, S. *et al.* Renin-Angiotensin System Inhibitors, Type 2 Diabetes and Fibrosis Progression: An Observational Study in Patients with Nonalcoholic Fatty Liver Disease. *PloS one* **11**, e0163069, doi:10.1371/journal.pone.0163069 (2016).
- 159 Kim, G., Kim, J., Lim, Y. L., Kim, M. Y. & Baik, S. K. Renin-angiotensin system inhibitors and fibrosis in chronic liver disease: a systematic review. *Hepatol Int* **10**, 819-828, doi:10.1007/s12072-016-9705-x (2016).
- 160 Facciorusso, A. *et al.* Angiotensin receptor blockers improve survival outcomes after radiofrequency ablation in hepatocarcinoma patients. *Journal of gastroenterology and hepatology* **30**, 1643-1650, doi:10.1111/jgh.12988 (2015).
- 161 Yoshiji, H. *et al.* Combination of vitamin K2 and angiotensin-converting enzyme inhibitor ameliorates cumulative recurrence of hepatocellular carcinoma. *J Hepatol* **51**, 315-321, doi:10.1016/j.jhep.2009.04.011 (2009).
- 162 Lanaya, H. *et al.* EGFR has a tumour-promoting role in liver macrophages during hepatocellular carcinoma formation. *Nature cell biology* **16**, 972-981, 971-977, doi:10.1038/ncb3031 (2014).
- 163 Fuchs, B. C. *et al.* Epidermal growth factor receptor inhibition attenuates liver fibrosis and development of hepatocellular carcinoma. *Hepatology* **59**, 1577-1590, doi:10.1002/hep.26898 (2014).
- 164 Okita, K. *et al.* Survey of survival among patients with hepatitis C virus-related hepatocellular carcinoma treated with peretinoin, an acyclic retinoid, after the completion of a randomized, placebo-controlled trial. *Journal of gastroenterology* **50**, 667-674, doi:10.1007/s00535-014-0996-1 (2015).
- 165 Llovet, J. M. *et al.* Sorafenib in advanced hepatocellular carcinoma. *The New England journal of medicine* **359**, 378-390, doi:10.1056/NEJMoa0708857 (2008).
- 166 Matsui, J. *et al.* Multi-kinase inhibitor E7080 suppresses lymph node and lung metastases of human mammary breast tumor MDA-MB-231 via inhibition of vascular endothelial growth factor-receptor (VEGF-R) 2 and VEGF-R3 kinase. *Clin Cancer Res* **14**, 5459-5465, doi:10.1158/1078-0432.CCR-07-5270 (2008).
- 167 Kudo, M. *et al.* Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet* **391**, 1163-1173, doi:10.1016/S0140-6736(18)30207-1 (2018).
- 168 Wilhelm, S. M. *et al.* Regorafenib (BAY 73-4506): a new oral multikinase inhibitor of angiogenic, stromal and oncogenic receptor tyrosine kinases with potent preclinical antitumor activity. *Int J Cancer* **129**, 245-255, doi:10.1002/ijc.25864 (2011).
- 169 Bruix, J. *et al.* Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **389**, 56-66, doi:10.1016/S0140-6736(16)32453-9 (2017).
- 170 Abou-Alfa, G. K. *et al.* Cabozantinib in Patients with Advanced and Progressing Hepatocellular Carcinoma. *The New England journal of medicine* **379**, 54-63, doi:10.1056/NEJMoa1717002 (2018).

- 171 El-Khoueiry, A. B. *et al.* Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet* **389**, 2492-2502, doi:10.1016/S0140-6736(17)31046-2 (2017).
- 172 Company, B.-M. S. *Bristol-Myers Squibb Announces Results from CheckMate -459 Study Evaluating Opdivo (nivolumab) as a First-Line Treatment for Patients with Unresectable Hepatocellular Carcinoma* <<https://news.bms.com/press-release/bmy/bristol-myers-squibb-announces-results-checkmate-459-study-evaluating-opdivo-nivol>> (2019).
- 173 EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol* **69**, 182-236, doi:10.1016/j.jhep.2018.03.019 (2018).
- 174 Heimbach, J. K. *et al.* AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology* **67**, 358-380, doi:10.1002/hep.29086 (2018).
- 175 Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Experimental & molecular medicine* **50**, 96, doi:10.1038/s12276-018-0071-8 (2018).
- 176 Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nature methods* **6**, 377-382, doi:10.1038/nmeth.1315 (2009).
- 177 Deng, Q., Ramskold, D., Reinius, B. & Sandberg, R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343**, 193-196, doi:10.1126/science.1245316 (2014).
- 178 Ziegenhain, C. *et al.* Comparative Analysis of Single-Cell RNA Sequencing Methods. *Molecular cell* **65**, 631-643 e634, doi:10.1016/j.molcel.2017.01.023 (2017).
- 179 Ramskold, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nature biotechnology* **30**, 777-782, doi:10.1038/nbt.2282 (2012).
- 180 Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature methods* **10**, 1096-1098, doi:10.1038/nmeth.2639 (2013).
- 181 Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352-356, doi:10.1038/nature21065 (2017).
- 182 Halpern, K. B. *et al.* Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. *Nature biotechnology* **36**, 962-970, doi:10.1038/nbt.4231 (2018).
- 183 Ben-Moshe, S. & Itzkovitz, S. Spatial heterogeneity in the mammalian liver. *Nature reviews. Gastroenterology & hepatology* **16**, 395-410, doi:10.1038/s41575-019-0134-x (2019).
- 184 MacParland, S. A. *et al.* Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nature communications* **9**, 4383, doi:10.1038/s41467-018-06318-7 (2018).
- 185 Furuyama, K. *et al.* Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. *Nat Genet* **43**, 34-41, doi:10.1038/ng.722 (2011).
- 186 Font-Burgada, J. *et al.* Hybrid Periportal Hepatocytes Regenerate the Injured Liver without Giving Rise to Cancer. *Cell* **162**, 766-779, doi:10.1016/j.cell.2015.07.026 (2015).

- 187 Wang, B., Zhao, L., Fish, M., Logan, C. Y. & Nusse, R. Self-renewing diploid Axin2(+) cells fuel homeostatic renewal of the liver. *Nature* **524**, 180-185, doi:10.1038/nature14863 (2015).
- 188 Tarlow, B. D. *et al.* Bipotential adult liver progenitors are derived from chronically injured mature hepatocytes. *Cell Stem Cell* **15**, 605-618, doi:10.1016/j.stem.2014.09.008 (2014).
- 189 Deng, X. *et al.* Chronic Liver Injury Induces Conversion of Biliary Epithelial Cells into Hepatocytes. *Cell Stem Cell* **23**, 114-122 e113, doi:10.1016/j.stem.2018.05.022 (2018).
- 190 Schmelzer, E. *et al.* Human hepatic stem cells from fetal and postnatal donors. *The Journal of experimental medicine* **204**, 1973-1987, doi:10.1084/jem.20061603 (2007).
- 191 Lowes, K. N., Brennan, B. A., Yeoh, G. C. & Olynyk, J. K. Oval cell numbers in human chronic liver diseases are directly related to disease severity. *Am J Pathol* **154**, 537-541, doi:10.1016/S0002-9440(10)65299-6 (1999).
- 192 Segal, J. M. *et al.* Single cell analysis of human foetal liver captures the transcriptional profile of hepatobiliary hybrid progenitors. *Nature communications* **10**, 3350, doi:10.1038/s41467-019-11266-x (2019).
- 193 Pepe-Mooney, B. J. *et al.* Single-Cell Analysis of the Liver Epithelium Reveals Dynamic Heterogeneity and an Essential Role for YAP in Homeostasis and Regeneration. *Cell Stem Cell* **25**, 23-38 e28, doi:10.1016/j.stem.2019.04.004 (2019).
- 194 Planas-Paz, L. *et al.* YAP, but Not RSPO-LGR4/5, Signaling in Biliary Epithelial Cells Promotes a Ductular Reaction in Response to Liver Injury. *Cell Stem Cell* **25**, 39-53 e10, doi:10.1016/j.stem.2019.04.005 (2019).
- 195 Kaneko, K., Kamimoto, K., Miyajima, A. & Itoh, T. Adaptive remodeling of the biliary architecture underlies liver homeostasis. *Hepatology* **61**, 2056-2066, doi:10.1002/hep.27685 (2015).
- 196 Clapper, J. R. *et al.* Diet-induced mouse model of fatty liver disease and nonalcoholic steatohepatitis reflecting clinical disease progression and methods of assessment. *American journal of physiology. Gastrointestinal and liver physiology* **305**, G483-495, doi:10.1152/ajpgi.00079.2013 (2013).
- 197 Xiong, X. *et al.* Landscape of Intercellular Crosstalk in Healthy and NASH Liver Revealed by Single-Cell Secretome Gene Analysis. *Molecular cell* **75**, 644-660 e645, doi:10.1016/j.molcel.2019.07.028 (2019).
- 198 Krenkel, O., Hundertmark, J., Ritz, T. P., Weiskirchen, R. & Tacke, F. Single Cell RNA Sequencing Identifies Subsets of Hepatic Stellate Cells and Myofibroblasts in Liver Fibrosis. *Cells* **8**, doi:10.3390/cells8050503 (2019).
- 199 Ramachandran, P. *et al.* Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature*, doi:10.1038/s41586-019-1631-3 (2019).
- 200 Zheng, C. *et al.* Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* **169**, 1342-1356 e1316, doi:10.1016/j.cell.2017.05.035 (2017).
- 201 Zhang, Q. *et al.* Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas. *Gut* **68**, 2019-2031, doi:10.1136/gutjnl-2019-318912 (2019).

- 202 Ho, D. W. *et al.* Single-cell transcriptomics reveals the landscape of intra-tumoral heterogeneity and stemness-related subpopulations in liver cancer. *Cancer letters* **459**, 176-185, doi:10.1016/j.canlet.2019.06.002 (2019).
- 203 Duan, M. *et al.* Diverse modes of clonal evolution in HBV-related hepatocellular carcinoma revealed by single-cell genome sequencing. *Cell research* **28**, 359-373, doi:10.1038/cr.2018.11 (2018).
- 204 D'Avola, D. *et al.* High-density single cell mRNA sequencing to characterize circulating tumor cells in hepatocellular carcinoma. *Scientific reports* **8**, 11570, doi:10.1038/s41598-018-30047-y (2018).
- 205 Hou, Y. *et al.* Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell research* **26**, 304-319, doi:10.1038/cr.2016.23 (2016).
- 206 Xue, R. *et al.* Genomic and Transcriptomic Profiling of Combined Hepatocellular and Intrahepatic Cholangiocarcinoma Reveals Distinct Molecular Subtypes. *Cancer Cell* **35**, 932-947 e938, doi:10.1016/j.ccell.2019.04.007 (2019).
- 207 van den Brink, S. C. *et al.* Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nature methods* **14**, 935-936, doi:10.1038/nmeth.4437 (2017).
- 208 Kulik, L. & El-Serag, H. B. Epidemiology and Management of Hepatocellular Carcinoma. *Gastroenterology* **156**, 477-491 e471, doi:10.1053/j.gastro.2018.08.065 (2019).
- 209 Hagstrom, H. *et al.* Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. *J Hepatol* **67**, 1265-1273, doi:10.1016/j.jhep.2017.07.027 (2017).
- 210 Connolly, J. J., Ooka, K. & Lim, J. K. Future Pharmacotherapy for Non-alcoholic Steatohepatitis (NASH): Review of Phase 2 and 3 Trials. *Journal of clinical and translational hepatology* **6**, 264-275, doi:10.14218/JCTH.2017.00056 (2018).
- 211 Rotman, Y. & Sanyal, A. J. Current and upcoming pharmacotherapy for non-alcoholic fatty liver disease. *Gut* **66**, 180-190, doi:10.1136/gutjnl-2016-312431 (2017).
- 212 Friedman, S. L., Neuschwander-Tetri, B. A., Rinella, M. & Sanyal, A. J. Mechanisms of NAFLD development and therapeutic strategies. *Nature medicine* **24**, 908-922, doi:10.1038/s41591-018-0104-9 (2018).
- 213 Kazankov, K. *et al.* The role of macrophages in nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Nature reviews. Gastroenterology & hepatology* **16**, 145-159, doi:10.1038/s41575-018-0082-x (2019).
- 214 Junttila, M. R. & de Sauvage, F. J. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* **501**, 346-354, doi:10.1038/nature12626 (2013).
- 215 Lu, P. F., Weaver, V. M. & Werb, Z. The extracellular matrix: A dynamic niche in cancer progression. *J Cell Biol* **196**, 395-406, doi:10.1083/jcb.201102147 (2012).
- 216 Aizarani, N. *et al.* A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* **572**, 199-204, doi:10.1038/s41586-019-1373-2 (2019).

## 7. SUPPLEMENTARY MATERIAL

Other publications during my PhD:

---

7.1 Hamdane N, Jühling F, Crouchet E, El Saghire H, Thumann C, Oudot MA, Bandiera S, **Saviano A** et al. HCV-Induced Epigenetic Changes Associated With Liver Cancer Risk Persist After Sustained Virologic Response. *Gastroenterology*. 2019 Jun;156(8):2313-2329.e7.

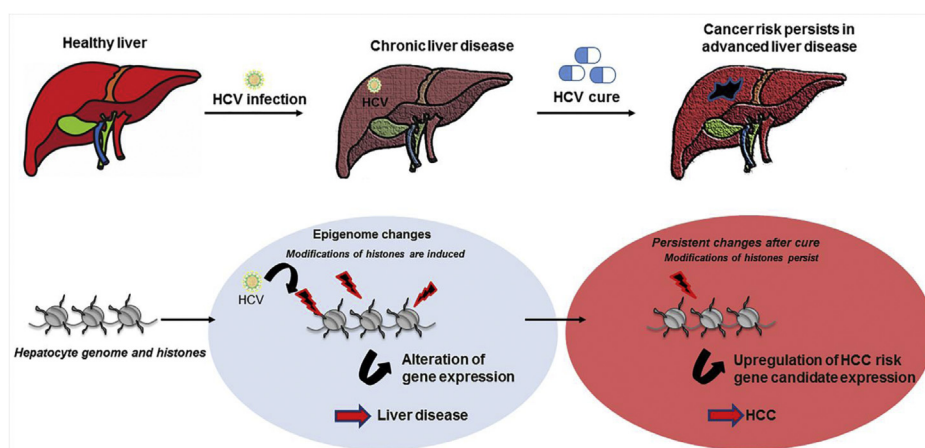
# BASIC AND TRANSLATIONAL—LIVER

## HCV-Induced Epigenetic Changes Associated With Liver Cancer Risk Persist After Sustained Virologic Response



Nouridine Hamdane,<sup>1,2,\*</sup> Frank Jühling,<sup>1,2,\*</sup> Emilie Crouchet,<sup>1,2</sup> Houssein El Saghire,<sup>1,2</sup> Christine Thumann,<sup>1,2</sup> Marine A. Oudot,<sup>1,2</sup> Simonetta Bandiera,<sup>1,2</sup> Antonio Saviano,<sup>1,2,11</sup> Clara Ponsolles,<sup>1,2</sup> Armando Andres Roca Suarez,<sup>1,2</sup> Shen Li,<sup>3</sup> Naoto Fujiwara,<sup>4</sup> Atsushi Ono,<sup>4,13,14</sup> Irwin Davidson,<sup>5</sup> Nabeel Bardeesy,<sup>6</sup> Christian Schmidl,<sup>7,8</sup> Christoph Bock,<sup>7,9,10</sup> Catherine Schuster,<sup>1,2</sup> Joachim Lupberger,<sup>1,2</sup> François Habersetzer,<sup>1,11</sup> Michel Doffoël,<sup>11</sup> Tullio Piardi,<sup>12</sup> Daniele Sommacale,<sup>12</sup> Michio Imamura,<sup>13</sup> Takuro Uchida,<sup>13</sup> Hideki Ohdan,<sup>14</sup> Hiroshi Aikata,<sup>13</sup> Kazuaki Chayama,<sup>13</sup> Tujana Boldanova,<sup>15,16</sup> Patrick Pessaux,<sup>1,11</sup> Bryan C. Fuchs,<sup>3</sup> Yujin Hoshida,<sup>4</sup> Mirjam B. Zeisel,<sup>1,2,17</sup> François H. T. Duong,<sup>1,2,15</sup> and Thomas F. Baumert<sup>1,2,11,18</sup>

<sup>1</sup>INSERM U1110, Institut de Recherche sur les Maladies Virales et Hépatiques, Strasbourg, France; <sup>2</sup>Université de Strasbourg, Strasbourg, France; <sup>3</sup>Division of Surgical Oncology, Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, Massachusetts; <sup>4</sup>Liver Tumor Translational Research Program, Harold C. Simmons Comprehensive Cancer Center, Division of Digestive and Liver Diseases, University of Texas Southwestern Medical Center, Dallas, Texas; <sup>5</sup>Department of Functional Genomics and Cancer, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/UDS, Illkirch, France; <sup>6</sup>Center for Cancer Research, Massachusetts General Hospital; Departments of Medicine, Harvard Medical School, Boston, Massachusetts; <sup>7</sup>CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria; <sup>8</sup>Regensburg Centre for Interventional Immunology (RCI) and University Medical Center of Regensburg, Regensburg, Germany; <sup>9</sup>Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria; <sup>10</sup>Max Planck Institute for Informatics, Saarbrücken, Germany; <sup>11</sup>Institut Hospitalo-Universitaire, Pôle Hépatite-digestif, Nouvel Hôpital Civil, Strasbourg, France; <sup>12</sup>General, Digestive, and Endocrine Surgery Unit, Hôpital Robert Debré, Centre Hospitalier Universitaire de Reims, Université de Reims Champagne-Ardenne, Reims, France; <sup>13</sup>Department of Gastroenterology and Metabolism, Applied Life Sciences, Institute of Biomedical & Health Sciences, Hiroshima University, Hiroshima, Japan; <sup>14</sup>Department of Gastroenterological and Transplant Surgery, Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan; <sup>15</sup>Department of Biomedicine, University Hospital Basel, University of Basel, Basel, Switzerland; <sup>16</sup>Division of Gastroenterology and Hepatology, University Hospital Basel, University of Basel, Basel, Switzerland; <sup>17</sup>INSERM U1052, CNRS UMR 5286, Cancer Research Center of Lyon (CRCL), Université de Lyon (UCBL), Lyon, France; and <sup>18</sup>Institut Universitaire de France (IUF), Paris, France



Gastroenterology

See editorial on page 2130. See Covering the Cover synopsis on 2118.

**BACKGROUND & AIMS:** Chronic hepatitis C virus (HCV) infection is an important risk factor for hepatocellular

carcinoma (HCC). Despite effective antiviral therapies, the risk for HCC is decreased but not eliminated after a sustained virologic response (SVR) to direct-acting antiviral (DAA) agents, and the risk is higher in patients with advanced fibrosis. We investigated HCV-induced epigenetic alterations that might affect risk for HCC after DAA treatment in patients and mice

with humanized livers. **METHODS:** We performed genome-wide ChIPmentation-based ChIP-Seq and RNA-seq analyses of liver tissues from 6 patients without HCV infection (controls), 18 patients with chronic HCV infection, 8 patients with chronic HCV infection cured by DAA treatment, 13 patients with chronic HCV infection cured by interferon therapy, 4 patients with chronic hepatitis B virus infection, and 7 patients with nonalcoholic steatohepatitis in Europe and Japan. HCV-induced epigenetic modifications were mapped by comparative analyses with modifications associated with other liver disease etiologies. uPA/SCID mice were engrafted with human hepatocytes to create mice with humanized livers and given injections of HCV-infected serum samples from patients; mice were given DAAs to eradicate the virus. Pathways associated with HCC risk were identified by integrative pathway analyses and validated in analyses of paired HCC tissues from 8 patients with an SVR to DAA treatment of HCV infection. **RESULTS:** We found chronic HCV infection to induce specific genome-wide changes in H3K27ac, which correlated with changes in expression of mRNAs and proteins. These changes persisted after an SVR to DAAs or interferon-based therapies. Integrative pathway analyses of liver tissues from patients and mice with humanized livers demonstrated that HCV-induced epigenetic alterations were associated with liver cancer risk. Computational analyses associated increased expression of SPHK1 with HCC risk. We validated these findings in an independent cohort of patients with HCV-related cirrhosis ( $n = 216$ ), a subset of which ( $n = 21$ ) achieved viral clearance. **CONCLUSIONS:** In an analysis of liver tissues from patients with and without an SVR to DAA therapy, we identified epigenetic and gene expression alterations associated with risk for HCC. These alterations might be targeted to prevent liver cancer in patients treated for HCV infection.

**Keywords:** Biomarker; Biopsy; Chemoprevention; Sox9.

**C**hronic hepatitis C virus (HCV) infection is a leading cause of hepatocellular carcinoma (HCC), the second most common and fastest rising cause of cancer-related death.<sup>1</sup> The development of direct-acting antivirals (DAAs) with cure rates of higher than 90% has been a major breakthrough in the management of patients with chronic HCV infection. However, although viral cure decreases the overall HCC risk in HCV-infected patients, it does not eliminate virus-induced HCC risk, especially in patients with advanced fibrosis.<sup>2,3</sup> Furthermore, convenient biomarkers to robustly predict HCC risk after viral cure and strategies for HCC prevention are absent.<sup>2</sup> These unexpected findings pose new challenges for patient management.<sup>4-6</sup>

Despite more than 2 decades of intensive research efforts, the pathogenesis of HCV-induced HCC and the HCC risk after DAA cure are still incompletely understood.<sup>6,7</sup> Although HCV is an RNA virus with little potential for integrating its genetic material into the host genome, HCV contributes to hepatocarcinogenesis through a direct and an indirect way. HCV-mediated liver disease and carcinogenesis are considered multistep processes that include chronic infection-driven hepatic inflammation and progressive liver fibrogenesis with formation of neoplastic clones

## WHAT YOU NEED TO KNOW

### BACKGROUND AND CONTEXT

Despite effective antiviral therapies, the risk for HCC is not eliminated following a sustained virologic response to direct-acting antiviral (DAA) agents, and risk is higher in patients with advanced fibrosis.

### NEW FINDINGS

In an analysis of liver tissues from patients with and without a sustained virologic response to DAA therapy, and from HCV-infected mice with humanized livers, the authors identified epigenetic and gene expression alterations associated with risk for HCC.

### LIMITATIONS

This was a retrospective analysis of liver tissues from patients and mice.

### IMPACT

The epigenetic alterations identified in this study might be targeted to prevent liver cancer in patients treated for HCV infection.

that arise and progress in the carcinogenic tissue microenvironment.<sup>4,6,8</sup> A 186-gene expression signature in liver tissue of HCV-infected patients has been associated with HCC risk and mortality, suggesting that virus-induced transcriptional reprogramming in the liver could play a functional role in hepatocarcinogenesis.<sup>9,10</sup>

Epigenetic modifications of histones can lead to chromatin opening and compacting and play a major role in gene regulation in health and disease.<sup>11</sup> Although epigenetic changes have been identified in established HCC,<sup>12</sup> their role in viral hepatocarcinogenesis remains largely unknown.

## Methods

### Human Subjects

Liver tissues from patients undergoing surgical resection or biopsy examination were collected at the Gastroenterology and Hepatology Clinic of the Hiroshima University Hospital (Hiroshima, Japan), the Basel University Hospital (Basel, Switzerland), the Centre Hospitalier Universitaire de Reims (Reims, France), and the Hôpitaux Universitaires de Strasbourg (Strasbourg, France). Protocols for patient tissue collection

\*Authors share co-first authorship.

**Abbreviations used in this paper:** DAA, direct-acting antiviral; FC, fold change; GSEA, gene set enrichment analysis; HCC, hepatocellular carcinoma; HBV, hepatitis B virus; HCV, hepatitis C virus; IFN, interferon; mTOR, mammalian target of rapamycin; NASH, nonalcoholic steatohepatitis; PCA, principal component analysis; PLS, prognostic liver signature; SVR, sustained virologic response; TNF $\alpha$ , tumor necrosis factor  $\alpha$ ; TSG, tumor suppressor gene.

 Most current article

© 2019 by the AGA Institute. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

0016-5085

<https://doi.org/10.1053/j.gastro.2019.02.038>



were reviewed and approved by the hospital ethics committees. Written and informed consent was obtained from all patients. Eligible patients were identified by a systematic review of patient charts. Histopathologic grading and staging of HCV liver biopsy specimens, according to the METAVIR classification system, were performed at the pathology institutes of the respective university hospitals. Overall, we analyzed liver tissue from 6 noninfected control patients, 18 patients with chronic HCV infection, 8 patients with DAA-cured chronic HCV, 13 patients with interferon (IFN)-cured chronic HCV, 4 patients with hepatitis B virus (HBV) infection, and 7 patients with nonalcoholic steatohepatitis (NASH). Furthermore, we studied 8 paired HCC samples with HCV-induced liver disease (Table 1).

### *HCV Infection of Human Hepatocyte Chimeric Mice and DAA Treatment*

cDNA-uPA<sup>+/+</sup>/SCID<sup>+/+</sup> (uPA/SCID) mice were engrafted with human hepatocytes and intravenously inoculated with serum samples containing approximately 10<sup>5</sup> HCV particles. HCV-infected mice were treated with a combination of MK-7009 and BMS-788329 DAAs.<sup>13</sup> Elimination of HCV in treated mice was confirmed by the absence of HCV viremia 12 weeks after cessation of therapy. See the [Supplementary Materials](#) for further details.

### *ChIPmentation-Based ChIP-Seq*

ChIPmentation-based ChIP-Seq on liver tissue using H3K27ac antibody (number 39134, Activ Motif, La Hulpe, Belgium) was performed as described previously<sup>14</sup> and adapted as follows. To perform ChIP-Seq on human and mouse livers, tissues were cut in small pieces of 2–3 mm, crosslinked with 0.4% formaldehyde for 10 minutes at room temperature, and quenched with glycine 125 mmol/L for 5 minutes at room temperature. Then, tissue was homogenized using a glass potter and ChIPmentation was performed as described previously.<sup>14</sup>

### *Processing of Raw ChIPmentation Data*

Reads were aligned to the human genome (hg19) and peaks were called in uniquely mapped reads using MACS2.<sup>8</sup> Peaks within all samples were intersected and used for counting reads if they overlapped in at least 2 samples. Read counts of genes were defined as the sum of all reads in peak regions overlapping the gene body or the promoter region, that is, the region up to 1500 bp ahead of the transcription start site. See the [Supplementary Materials](#) for further details.

### *RNA Extraction and Next-Generation Sequencing*

Liver tissues were lysed in TRI-reagent (Molecular Research Center; Cincinnati, OH) and RNA was purified using Direct-zol RNA MiniPrep (Zymo Research, Irvine, CA) or RNeasy kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. RNA quantity and quality were assessed using NanoDrop (Thermo Scientific, Waltham, MA) and Bioanalyzer 2100 (Illumina, San Diego, CA). Libraries of extracted RNA were prepared and sequenced as described previously.<sup>3,9</sup>

### *Processing of RNA-Seq Data*

Reads were counted with htseq-count, and a differentially expression analysis was performed with DESeq2 applying GENCODE 19.<sup>15</sup> Reads were taken from our RNA-Seq experiments as described earlier and from external sources: RNA-Seq from infected (low ISG) vs control patients was retrieved from the GEO dataset GSE84346 (low ISG samples). See the [Supplementary Materials](#) for further details.

### *Pathway Enrichment and Correlation Analyses*

Pathway enrichment analyses were performed using gene set enrichment analysis (GSEA) with all gene sets included in MSigDB 6.0.<sup>16</sup> We used the pre-ranked version of GSEA and genes were ranked for *P* values of differential expression and modification analyses. Figures showing enriched pathways and gene sets, Spearman correlations, and oncogene log<sub>2</sub> fold change (FC) were drawn using ggplot2 and the R environment (R Foundation, Vienna, Austria). Gene network analysis was performed based on 3 MSigDB subsets: Hallmark gene sets, curated gene sets, and gene ontology gene sets. See the [Supplementary Materials](#) for further details.

### *Western Blot*

Expression of SPHK1 and SOX9 proteins was assessed by western blot and quantified using ImageJ software (National Institutes of Health, Bethesda, MD). See the [Supplementary Materials](#) for further details.

### *Association of Hepatic Gene Expression With Prognostic Cox Score for Overall Death*

Prognostic association of hepatic gene expression was determined using the Cox score for time to overall death in HCV-infected patients with advanced liver disease and HCC as previously described.<sup>17</sup>

### *Gene Expression and Assessment of HCC Risk in HCV Cohorts*

Patients with early-stage HCV cirrhosis (*n* = 216<sup>10</sup>; GSE15654) and a subgroup of patients who had achieved a sustained virologic response (SVR) before the biopsy (*n* = 21) were classified into SPHK1-high and -low expression groups based on the cutoff value of 1 sample standard deviation above the mean. Cumulative probabilities of HCC development were calculated using the Kaplan-Meier procedure and compared by log-rank test.

### *Data Availability*

The Sequence Read Archive accession number for the data reported in this study is SRP170244.

## **Results**

### *Virus-Induced Modifications of Histone Mark H3K27ac Persist in Human Liver After DAA Cure in HCV-Infected Patients*

To investigate whether chronic HCV infection triggers persistent epigenetic modifications after cure, we performed a genome-wide analysis using ChIPmentation-based

**Table 1.** Characteristics of Studied Patients

|              | Biopsy ID | Sex | Age | Diagnosis                             | Viral genotype | Viral load (IU/mL) | METAVIR grade | METAVIR stage | Antiviral treatment         |
|--------------|-----------|-----|-----|---------------------------------------|----------------|--------------------|---------------|---------------|-----------------------------|
| Controls     | C1        | F   | 55  | Minimal hepatitis                     | N/A            | N/A                | N/A           | F0            | N/A                         |
|              | C2        | M   | 46  | Minimal hepatitis                     | N/A            | N/A                | N/A           | F0            | N/A                         |
|              | C3        | F   | 40  | Lobular hepatitis                     | N/A            | N/A                | N/A           | F0            | N/A                         |
|              | C4        | F   | 53  | Minimal hepatitis                     | N/A            | N/A                | N/A           | F0            | N/A                         |
|              | C5        | M   | 56  | Lobular hepatitis                     | N/A            | N/A                | N/A           | F0            | N/A                         |
|              | C6        | F   | 58  | Minimal hepatitis                     | N/A            | N/A                | N/A           | F0            | N/A                         |
|              | C7        | F   | 51  | Chronic indeterminate hepatitis       | N/A            | N/A                | N/A           | F3            | N/A                         |
|              | C8        | F   | 37  | Acute partially cholestatic hepatitis | N/A            | N/A                | N/A           | F0            | N/A                         |
|              | C9        | F   | 44  | Cholestatic hepatitis                 | N/A            | N/A                | N/A           | F1            | N/A                         |
|              | C10       | M   | 78  | Adjacent liver from CCM resection     | N/A            | N/A                | N/A           | N/A           | N/A                         |
|              | C11       | F   | 58  | Adjacent liver from CCM resection     | N/A            | N/A                | N/A           | N/A           | N/A                         |
|              | C12       | F   | 70  | Adjacent liver from CCM resection     | N/A            | N/A                | N/A           | N/A           | N/A                         |
|              | C13       | M   | 63  | Adjacent liver from CCM resection     | N/A            | N/A                | N/A           | N/A           | N/A                         |
|              | C14       | M   | 70  | Adjacent liver from CCM resection     | N/A            | N/A                | N/A           | N/A           | N/A                         |
|              | C15       | F   | 69  | Adjacent liver from CCM resection     | N/A            | N/A                | N/A           | N/A           | N/A                         |
|              | C16       | M   | 53  | Adjacent liver from CCM resection     | N/A            | N/A                | N/A           | N/A           | N/A                         |
|              | C17       | M   | 71  | Adjacent liver from CCM resection     | N/A            | N/A                | N/A           | N/A           | N/A                         |
| HBV          | B1        | F   | 46  | HBV                                   | N/A            | N/A                | N/A           | F4            | NUC                         |
|              | B2        | M   | 65  | HBV and HCC                           | N/A            | N/A                | N/A           | F4            | NUC                         |
|              | B3        | M   | 57  | HBV and HCC                           | N/A            | N/A                | N/A           | F4            | NUC                         |
|              | B4        | M   | 58  | HBV and HCC                           | N/A            | N/A                | N/A           | F4            | NUC                         |
| NASH         | N1        | M   | 27  | NASH and HCC                          | N/A            | N/A                | N/A           | F4            | N/A                         |
|              | N2        | M   | 63  | NASH and HCC                          | N/A            | N/A                | N/A           | F4            | N/A                         |
|              | N3        | M   | 73  | NASH and HCC                          | N/A            | N/A                | N/A           | F4            | N/A                         |
|              | N4        | M   | 76  | NASH and HCC                          | N/A            | N/A                | N/A           | F4            | N/A                         |
|              | N5        | F   | 65  | NASH and HCC                          | N/A            | N/A                | N/A           | F4            | N/A                         |
|              | N6        | F   | 47  | NASH and HCC                          | N/A            | N/A                | N/A           | F4            | N/A                         |
|              | N7        | F   | 68  | NASH and HCC                          | N/A            | N/A                | N/A           | F4            | N/A                         |
| HCV infected | H1        | F   | 62  | Chronic HCV                           | 1a             | 5140000            | A1            | F1            | Naïve                       |
|              | H2        | M   | 44  | Chronic HCV                           | 1a             | 7.41E + 06         | A1            | F2            | Naïve                       |
|              | H3        | F   | 23  | Chronic HCV                           | 3a             | 2.46E + 02         | A2            | F2            | Naïve                       |
|              | H4        | F   | 60  | Chronic HCV                           | 2              | 2.70E + 06         | A2            | F2            | Naïve                       |
|              | H5        | M   | 23  | Chronic HCV                           | 1a             | 1.76E + 06         | A1            | F1            | Intolerant to Peg-IFN/RBV   |
|              | H6        | M   | 48  | Chronic HCV                           | 1a             | 5.93E + 06         | A1            | F2            | Naïve                       |
|              | H7        | F   | 38  | Chronic HCV                           | 1b             | 7.95E + 05         | A1            | F2            | Naïve                       |
|              | H8        | M   | 58  | Chronic HCV                           | 4              | 4.08E + 06         | A3            | F2            | Nonresponder to Peg-IFN/RBV |
|              | H9        | M   | 52  | Chronic HCV                           | 1a             | 6.60E + 05         | A3            | F3            | Naïve                       |
|              | H10       | M   | 54  | Chronic HCV and HCC                   | 1b             | 4.40E + 04         | A1            | F4            | Relapse to SOF/DCV/RBV      |
|              | H11       | M   | 68  | Chronic HCV and HCC                   | 2a             | 2.51E + 05         | A3            | F3            | Naïve                       |

Table 1. Continued

|           | Biopsy ID        | Sex | Age | Diagnosis           | Viral genotype | Viral load (IU/mL) | METAVIR grade | METAVIR stage | Antiviral treatment         |
|-----------|------------------|-----|-----|---------------------|----------------|--------------------|---------------|---------------|-----------------------------|
|           | H12              | M   | 51  | Chronic HCV         | 3a             | 3.30E + 06         | A2            | F1            | Naïve                       |
|           | H13              | M   | 54  | Chronic HCV         | 4              | 3.31E + 06         | A2            | F1            | Naïve                       |
|           | H14              | F   | 48  | Chronic HCV         | 3a             | 1.15E + 06         | A3            | F4            | Naïve                       |
|           | H15              | M   | 65  | Chronic HCV         | 1b             | 2.25E + 06         | A2            | F4            | Naïve                       |
|           | H16              | M   | 81  | Chronic HCV and HCC | 1b             | 1.85E + 06         | A1            | F1            | Nonresponder to Peg-IFN/RBV |
|           | H17              | M   | 51  | Chronic HCV and HCC | 3a             | 3.79E + 06         | A2            | F4            | Relapse to SOF/RBV          |
|           | H18              | F   | 71  | Chronic HCV and HCC | 1b             | 3.93E + 06         | A1            | F1            | Naïve                       |
|           | H19              | F   | 49  | Chronic HCV         | 3a             | 3.50E + 06         | A3            | F4            | Naïve                       |
|           | H20              | M   | 34  | Chronic HCV         | N/A            | 2.21E + 06         | A3            | F4            | Naïve                       |
|           | H21              | M   | 53  | Chronic HCV         | 1              | 1.35E + 06         | A3            | F4            | Naïve                       |
|           | H22              | F   | 62  | Chronic HCV         | N/A            | 6.10E + 06         | A3            | F4            | Naïve                       |
|           | H23              | F   | 59  | Chronic HCV         | 4              | 2.68E + 06         | A3            | F4            | Naïve                       |
|           | H24 <sup>a</sup> | M   | 79  | Chronic HCV and HCC | 1b             | 2.00E + 06         | A2            | F2            | N/A                         |
|           | H25 <sup>a</sup> | M   | 56  | Chronic HCV and HCC | 1b             | 2.00E + 06         | A3            | F4            | N/A                         |
|           | H26 <sup>a</sup> | F   | 79  | Chronic HCV and HCC | 1b             | 5.01E + 05         | A2            | F4            | N/A                         |
|           | H27 <sup>a</sup> | M   | 85  | Chronic HCV and HCC | 1b             | 3.16E + 05         | A3            | F3            | N/A                         |
|           | H28 <sup>a</sup> | M   | 64  | Chronic HCV and HCC | 2b             | 1.00E + 07         | A2            | F4            | N/A                         |
|           | H29 <sup>a</sup> | F   | 76  | Chronic HCV and HCC | 1b             | 6.31E + 06         | A2            | F4            | N/A                         |
|           | H30 <sup>a</sup> | F   | 84  | Chronic HCV and HCC | 1b             | 5.01E + 04         | A2            | F3            | N/A                         |
|           | H31 <sup>a</sup> | M   | 61  | Chronic HCV and HCC | 1b             | 3.98E + 04         | A2            | F2            | N/A                         |
| HCV cured | D1 <sup>a</sup>  | M   | 65  | Cured HCV and HCC   | 1b             | Undetectable       | A0            | F2            | SOF/DCV                     |
|           | D2 <sup>a</sup>  | M   | 58  | Cured HCV and HCC   | 1a             | Undetectable       | A0            | F4            | SOF/LDV                     |
|           | D3 <sup>a</sup>  | F   | 79  | Cured HCV and HCC   | 1b             | Undetectable       | A2            | F4            | DCV/ASV                     |
|           | D4 <sup>a</sup>  | M   | 63  | Cured HCV and HCC   | 2a             | Undetectable       | A2            | F4            | SOF/RBV                     |
|           | D5 <sup>a</sup>  | M   | 69  | Cured HCV and HCC   | 1b             | Undetectable       | A2            | F3            | DCV/ASV                     |
|           | D6 <sup>a</sup>  | M   | 73  | Cured HCV and HCC   | 1b             | Undetectable       | A2            | F3            | DCV/ASV                     |
|           | D7 <sup>a</sup>  | M   | 75  | Cured HCV and HCC   | 1b             | Undetectable       | A2            | F3            | SOF/LDV                     |
|           | D8 <sup>a</sup>  | F   | 75  | Cured HCV and HCC   | 1b             | Undetectable       | A2            | F3            | SOF/LDV                     |
|           | D9 <sup>a</sup>  | M   | 71  | Cured HCV and HCC   | 1B             | Undetectable       | A3            | F2            | DCV/ASV                     |
|           | D10 <sup>a</sup> | M   | 73  | Cured HCV and HCC   | 1B             | Undetectable       | A2            | F3            | DCV/ASV                     |
|           | D11 <sup>a</sup> | F   | 76  | Cured HCV and HCC   | 1B             | Undetectable       | A2            | F2            | DCV/ASV                     |
|           | D12 <sup>a</sup> | M   | 61  | Cured HCV and HCC   | 2A             | Undetectable       | A2            | F3            | SOF/RBV                     |
|           | D13 <sup>a</sup> | F   | 71  | Cured HCV and HCC   | 1B             | Undetectable       | A2            | F4            | DCV/ASV                     |
|           | D14 <sup>a</sup> | M   | 79  | Cured HCV and HCC   | 1B             | Undetectable       | N/A           | N/A           | DCV/ASV                     |
|           | D15 <sup>a</sup> | M   | 64  | Cured HCV and HCC   | 1B             | Undetectable       | A2            | F3            | SOF/LDV                     |
|           | D16 <sup>a</sup> | M   | 78  | Cured HCV and HCC   | 1B             | Undetectable       | A1            | F1            | SOF/LDV                     |
|           | I1               | M   | 68  | Cured HCV and HCC   | 2A             | Undetectable       | N/A           | F3            | Peg-IFN/RBV                 |
|           | I2               | M   | 61  | Cured HCV and HCC   | 2A             | Undetectable       | A2            | F4            | Peg-IFN/RBV                 |
|           | I3               | F   | 74  | Cured HCV and HCC   | 2B             | Undetectable       | A2            | F3            | IFN/RBV                     |
|           | I4               | M   | 69  | Cured HCV and HCC   | 1B             | Undetectable       | A1            | F2            | Peg-IFN/RBV                 |
|           | I5               | M   | 66  | Cured HCV and HCC   | 2B             | Undetectable       | A2            | F4            | IFN                         |

Table 1. Continued

| Biopsy ID | Sex | Age | Diagnosis         | Viral genotype | Viral load (IU/mL) | METAVIR grade | METAVIR stage | Antiviral treatment |
|-----------|-----|-----|-------------------|----------------|--------------------|---------------|---------------|---------------------|
| I6        | F   | 68  | Cured HCV and HCC | 1B             | Undetectable       | A2            | F2            | IFN                 |
| I7        | M   | 54  | Cured HCV and HCC | 1B             | Undetectable       | A2            | F4            | Peg-IFN/RBV         |
| I8        | M   | 66  | Cured HCV and HCC | 1B             | Undetectable       | A1            | F3            | IFN                 |
| I9        | M   | 74  | Cured HCV and HCC | 2A             | Undetectable       | A2            | F1            | Peg-IFN             |
| I10       | M   | 80  | Cured HCV and HCC | 1B             | Undetectable       | A1            | F2            | Peg-IFN/RBV         |
| I11       | F   | 77  | Cured HCV and HCC | 1B             | Undetectable       | A1            | F4            | Peg-IFN/RBV         |
| I12       | M   | 70  | Cured HCV and HCC | 1B             | Undetectable       | A1            | F1            | Peg-IFN             |
| I13       | M   | 65  | Cured HCV and HCC | 1B             | Undetectable       | A2            | F2            | Peg-IFN/RBV         |

NOTE. Biopsy identification number, sex, age, pathologic diagnosis, HCV genotype and load, antiviral treatment (for HCV-infected and HCV-cured patients), and METAVIR grade (when applicable) and score are presented.

ASV, asunaprevir; CCM, colon cancer metastasis; DCV, daclatasvir; F, female; IU, international unit; LDV, ledipasvir; M, male; N/A, not applicable; NUC, nucleos(t)ide analogues; Peg, pegylated; RBV, ribavirin; SOF, sofosbuvir.

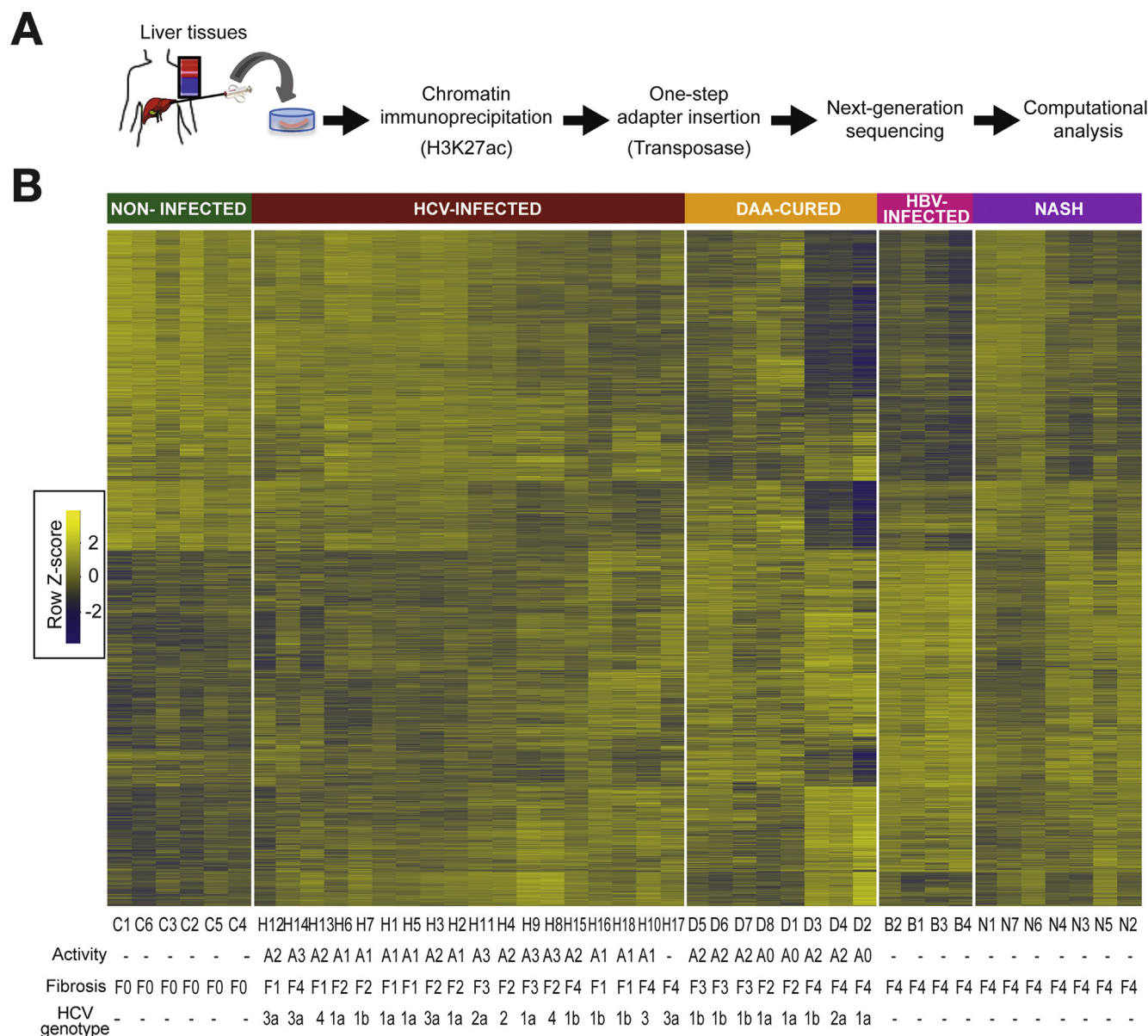
<sup>a</sup>Paired analysis of HCC and nontumor tissue.

ChIP-Seq<sup>14</sup> profiling the well-characterized histone modification H3K27ac in liver tissues from 18 patients with chronic HCV infection, 21 patients with DAA- or IFN-based curative therapy, and 6 noninfected controls (Figure 1A and Table 1). The H3K27ac modification is associated with active promoters and enhancers and with activation of transcription.<sup>18</sup> We observed significant changes in specific H3K27ac modifications in HCV-infected patients compared with noninfected controls (Figures 1B and Supplementary Figure 1). To study whether these were etiology specific, we performed comparative analyses of liver tissues with chronic HCV infection (n = 18), chronic HBV infection (n = 4), and NASH (n = 7). Using principal component analysis (PCA), we found that the distribution of H3K27ac changes in the epigenome of livers of noninfected, HCV-infected, HBV-infected, and NASH samples formed distinct clusters on the PCA plot, suggesting that an important part of the changes are etiology specific (Figure 2A). Next, we performed a correlation analysis of H3K27ac changes among HCV-infected, HBV-infected, and NASH samples. Our data showed a positive correlation of H3K27ac changes (Figure 2B) among patients with NASH ( $r = 0.83$ ;  $P < 10^{-10}$ ), or patients with HBV infection ( $r = 0.79$ ;  $P < 10^{-10}$ ), or HCV infection, suggesting that some epigenetic modifications are shared among etiologies. To analyze the impact of epigenetic changes in genes related to immune responses, we extracted immune-related genes from MSigDB and performed a restricted correlation study that showed lower correlation coefficients (NASH vs HCV,  $r = 0.75$ ,  $P < 10^{-10}$ ; HBV vs HCV,  $r = 0.62$ ,  $P < 10^{-10}$ ) compared with analyses composed of all genes (Supplementary Figure 2). These findings suggest that epigenetic modifications in immune genes associated with inflammatory responses are only partly responsible for the similarities between etiologies.

Recent studies have reported a correlation between fibrosis and an increased incidence of HCC.<sup>6</sup> However, the molecular mechanism of fibrosis-induced HCC is not well understood. Our comparative analysis showed that H3K27ac modifications, separated based on fibrosis score along the primary component (dimension 1), accounted for 42% of the variation between samples. This suggests that a substantial fraction of the observed H3K27ac alterations is related to liver fibrosis. Interestingly, we did not observe any significant correlation between these epigenetic changes and the activity score (ie, reflecting liver inflammation), suggesting that aberrant H3K27 acetylation is less dependent of necro-inflammatory activity but rather dependent on the fibrosis stage (Figure 1B).

By comparing H3K27ac modifications in liver tissue with chronic HCV infection before DAA treatment and in liver tissue with successful DAA cure, we studied whether epigenetic changes persisted in cured patients. Interestingly, we found a significant and positive correlation of H3K27ac modifications after comparing HCV-infected and DAA-cured samples ( $r = 0.87$ ;  $P < 10^{-10}$ ; Figure 2C). A comparative analysis showed a strong positive correlation between epigenetic changes in liver samples of DAA-cured and IFN-cured patients ( $r = 0.91$ ;  $P < 10^{-10}$ ; Supplementary Figure 1B), suggesting that HCV-





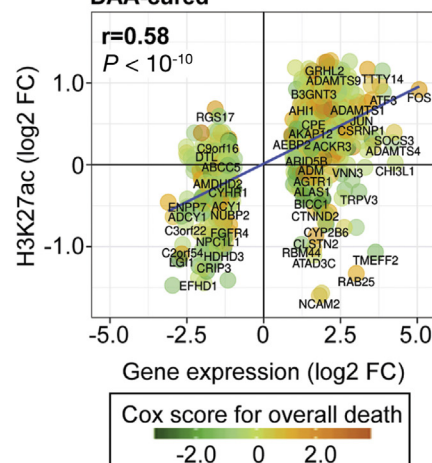
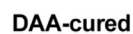
**Figure 1.** HCV-induced epigenetic changes persist after HCV clearance in patient-derived liver tissue. (A) Approach: HCV-induced H3K27ac histone modifications were measured genome-wide using a ChIPmentation-based ChIP-Seq protocol optimized for low input material such as patient-derived liver biopsy samples and resections. (B) Unsupervised clustering of normalized read counts in ChIP-Seq peaks of 12,700 genes linked with significant ( $P < .05$ ) H3K27ac modifications in HCV-infected ( $n = 18$ ), DAA-cured ( $n = 8$ ), HBV-infected ( $n = 4$ ), or NASH ( $n = 7$ ) vs noninfected control ( $n = 6$ ) patients.

induced epigenetic changes persist after DAA- and IFN-based therapies.

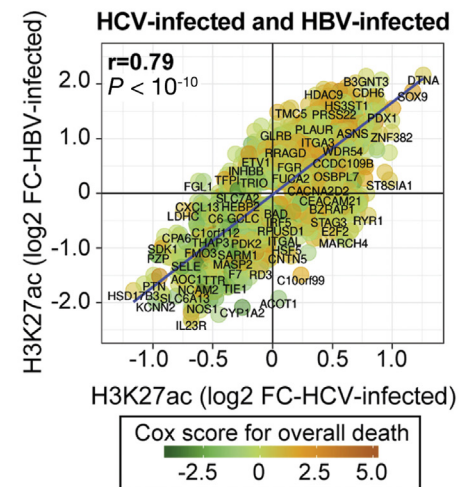
To address the potential clinical relevance, we next analyzed genes that were epigenetically modulated by HCV infection by integrating ChIP-Seq data and by assigning a gene expression-based Cox score for overall death based on the clinical outcome of a cohort of 216 HCV-induced cirrhotic patients who later developed HCC.<sup>10</sup> We chose this score because it has been shown to robustly predict clinical outcome of patients with advanced HCV liver disease.<sup>10</sup> Importantly, we found that persistent H3K27ac

modifications were linked with genes associated with a high Cox score for overall death in HCV-infected patients and advanced liver disease<sup>17</sup> (Figure 2C), confirming the clinical impact of these findings. Next, we compared H3K27ac enrichment and transcriptomic changes in HCV-infected and in DAA-cured patients. We found a positive correlation between H3K27ac and gene expression changes in HCV-infected and DAA-cured patients ( $r = 0.73$ ;  $P < 10^{-10}$  and  $r = 0.58$ ;  $P < 10^{-10}$ , respectively; Figure 2D), supporting the functional relevance of these epigenetic changes for the deregulation of gene transcription that persists after cure.

**A**



# B



## Persistent Epigenetic Changes Are Associated With Liver Carcinogenesis After Cure

Epigenetic regulation is an indispensable process for normal development and preservation of tissue-specific gene expression profiles. Thus, any perturbation in the epigenetic landscape can lead to shifted gene function and malignant cellular transformation. We addressed the potential functional role of the observed alterations for virus-induced liver disease and hepatocarcinogenesis by performing a pathway enrichment analysis of genes associated with H3K27ac changes in liver tissues from HCV-infected and cured patients. We found that chronic HCV infection induces significant epigenetic H3K27ac changes on genes that belong to pathways related to tumor necrosis factor  $\alpha$  (TNF $\alpha$ ), inflammatory response, and interleukin 2 and signal transducer and activator of transcription 5 signaling (Figure 3A). Furthermore, we observed lower levels of H3K27ac within genes related to pathways associated with coagulation and metabolism, such as oxidative phosphorylation, fatty acid metabolism, or adipogenesis (Figure 3A). Remarkably, several altered pathways persisted after cure (eg, TNF $\alpha$  signaling, inflammatory response, G2M checkpoint, epithelial-mesenchymal transition, and phosphoinositide 3-kinase, Akt, and mammalian target of rapamycin [mTOR]; Figure 3A). We also observed lower levels of H3K27ac mapping to genes related to oxidative phosphorylation pathways (Figure 3A). Overall, our data provide evidence supporting a functional role for H3K27ac changes in establishing gene expression patterns that persist after cure and contribute to carcinogenesis.

We proceeded to study the impact of fibrosis on persistence of epigenetic modifications. Our analysis showed that H3K27ac changes observed in HCV-infected patients were partly reversed in cured patients with stage F2–3 fibrosis. This group shared 2259 of the 5318 (42.5%) modified genes in the HCV-infected group (Figure 3B). In contrast, in DAA-cured patients with advanced liver disease (F4), the HCV-induced H3K27ac changes largely persisted. The HCV-infected group shared nearly all modified genes (96.6%, 5140 of 5318 genes) with F4 cured patients (Figure 3B). Collectively, we identified significant changes of H3K27ac levels on 2193 genes persisting in the 2 DAA-cured patient groups (Figure 3B and Supplementary Table 1). Among these candidates, we identified oncogenes and tumor suppressor genes (TSGs) that are associated with, respectively, increased or decreased levels of H3K27ac (Figure 3C). These alterations were even more pronounced in patients with

advanced fibrosis (Figure 3C), correlating with an enhanced risk for developing HCC in F4 vs F2–F3.<sup>2,3</sup> Importantly, we found a clear correlation between transcriptomic and epigenomic changes of the identified oncogenes and TSGs, supporting the biological relevance of the findings (Figure 3D). Among these oncogenes was *SPHK1*, a lipid kinase mediating the phosphorylation of sphingosine to form SP1, which is a major regulator of cell apoptosis inhibition and proliferation promotion. SPHK1 and SP1 play key roles in the TNF $\alpha$  and nuclear factor  $\kappa$ B signaling pathways.<sup>19</sup> SPHK1 expression is increased and associated with tumor size and progression in patients with HCC.<sup>20</sup> Among the TSGs with significantly decreased H3K27ac level in HCV-infected patient livers were *PTPRD*, *TSC2*, and the major regulator of DNA repair, *BRCA1*. PTPRD has been identified as a candidate tumor suppressor in the liver impaired by HCV infection.<sup>21</sup> TSC2 has been reported to be a negative regulator of the mTOR signaling pathway. Its down-regulation is associated with metabolic defects, liver disease progression, and carcinogenesis.<sup>7</sup> Collectively, the overexpressed oncogenes and down-regulated TSGs that are enriched or decreased for the H3K27ac mark in chronic HCV infection, respectively, are involved in processes that favor carcinogenesis.

To further confirm that the persistent H3K27ac changes are linked to HCC risk, we referred to the genes of the recently reported 186-gene prognostic liver signature (PLS) and a 32-gene subset thereof for predicting liver disease progression, HCC development, and death for all HCC etiologies.<sup>9,17,22</sup> We analyzed functional links, that is, commonly shared pathways in MsigDB, among the 32-gene set, the 2193 genes with persistent epigenetic and transcriptional modifications, and the hallmarks of cancer.<sup>23</sup> We found that 1411 of the identified genes are closely connected to the PLS through shared pathways. Then, we assigned categories related to the hallmarks of cancer to the deregulated genes to understand the pathophysiologic impact of chronic HCV infection. Our analyses showed that approximately 900 genes of the genes with epigenetic modifications are directly linked with carcinogenesis. A network of these genes associated with at least 1 hallmark of cancer is shown in Figure 3E.

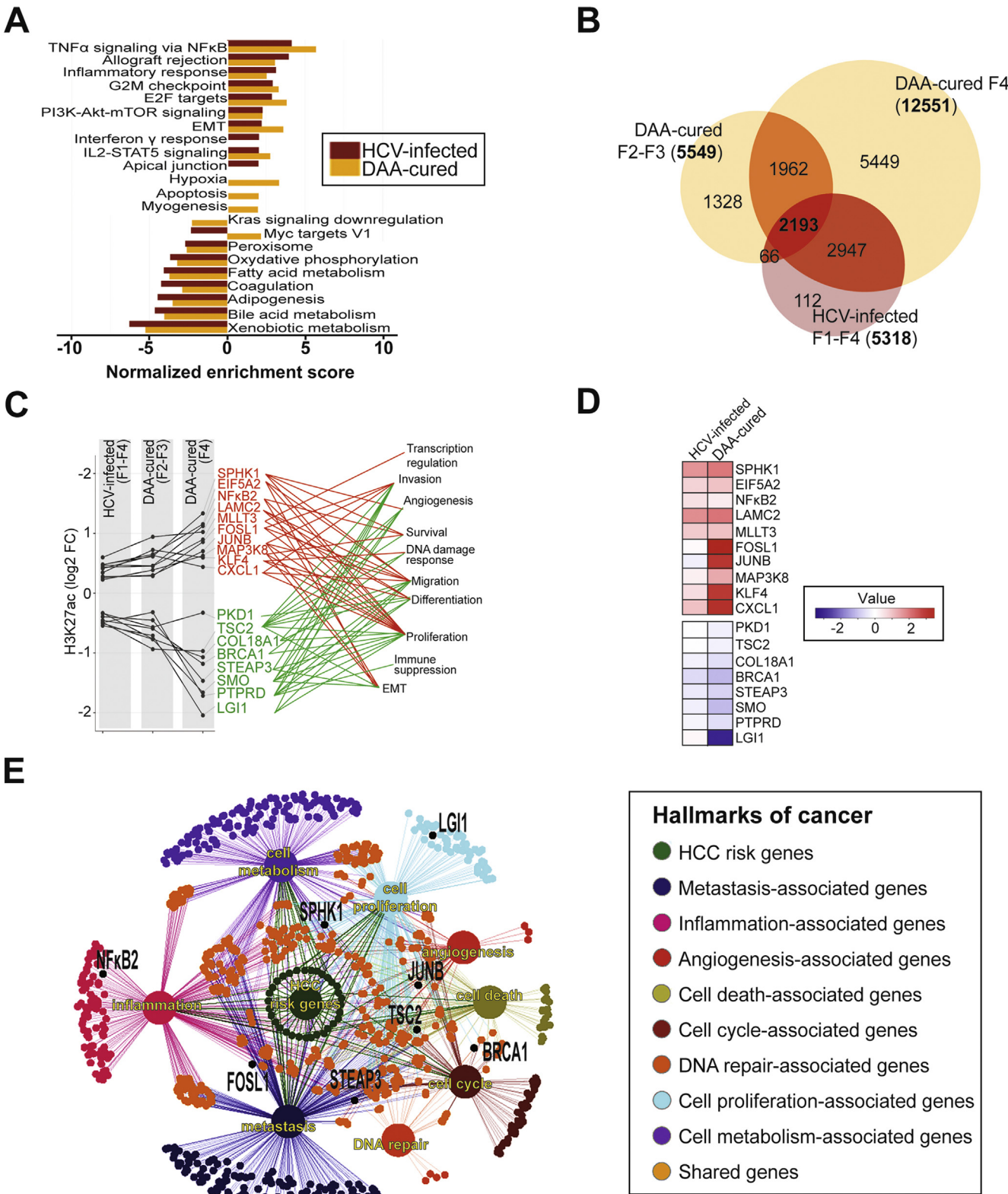
Next, we investigated whether H3K27ac alterations persist in cancer tissues after cure. We performed pairwise comparison of HCC and adjacent nontumorous tissue from the individual DAA-treated patient. We found a genome-wide H3K27ac enrichment in adjacent nontumorous and in tumorous tissues compared with noninfected samples (Figure 4). Deeper analysis showed that 52% of H3K27ac enriched genes are specific to tumorous tissues, 31% are

**Figure 2.** HCV-infection induces specific epigenetic changes in the liver of HCV-infected patients. (A) PCA for control, noninfected, HCV-infected, DAA-cured, IFN-cured, HBV-infected, and NASH patient samples. Comparative analysis of epigenetic modifications separated based on fibrosis score along the primary component (dimension 1). (B) H3K27ac modifications among HCV-infected patients correlate (Spearman rank correlation coefficients and *P* values) with H3K27ac modifications among NASH or HBV-infected patients. Common H3K27ac modifications were analyzed. Prognostic association of hepatic gene expression was determined by using Cox score for time to overall death in a cohort of patients as previously described.<sup>17</sup> (C) HCV-induced and persistent epigenetic changes after DAA cure in patient-derived liver tissue are associated with a decreased survival and death. H3K27ac modifications among HCV-infected correlate with persistent H3K27ac modifications among DAA-cured patients. (D) H3K27ac modifications correlate with significantly differentially expressed genes in HCV-infected and DAA-cured patients.

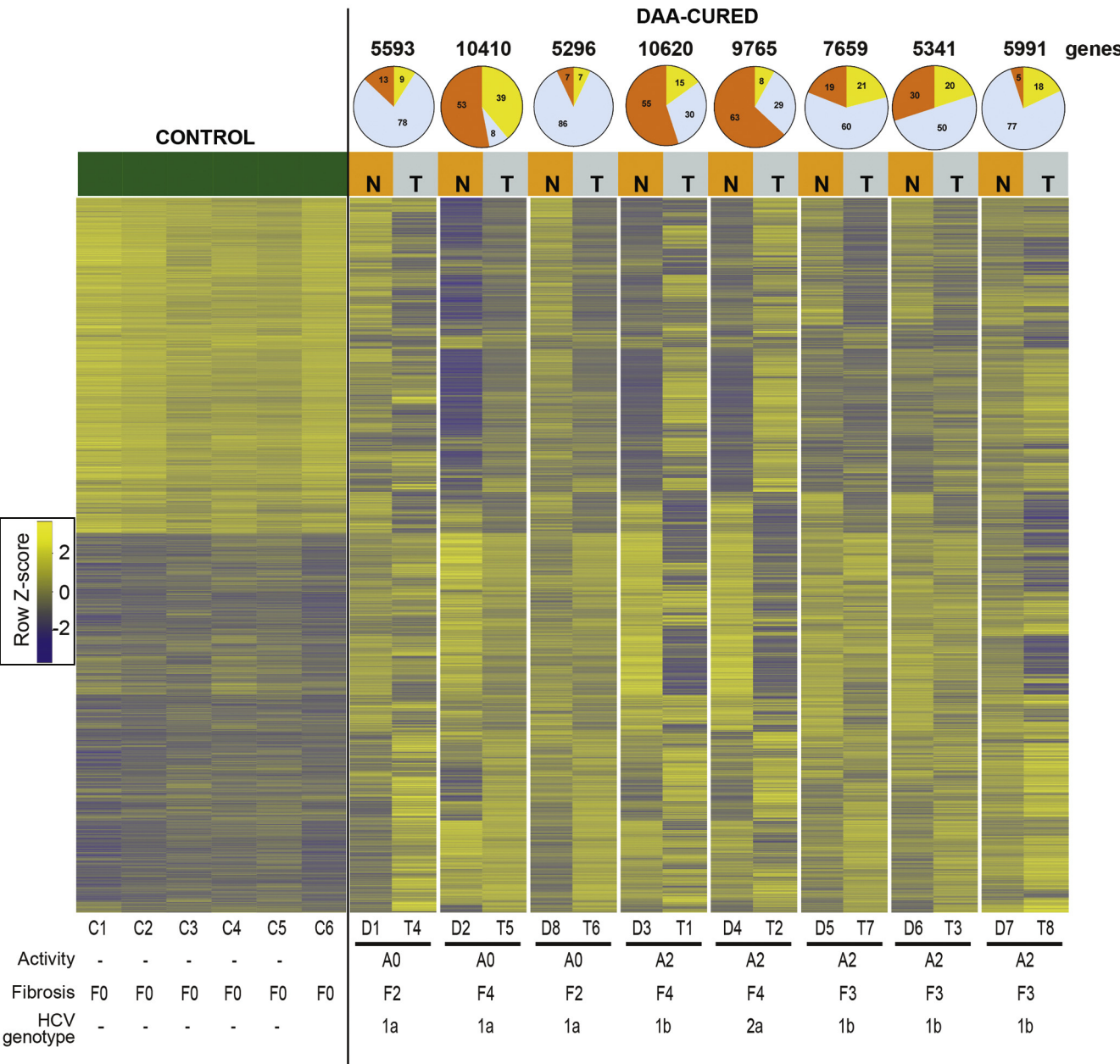


specific to adjacent nontumorous tissues, and 17% are common to the paired tissue. These data suggest that epigenetic alterations persist from advanced fibrosis to HCC and therefore could play a pathogenic role in

hepatocarcinogenesis before and after cure. Furthermore, the presence of epigenetic modifications in adjacent tumor tissue suggests that the epigenetic modifications might precede hepatocarcinogenesis.

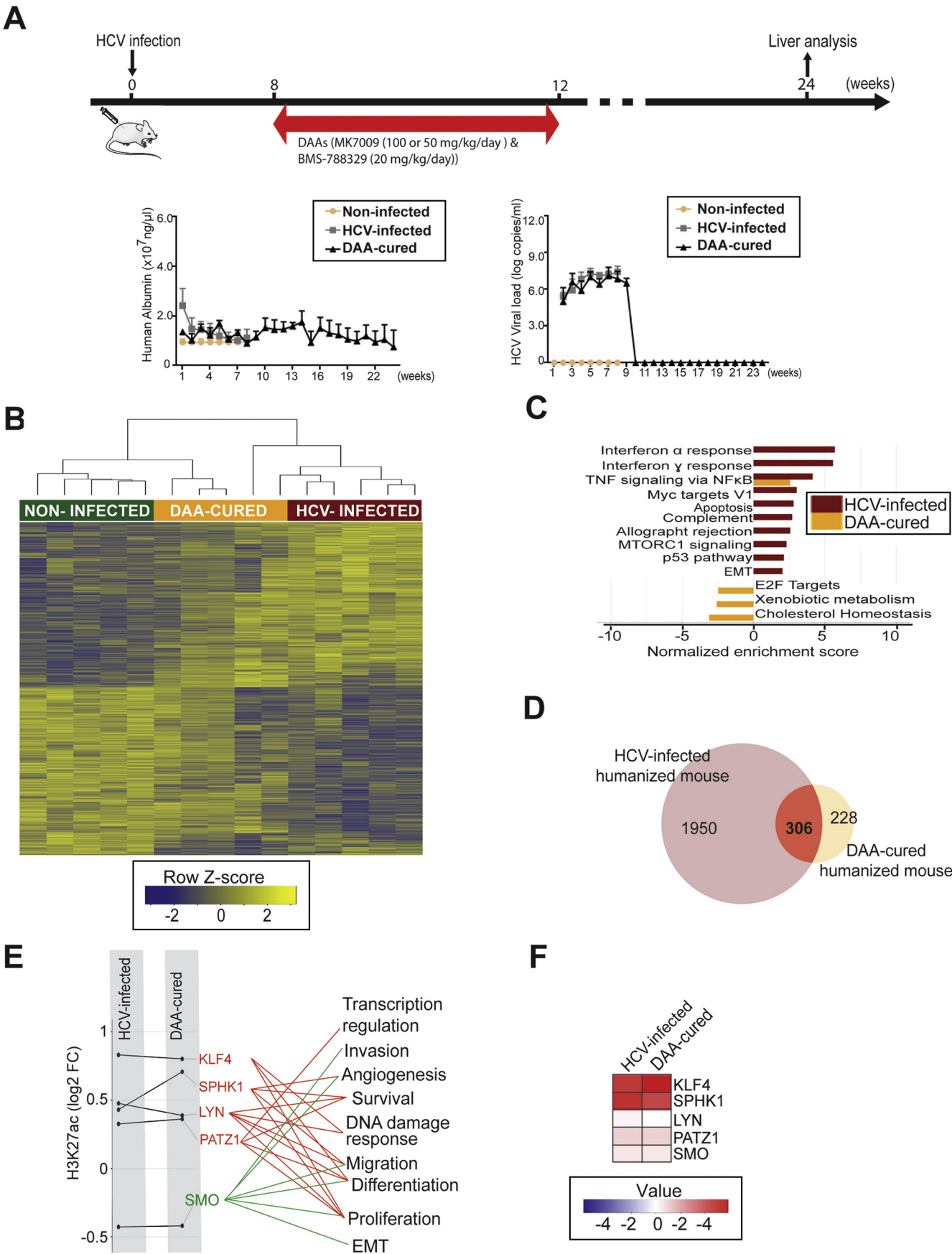






**Figure 4.** HCV-induced epigenetic changes persisting after DAA-based cure are present in the tumor tissue of patients with DAA-cured HCC. H3K27ac modifications from patient-derived resections of tumor and nontumor adjacent paired tissue samples. Similar to the analysis shown in Figure 1B, we performed an unsupervised clustering of normalized read counts in ChIP-Seq peaks of 7609 genes linked with significant ( $q < 0.05$ ) H3K27ac modifications in DAA-cured adjacent ( $n = 8$ ) or paired-tumor ( $n = 8$ ) tissues vs noninfected control patients ( $n = 6$ ). The proportions (percentages) of common (yellow) or distinct genes associated with changes in H3K27ac levels in tumor (blue) or nontumor paired-adjacent tissues (orange) are represented as a pie chart. N, nontumor; T, tumor.

**Figure 3.** Pathway analysis of epigenetic and transcriptional reprogramming in HCV-infected patients unravels candidate genes driving carcinogenesis after DAA cure. (A) Hallmark pathways significantly enriched for H3K27ac modifications in infected ( $n = 18$ ) or/and DAA-cured ( $n = 8$ ) compared with control ( $n = 6$ ) patient samples. A large overlap of enriched pathways persists in DAA-cured patients. (B) Venn diagram showing HCC risk gene candidates as the overlap of significantly modified genes in HCV-infected (F1–F4) and DAA-cured (F2–F3 and F4) patients derived from the ChIP-Seq experiment shown in Figure 1B. (C) Oncogenes (red) and TSGs (green) from the 2193 potential HCC risk gene candidates, with their biological functions indicated. (D) Heat map depicting transcriptional changes of the oncogenes and TSGs described in C in HCV-infected and DAA-cured patients. (E) Genes with persistent HCV-induced H3K27ac modifications after DAA cure, linked with the 32-gene prognostic liver signature predicting HCC in HCV-infected patients,<sup>9,17</sup> and overlapped with the hallmarks of cancer. Oncogenes shown in D are highlighted in black. This network includes 910 potential HCC risk gene candidates, highlighting a strong enrichment for modifications linked to carcinogenesis. EMT, epithelial–mesenchymal transition; IL2, interleukin 2; PI3K, phosphoinositide 3-kinase; STAT5, signal transducer and activator of transcription 5.



### Identification of HCV-Specific Epigenetic and Transcriptional Modifications That Are Independent of Inflammation and Fibrosis Using a Human Liver Chimeric Mouse Model

In the HCV-infected patient livers, epigenetic and transcriptional changes are most likely due to direct HCV-hepatocyte interactions and indirect mechanisms caused by chronic inflammation and fibrosis. Furthermore, our analysis is based on bulk tissue in which hepatocyte-related changes are difficult to distinguish from those in non-parenchymal cells. To clarify which fraction of the observed changes is dependent on HCV-hepatocyte interactions, we applied an HCV-permissive human liver chimeric mouse model.<sup>13</sup> In this model HCV efficiently infects the engrafted human hepatocytes without detectable liver fibrosis and inflammation. Moreover, human-specific sequencing reads in the ChIP-Seq pipeline are hepatocyte related because in liver bulk tissue only engrafted hepatocytes are of human origin. HCV-infected animals were cured using a combination of DAAs. Measurements of human albumin and HCV viral load in animals confirmed the viability of the engrafted hepatocytes and viral cure, respectively (Figure 5A). Similar to the findings in patients, we observed significant changes in H3K27ac levels in HCV-infected mice persisting after DAA cure (Figure 5B). Kyoto Encyclopedia of Genes and Genomes network analysis showed that pathways of genes showing epigenetic alterations included TNF signaling by nuclear factor  $\kappa$ B, IFN $\alpha/\gamma$  responses, complement, apoptosis, and mTOR signaling (Figure 5C). We found a persistence of TNF signaling through the nuclear factor  $\kappa$ B pathway, whereas the other HCV-induced pathways (ie, apoptosis, mTORC1 signaling, and IFN $\alpha/\gamma$  response) were restored to basal level after DAA-mediated cure (Figure 5C).

By intersecting genes associated with significant H3K27ac modifications from infected and cured mice, we identified 306 genes with persistent H3K27ac modifications after cure (Figure 5D and Supplementary Table 2). We found *SPHK1* and *KLF4* oncogenes and *SMO* TSGs, previously identified in patient samples (Figure 3C), to be associated with increased or decreased level of H3K27ac, respectively, in DAA-cured mice (Figure 5E), supporting the biological relevance of the findings in humanized mice. Similar to the results obtained in patients, we found a strong correlation between transcriptomic and epigenomic changes (Figure 5F).

Next, we identified HCV-specific epigenetic modifications in hepatocytes that are associated with HCC development by integrative analysis of epigenomic and transcriptomic data from patient and mouse liver samples. A comparative analysis of genes with persistent H3K27ac modifications in patients and mice showed a set of 65 commonly modified genes ( $P = 2.94 \times 10^{-9}$ ; Figure 6A). Further analysis identified that some of these 65 genes have their transcripts significantly correlated to epigenetic changes after DAA cure in patients and humanized mice. We ranked their transcript expression based on the FC relative to the noninfected samples. This approach identified 38 genes that were enriched for H3K27ac and that are associated with a significant positive FC of their transcripts after HCV infection and DAA cure compared with noninfected samples (Figure 6B). We further studied the biological function of these 38 genes by performing gene set analysis and found that they are associated to KRas, TNF $\alpha$ , and interleukin 2 and signal transducer and activator of transcription 5 signaling or to p53, epithelial-mesenchymal transition, apoptosis, glycolysis, and inflammation pathways (Supplementary Figure 3). Because they were identified by integrative analysis of data from patients and immunodeficient humanized mice, we hypothesize that inflammation-related genes derive from the innate response of infected hepatocytes.

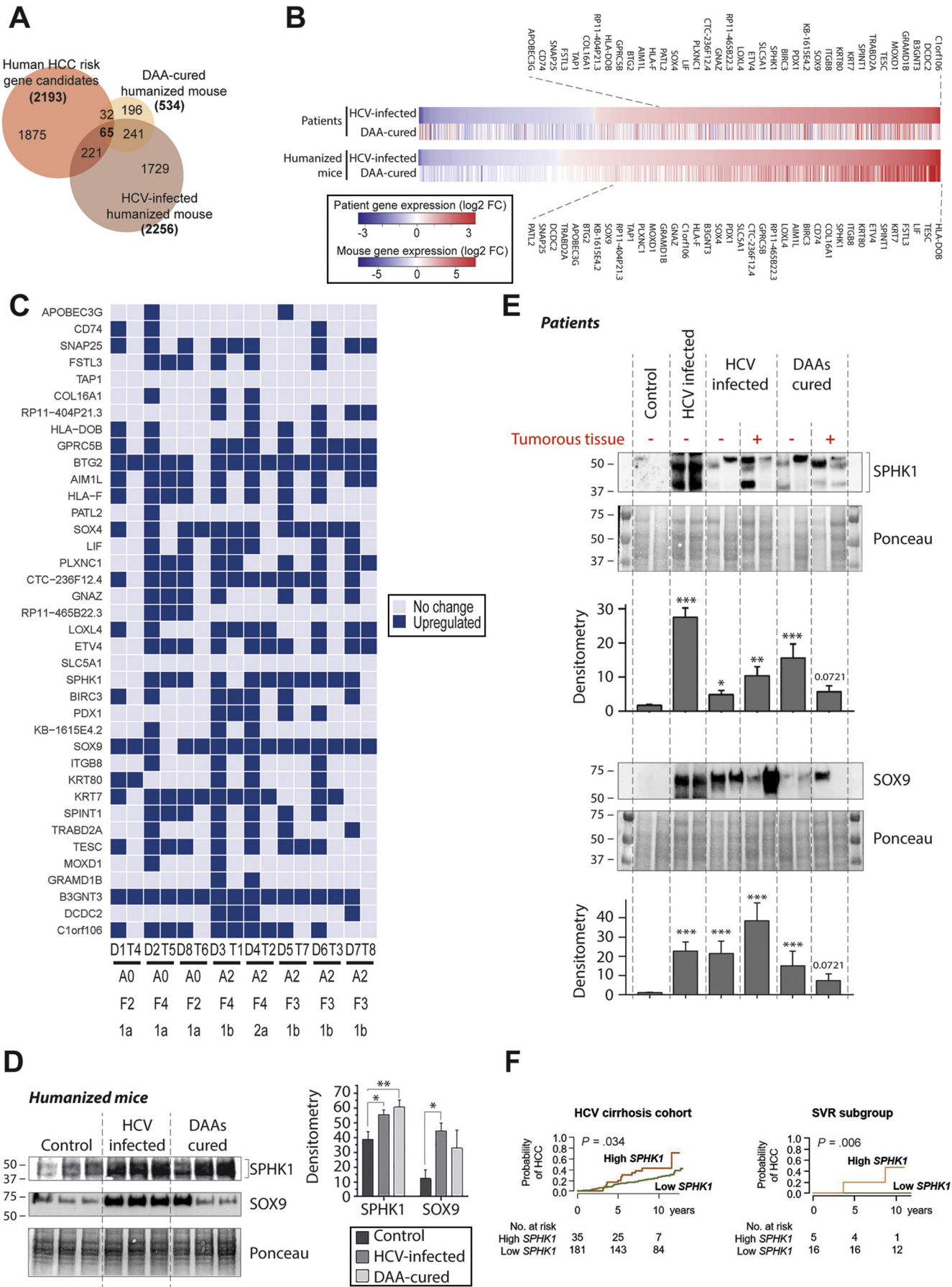
To obtain further evidence that these alterations play a role in hepatocarcinogenesis after cure, we compared their H3K27ac levels in paired liver tissues of nontumorous adjacent and HCC. We found that most of them already harbored changes in the nontumorous sample that remained in HCC tissue (Figure 6C). For instance, changes were observed for *SPHK1* in nontumorous tissue in 7 of 8 patients and persisted in HCC tissue in 4 patients. H3K27ac modifications in *SOX9*, a gene that is associated to ductular reaction, was found in nontumorous tissue in all DAA-cured patients and remained in HCC tissue in 7 of 8 patients.

### HCV and Hepatocyte-Specific Epigenetic Modifications Translate Into Liver Protein Expression Changes and Are Associated With HCC Development in HCV Cirrhosis and SVR Cohorts

To further validate the biological relevance of HCV-induced epigenetic and transcriptional changes, we studied

**Figure 5.** Analysis of H3K27ac changes in livers of HCV-infected humanized mice identifies virus-specific modifications in human hepatocytes. (A) Our experimental setup: uPA-SCID mice were infected with HCV for 8 weeks and cured with a combination of DAAs MK7009 (50 or 100 mg/kg/d) and BMS-788329 (20 mg/kg/d) for 16 weeks. Livers were analyzed at week 24 by ChIP-Seq and RNA-Seq. Human albumin level (left) and HCV viral load (right) were measured to monitor functional engrafted human hepatocytes and HCV clearance after DAA treatment, respectively. (B) Unsupervised clustering of normalized read counts in ChIP-Seq peaks of 2483 genes linked with significant ( $q < 0.05$ ) H3K27ac modifications in HCV-infected ( $n = 5$ ) or DAA-cured ( $n = 5$ ) vs noninfected control ( $n = 5$ ) mice. (C) Hallmark pathways significantly enriched for H3K27ac modifications in infected ( $n = 5$ ) or/and DAA-cured ( $n = 5$ ) compared with noninfected ( $n = 5$ ) mice samples. A significant overlap of enriched pathways persists in DAA-cured mice. (D) Venn diagram showing the HCV-induced and persistent genes with H3K27ac changes as the overlap of significantly modified genes in HCV-infected and DAA-cured mice derived from the ChIP-Seq experiment shown in B. (E) Oncogenes (red) and TSGs (green) with persistent HCV-induced H3K27ac modifications identified in the 306 HCV-induced and persistent genes with H3K27ac changes, with their biological functions indicated. (F) Heat map depicting transcriptional changes of the oncogenes and TSGs described in E in HCV-infected humanized and DAA-cured mice. EMT, epithelial-mesenchymal transition; NF $\kappa$ B, nuclear factor  $\kappa$ B.





whether the expression of the identified genes correlates with corresponding protein abundance. We quantified the protein expression of *SPHK1* and *SOX9* genes by immunoblotting in patient and mouse liver samples (Figure 6D and E and Supplementary Figures 4–6). We found increased SPHK1 and SOX9 protein levels at HCV infection that remained increased after DAA cure. Importantly, by comparing pairwise liver tissue from adjacent nontumorous areas and HCC, we found that the expression of SPHK1 and SOX9 were already increased in adjacent nontumorous tissue (Figure 6D and E), suggesting that the up-regulation of these proteins preceded tumor development.

To assess the potential of the expression of these genes as biomarkers to predict HCC risk, we assessed the association of *SPHK1* expression with the long-term probability to develop HCC over a decade in a cohort of patients with HCV cirrhosis ( $n = 216$ ), among which a subset of patients achieved SVR ( $n = 21$ ). We found that high expression of *SPHK1* is significantly associated with HCC risk in the 2 cohorts ( $P < .034$  for HCV cirrhosis and  $P < .006$  for SVR; Figure 6F), identifying a potential predictor of HCC risk post SVR.

## Discussion

Our study exposes a previously undiscovered paradigm showing that chronic HCV infection induces H3K27ac modifications that are associated with HCC risk and that persist after HCV cure. Thus far, only limited data have shown that HCV infection can induce epigenetic changes.<sup>24</sup> Previous attempts to connect specific histone marks to HCC development were inconclusive because of semi-quantitative approaches.<sup>25,26</sup> For the first time, our study provides an integrative genome-wide approach that combines analyses in patient liver tissue and a humanized animal model.

Long-term epigenetic alterations also were observed after Epstein-Barr virus infection<sup>27</sup> or after transient hyperglycemia.<sup>28</sup> Indeed, latent Epstein-Barr infection triggered persistent epigenetic reprogramming, possibly resulting in the establishment of immortal growth and cancer, whereas transient hyperglycemia resulted in persistent enrichment of H3K4me1 on the *p65* gene

promoter and subsequently in oxidative stress and increased cancer risk. Importantly, these data suggest that persistent epigenetic changes also can occur through environmental changes, independently from direct viral infection.

Epigenetic changes in patient liver tissue can result from infected hepatocytes and from virus-induced inflammatory or fibrotic responses in the liver microenvironment. Interestingly, PCA showed a clear correlation of epigenetic changes with fibrosis stage (Figure 2A), suggesting that HCV-induced histone modifications and fibrogenesis are interdependent from the progression of liver disease. Indeed, epigenetic changes are considered as orchestrating fibrogenesis,<sup>29</sup> including the activation of hepatic stellate cells. In contrast, the induction of fibrosis triggers a liver response to injury, implicating the epigenetic machinery to mediate the activation of dedicated genes,<sup>30</sup> and thereby enhancing HCV-established epigenetic changes. Because distinct epigenetic changes were found in patient liver tissue and humanized mouse liver tissue (Figures 3 and 5), where no necro-inflammatory response or fibrosis is present, it is likely that a fraction of the observed changes is caused by direct HCV-hepatocyte interactions. Collectively, our results suggest that direct virus-hepatocyte interactions and indirect mechanisms, such as disease-induced fibrosis mediated by the liver non-parenchymal cells, contribute to the observed epigenetic changes in the livers of HCV-infected patients. Importantly, our data provide a previously undiscovered mechanism for persistent HCC risk after DAA cure in advanced fibrosis and could explain why a small number of patients develop HCC even in the absence of fibrosis.<sup>2</sup> However, we point out that this mechanism is not exclusive, and many other factors most likely contribute to hepatocarcinogenesis after cure.

Although we did not perform extensive functional studies, our data provide evidence that HCV-induced H3K27ac modifications on specific genes are causal factors for HCC risk after DAA cure. Our hypothesis is strongly supported by (1) altered expression of genes known to promote and drive carcinogenesis, (2) the correlation of epigenetic changes with a clinical Cox score for overall death and a HCC risk score,<sup>17</sup> (3) the positive correlation between

**Figure 6.** Intersection of ChIP-Seq and RNA-Seq analyses from livers of patients and humanized mice uncovers HCV-induced persistent epigenetic changes associated with HCC risk after SVR. (A) Venn diagram showing the overlap of H3K27ac modifications between the human HCC risk gene candidates and significantly modified genes in HCV-infected and DAA-cured mice derived from the ChIP-Seq experiments shown in Figures 1B and 5B, respectively. (B) Expression data of genes with significant H3K27ac changes from livers of HCV-infected and DAA-cured patients ( $n = 32$ ) and mice ( $n = 15$ ) were intersected to uncover common genes with HCV-induced and persistent epigenetic and transcriptional changes after DAA. (C) Presence of epigenetic modifications on the 38 identified genes in pairwise liver tissues from DAA-cured patients. H3K27ac modifications (vs control liver samples) were assessed on the corresponding genes in nontumorous adjacent and HCC liver tissues from DAA-cured patients. *Dark blue squares* represent increased H3K27ac changes and *light blue squares* represent unchanged status. (D) Analysis of protein level of SPHK1 and SOX9 protein in control, HCV-infected, and DAA-cured mice by western blot. (E) Analysis of SPHK1 and SOX9 protein levels in control ( $n = 7$ ), HCV-infected (non-HCC and HCC;  $n = 8$ ) and DAA-cured (non-HCC and HCC;  $n = 8$ ) patients by western blot. One representative gel of 4 is shown. Graphs show quantification of western blot intensities in arbitrary units normalized to total protein level (Ponceau staining). Results show mean  $\pm$  standard error of the mean of integrated blot densities. (F) Probability of HCC development according to the gene expression level of SPHK1 among 216 patients with HCV-induced cirrhosis or 21 patients with HCC occurrence after HCV cure.

the magnitude of epigenetic changes and fibrosis stage, which is the strongest clinical risk factor for HCC,<sup>6</sup> and (4) the presence of H3K27ac modifications in HCC tumors of the same patients. Collectively, these findings suggest that epigenetic modifications precede hepatocarcinogenesis. Among the identified genes, functional knockout of *SOX9* has been reported to decrease liver cancer cell growth,<sup>31</sup> and *SPHK1* deletion decreased diethyl-nitrosamine-induced liver cancer in mice,<sup>32</sup> whereas ETS translocation variant 4 (*ETV4*) is up-regulated and is associated to HCC progression.<sup>33</sup> Importantly, extended analysis in additional cohorts showed that those genes that were epigenetically changed by HCV infection and that persisted after DAA cure predicted HCC risk in cohorts of patients with HCV cirrhosis and SVR (Figure 6C). Although we do not have experimental evidence that HCV-mediated modulation of *SPHK1* or *SOX9* gene expression is sufficient to promote cancer, our data combined with published knowledge on the role of these proteins in cancer biology<sup>31,32</sup> nevertheless suggest that *SPHK1* and *SOX9*, among additional tumor-associated proteins, participate in HCV-induced HCC. This strongly supports the hypothesis that H3K27ac alterations of the identified genes precede HCC onset.

Other well-known causes for HCC development are chronic HBV infection and NASH.<sup>2</sup> Interestingly, we found that H3K27ac modifications also are present in these etiologies (Figures 1B and 2B). In-depth analyses including PCA (Figures 2A and Supplementary Figure 2) showed etiology-independent and etiology-specific epigenetic profiles in liver disease.

Because of the difficulty of obtaining liver tissue after HCV cure, which was available only for patients with concomitant HCC, the number of patient tissues is limited. Because it is impossible to obtain healthy liver tissue for ethical reasons, the control samples from patients with nonviral minimal liver disease or adjacent tissue from patients undergoing surgery for metastasis for colorectal cancer exhibited heterogeneity. Furthermore, the H3K27ac mark constitutes only a part of the epigenetic gene regulation program. Nevertheless, the robust results obtained by clustering and statistical analyses combined with consistent results from patients of different cohorts and clinical centers and confirmation of the key concept in humanized mouse engrafted with hepatocytes from the same donor and infected with the same viral inoculum allowed arresting conclusions.

HCC is often asymptomatic and thus remains undiagnosed until the late stage. Therefore, there is an urgent medical need for biomarkers to predict HCC risk. A large body of literature has shown the association between the human epigenome and cancer development.<sup>34</sup> In this study, showing that HCV induces persistent epigenetic alterations after DAA cure provides a unique opportunity to uncover novel biomarkers for HCC risk, that is, from plasma through the detection of epigenetic changes of histones bound to circulating DNA complexes. Furthermore, by uncovering virus-induced epigenetic changes as therapeutic targets, our findings offer novel perspectives for HCC prevention—a key unmet medical need.

## Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at [www.gastrojournal.org](http://www.gastrojournal.org), and at <https://doi.org/10.1053/j.gastro.2019.02.038>.

## References

1. Ryerson AB, Ehemann CR, Altekruse SF, et al. Annual report to the nation on the status of cancer, 1975–2012, featuring the increasing incidence of liver cancer. *Cancer* 2016;122:1312–1337.
2. Fujiwara N, Friedman SL, Goossens N, Hoshida Y. Risk factors and prevention of hepatocellular carcinoma in the era of precision medicine. *J Hepatol* 2018;68:526–549.
3. Boldanova T, Suslov A, Heim MH, Necseulea A. Transcriptional response to hepatitis C virus infection and interferon-alpha treatment in the human liver. *EMBO Mol Med* 2017;9:816–834.
4. El-Serag HB, Kanwal F, Richardson P, Kramer J. Risk of hepatocellular carcinoma after sustained virological response in veterans with hepatitis C virus infection. *Hepatology* 2016;64:130–137.
5. Kanwal F, Kramer J, Asch SM, et al. Risk of hepatocellular cancer in HCV patients treated with direct-acting antiviral agents. *Gastroenterology* 2017;153:996–1005 e1.
6. van der Meer AJ, Feld JJ, Hofer H, et al. Risk of cirrhosis-related complications in patients with advanced fibrosis following hepatitis C virus eradication. *J Hepatol* 2017;66:485–493.
7. Sajankila SP, Manthena PV, Adhikari S, et al. Suppression of tumor suppressor Tsc2 and DNA repair glycosylase Nth1 during spontaneous liver tumorigenesis in Long-Evans Cinnamon rats. *Mol Cell Biochem* 2010;338:233–239.
8. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;9:R137.
9. Nakagawa S, Wei L, Song WM, et al. Molecular liver cancer prevention in cirrhosis by organ transcriptome analysis and lysophosphatidic acid pathway inhibition. *Cancer Cell* 2016;30:879–890.
10. Hoshida Y, Villanueva A, Sangiovanni A, et al. Prognostic gene expression signature for patients with hepatitis C-related early-stage cirrhosis. *Gastroenterology* 2013;144:1024–1030.
11. Jones PA, Issa J-PJ, Baylin S. Targeting the cancer epigenome for therapy. *Nat Rev Genet* 2016;17:630.
12. Cai MY, Hou JH, Rao HL, et al. High expression of H3K27me3 in human hepatocellular carcinomas correlates closely with vascular invasion and predicts worse prognosis in patients. *Mol Med* 2011;17:12–20.
13. Uchida T, Hiraga N, Imamura M, et al. Elimination of HCV via a non-ISG-mediated mechanism by vaniprevir and BMS-788329 combination therapy in human hepatocyte chimeric mice. *Virus Res* 2016;213:62–68.
14. Schmidl C, Rendeiro AF, Sheffield NC, Bock C. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat Methods* 2015;12:963–965.



15. Harrow J, Denoeud F, Frankish A, et al. GENCODE: producing a reference annotation for ENCODE. *Genome Biol* 2006;7(Suppl 1):S41–S49.
16. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–15550.
17. Hoshida Y, Villanueva A, Kobayashi M, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med* 2008;359:1995–2004.
18. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? *Mol Cell* 2013;49:825–837.
19. Rohrbach T, Maceyka M, Spiegel S. Sphingosine kinase and sphingosine-1-phosphate in liver pathobiology. *Crit Rev Biochem Mol Biol* 2017;52:543–553.
20. Cai H, Xie X, Ji L, et al. Sphingosine kinase 1: A novel independent prognosis biomarker in hepatocellular carcinoma. *Oncol Lett* 2017;13:2316–2322.
21. Van Renne N, Roca Suarez AA, Duong FHT, et al. miR-135a-5p-mediated downregulation of protein tyrosine phosphatase receptor delta is a candidate driver of HCV-associated hepatocarcinogenesis. *Gut* 2018;67:953–962.
22. King LY, Canasto-Chibuque C, Johnson KB, et al. A genomic and clinical prognostic index for hepatitis C-related early-stage cirrhosis that predicts clinical deterioration. *Gut* 2015;64:1296–1302.
23. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646–674.
24. Okamoto Y, Shinjo K, Shimizu Y, et al. Hepatitis virus infection affects DNA methylation in mice with humanized livers. *Gastroenterology* 2014;146:562–572.
25. Hayashi A, Yamauchi N, Shibahara J, et al. Concurrent activation of acetylation and tri-methylation of H3K27 in a subset of hepatocellular carcinoma with aggressive behavior. *PLoS One* 2014;9:e91330.
26. Ruthenburg AJ, Allis CD, Wysocka J. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* 2007;25:15–30.
27. Scott RS. Epstein-Barr virus: a master epigenetic manipulator. *Curr Opin Virol* 2017;26:74–80.
28. El-Osta A, Brasacchio D, Yao D, et al. Transient high glucose causes persistent epigenetic changes and altered gene expression during subsequent normoglycemia. *J Exp Med* 2008;205:2409–2417.
29. Moran-Salvador E, Mann J. Epigenetics and liver fibrosis. *Cell Mol Gastroenterol Hepatol* 2017;4:125–134.
30. Bae WK, Kang K, Yu JH, et al. The methyltransferases enhancer of zeste homolog (EZH) 1 and EZH2 control hepatocyte homeostasis and regeneration. *FASEB J* 2015;29:1653–1662.
31. Richtig G, Aigelsreiter A, Schwarzenbacher D, et al. SOX9 is a proliferation and stem cell factor in hepatocellular carcinoma and possess widespread prognostic significance in different cancer types. *PLoS One* 2017;12:e0187814.
32. Chen J, Qi Y, Zhao Y, et al. Deletion of sphingosine kinase 1 inhibits liver tumorigenesis in diethylnitrosamine-treated mice. *Oncotarget* 2018;9:15635–15649.
33. Kim E, Kim D, Lee J-S, et al. Capicua suppresses hepatocellular carcinoma progression by controlling the ETV4–MMP1 axis. *Hepatology* 2018;67:2287–2301.
34. Wilson CL, Mann DA, Borthwick LA. Epigenetic reprogramming in liver fibrosis and cancer. *Adv Drug Deliv Rev* 2017;121:124–132.

Author names in bold designate shared co-first authorship.

Received December 10, 2018. Accepted February 27, 2019.

#### Reprint requests

Address requests for reprints to: Thomas F. Baumert, MD, INSERM U1110, University of Strasbourg, 3 rue Koeberlé, F-67000 Strasbourg, France. e-mail: [thomas.baumert@unistra.fr](mailto:thomas.baumert@unistra.fr).

#### Acknowledgments

We acknowledge the Centre de Ressources Biologiques (Biological Resource Centre; Strasbourg, France) for the management of patient-derived liver tissues. We acknowledge the work of the IGBMC high-throughput sequencing facility and the Quantitative Genomics Facility at the Department of Biosystems Science and Engineering at the scientific central facilities of ETH Zurich. The IGBMC high-throughput sequencing facility is a member of the France Génomique consortium (ANR10-INBS-09-08). The laboratory of Irwin Davidson is an *équipe labellisée* of the Ligue Nationale contre le Cancer.

Author contributions: Thomas F. Baumert initiated and coordinated the study. Thomas F. Baumert, Nouridine Hamdane, Frank Jühling, Mirjam B. Zeisel, Yujin Hoshida, and Bryan C. Fuchs designed experiments and analyzed data. Nouridine Hamdane, Emilie Crouchet, Christine Thumann, Marine A. Oudot, Clara Ponsolles, Armando Andres Roca Suarez, and Shen Li performed experiments. Frank Jühling performed computational analyses. Nabeel Bardeesy contributed to the concept and approach of the study. Christian Schmid and Christoph Bock performed sequencing of ChIPmentation experiments. Irwin Davidson, Houssein El Saghiere, Antonio Saviano, Naoto Fujiwara, Catherine Schuster, Atsushi Ono, and Yujin Hoshida analyzed data. Patrick Pessaux, Michio Imamura, Takuro Uchida, Hideki Ohdan, Hiroshi Aikata, Kazuaki Chayama, Tullio Piardi, Daniele Sommacale, François Habersetzer, Michel Doffoël, and Tujana Boldanova provided clinical liver tissue samples. Nouridine Hamdane, Frank Jühling, François H.T. Duong, Bryan C. Fuchs, Joachim Lupberger, Mirjam B. Zeisel, and Thomas F. Baumert wrote the manuscript.

#### Conflicts of interest

Authors declare no conflict of interest.

#### Funding

This work was supported by the ARC (Paris) and the Institut Hospitalo-Universitaire (Strasbourg; TheraHCC IHUARC IHU201301187 to Thomas F. Baumert), the Foundation of the University of Strasbourg and Roche Institute (HEPKIN to Thomas F. Baumert and Yujin Hoshida), the Agence Nationale de Recherches sur le Sida et les Hépatites Virales (2017/1633 to Thomas F. Baumert), the U.S. Department of Defense (W81XWH-16-1-0363 to Yujin Hoshida and Thomas F. Baumert), the Cancéropôle du Grand-Est (AAP Emergence 2017 to Joachim Lupberger), the National Institutes of Health (DK099558 to Yujin Hoshida), and the Research Program on Hepatitis from the Japan Agency for Medical Research and Development, AMED (17fk0210104h0001 to Kazuaki Chayama). This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement 667273 (HEPCAR to Thomas F. Baumert and Joachim Lupberger). This project has received funding from the European Research Council under the European Union's Horizon 2020 Research and Innovation Program under grant 671231 (HEPCIR to Thomas F. Baumert and Yujin Hoshida). This work has been published under the framework of the LABEX ANR-10-LABX-0028\_HEPSYS, and PLAN CANCER 2014–2019 and benefits from a funding from the state managed by the French National Research Agency as part of the Investments for the Future Program, the French National Cancer Institute, and INSERM.

## Supplementary Methods

### *HCV Infection of Human Hepatocyte Chimeric Mouse and DAA Treatment*

cDNA-uPA<sup>+/+</sup>/SCID<sup>+/+</sup> (uPA/SCID) mice were created and human hepatocytes were transplanted as described previously.<sup>1</sup> Mice were intravenously inoculated with serum samples containing 10<sup>5</sup> HCV particles. The viremic serum was obtained from an HCV-infected (genotype 1b) DAA-naïve patient who provided written informed consent to participate in the study, according to the process approved by the ethical committee of the hospital and in accordance with the ethical guidelines of the 1975 Declaration of Helsinki. Blood sampling was done weekly, and serum samples were divided into small aliquots and stored in liquid nitrogen before measurement of HCV RNA. All animal protocols were performed in accordance with the Guide for the Care and Use of Laboratory Animals (<https://grants.nih.gov/grants/olaw/guide-for-the-care-and-use-of-laboratory-animals.pdf>). The experimental protocol was approved by the Ethics Review Committee for Animal Experimentation of the Graduate School of Biomedical Sciences at Hiroshima University (A14-195). Sixteen mice were divided into 3 groups: 6 mice were infected with HCV and treated with DAAs, 5 mice were infected with HCV but were not treated with DAAs, and 5 uninfected and untreated mice were used as controls. After the establishment of stable viremia, HCV-infected mice were treated with a combination of MK-7009 (vaniprevir; Merck Sharp & Dohme Corp, Kenilworth, NJ) and BMS-788329 (NS5A inhibitor; Bristol-Meyers Squibb, New York, NY) as described previously.<sup>2</sup> Elimination of HCV in 6 treated mice was confirmed by the absence of HCV viremia 12 weeks after cessation of therapy and by undetectable HCV RNA by reverse transcription-nested polymerase chain reaction from extracted liver tissue. Five viremic mice and 5 control mice were sacrificed at week 8. All liver samples were snap frozen and stored at -80°C before analysis.

### *Processing of Raw ChIPmentation Data*

Reads were aligned to the human genome (hg19) using HISAT2<sup>3</sup> reporting up to 100 alignments per read. Data from humanized mice were mapped similarly, but to an artificial genome consisting of all human (hg19) and mouse (mm10) chromosomes, and only reads mapping to human chromosomes were kept for further analysis. Sorting, indexing, and other basic operations on alignments were performed with samtools<sup>4</sup> and intersections of annotations and peaks with alignments were performed using bedtools intersect.<sup>5</sup> Peaks were called in uniquely mapped reads filtered for duplicates using MACS2<sup>6</sup> in standard mode and with corresponding input sequence data. Only samples with at least 10,000 peaks were used for further analyses. Peaks within all samples were intersected and used for counting reads if they overlapped in at least 2 samples. Close peak regions with a maximal distance of 500 bp were merged. Read counts of genes were defined as the sum of all reads in peak

regions overlapping the gene body or the promoter region, that is, the region up to 1500 bp ahead of the transcription start site.

### *Processing of RNA-Seq Data*

Raw reads of patient's samples had to be trimmed for primer and quality using cutadapt.<sup>7</sup> Reads were mapped using HISAT2<sup>3</sup> to the human genome hg19 (patients) or to hg19 and mm10 (humanized mice) as described earlier for raw ChIPmentation data. Reads were counted with htseq-count, and a differentially expression analysis was performed with DESeq2 applying GENCODE 19.<sup>8</sup> Reads were taken from our RNA-Seq experiments as described earlier and from external sources: RNA-Seq from infected vs control patients was taken from the GEO dataset GSE84346 (low ISG samples).

### *Pathway Enrichment and Correlation Analyses*

The full downstream ChIP-Seq analysis was based on read counts in ChIP-Seq peaks called as described earlier. Differentially modified genes (GENCODE 19 annotation) and log<sub>2</sub> FCs were identified using these peak read counts as input for edgeR.<sup>9</sup> Pathway enrichment analyses were performed using local javaGSEA with all gene sets included in MSigDB 6.0.<sup>10</sup> We used the pre-ranked version of javaGSEA and genes were ranked for *P* values of differential expression and modification analyses. Figures showing enriched pathways and gene sets, Spearman correlations, and oncogene log<sub>2</sub> FCs were drawn using ggplot2 and the R environment. Immune-related genes used for calculating correlations were selected from MSigDB by including only genes from pathways with the term "IMMUNE" in their title. Heat maps of gene expression and histone modifications were generated by applying the heatmap.2 function in combination with clustering through Spearman correlation included in the R package gplots. Gene network analysis was performed based on 3 MSigDB subsets: Hallmark gene sets, curated gene sets, and gene ontology gene sets. Genes were assigned with the hallmarks of cancer in case they were found in gene sets whose designation matches a corresponding term. Network figures were generated manually using Cytoscape.<sup>11</sup> Genes were defined as to be "connected to the PLS" in the case they shared at least 1 common pathway listed in MSigB 6.0 with at least 1 of the 32 PLS genes.

### *Western Blot and Antibodies*

The expression of SPHK1 and SOX9 proteins was assessed by western blot using polyclonal rabbit antibodies anti-SPHK1 (D1H1L; number 12071) and anti-SOX9 (D8G8H; number 82630) from Cell Signaling (Danvers, MA). Protein expression was quantified using ImageJ software. Because anti-SPHK1 antibody detects all 3 isoforms<sup>12</sup> of SPHK1 and it is only partially understood which isoform or which post-translational modification on the oncogene SPHK1 predominantly triggers carcinogenesis, all apparent bands were included in the densitometry analysis.

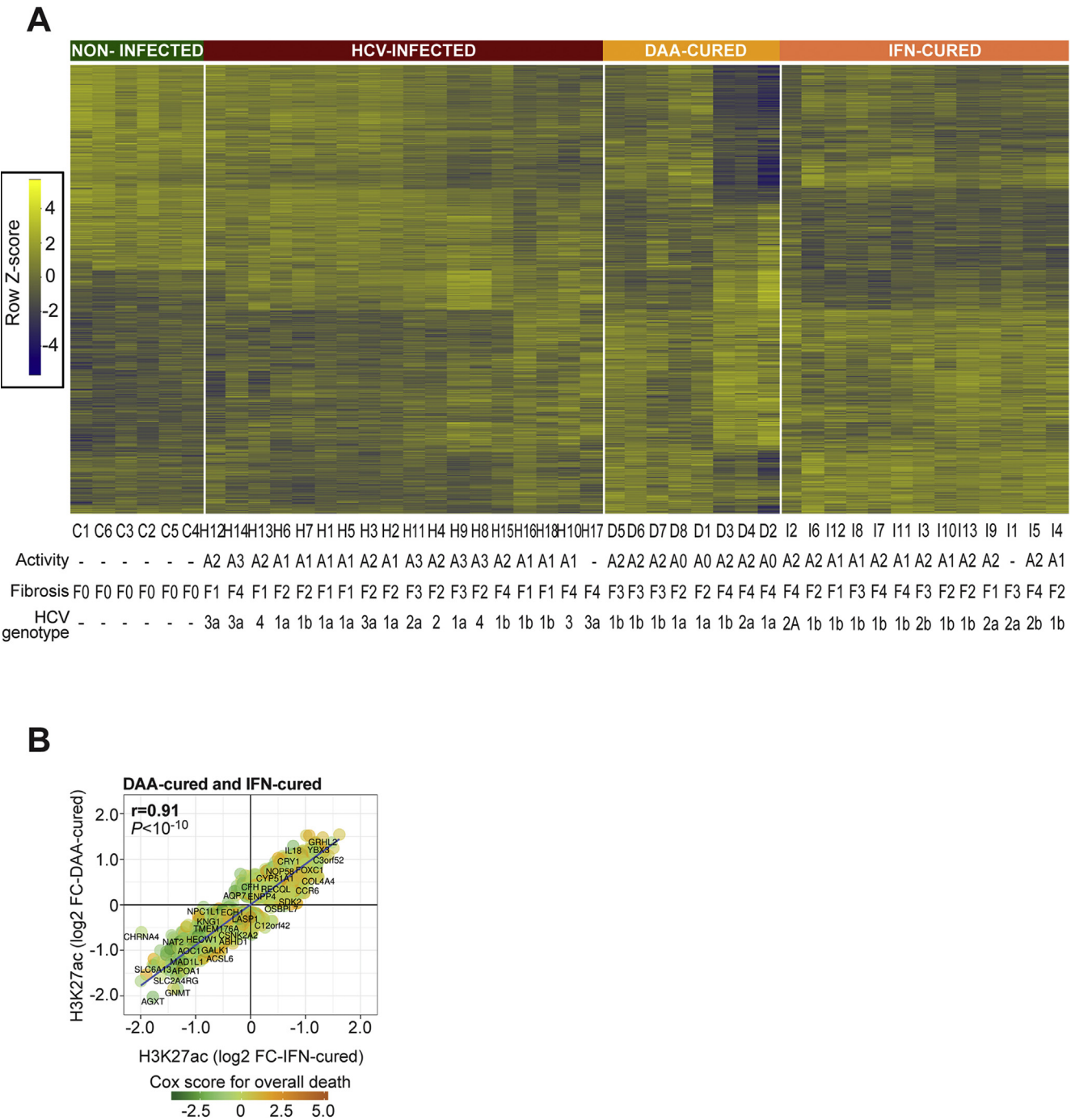


## References

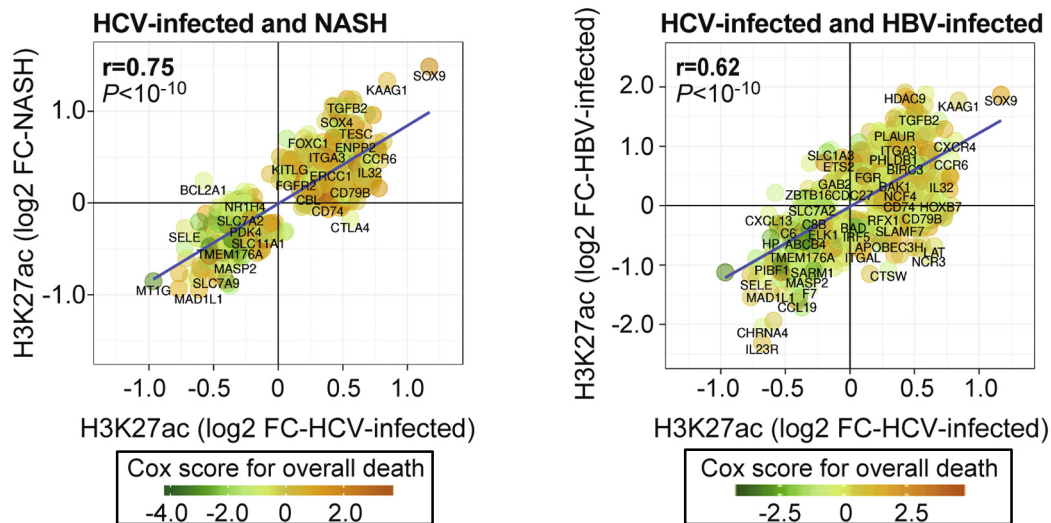
1. Tateno C, Kawase Y, Tobita Y, et al. Generation of novel chimeric mice with humanized livers by using hemizygous cDNA-uPA/SCID mice. *PLoS One* 2015; 10:e0142145.
2. Uchida T, Hiraga N, Imamura M, et al. Elimination of HCV via a non-ISG-mediated mechanism by sofosbuvir and BMS-788329 combination therapy in human hepatocyte chimeric mice. *Virus Res* 2016; 213:62–68.
3. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *NatMethods* 2015;12:357–360.
4. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics* 2011;27:2987–2993.
5. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841–842.
6. **Zhang Y, Liu T**, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;9:R137.
7. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011; <https://doi.org/10.14806/ej.17.1.200>.
8. **Harrow J, Denoeud F, Frankish A**, et al. GENCODE: producing a reference annotation for ENCODE. *Genome Biol* 2006;7(Suppl 1):S41–S49.
9. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–140.
10. **Subramanian A, Tamayo P**, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–15550.
11. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13:2498–2504.
12. **Hatoum D, Haddadi N**, Lin Y, et al. Mammalian sphingosine kinase (SphK) isoenzymes and isoform expression: challenges for SphK as an oncotarget. *Oncotarget* 2017;8:36898–36929.

---

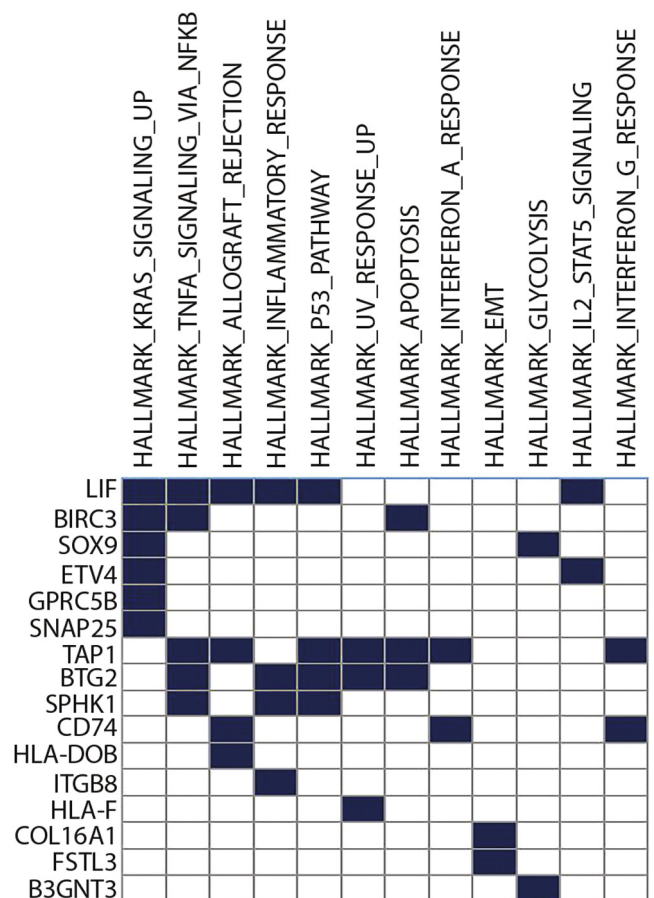
Author names in bold designate shared co-first authorship.



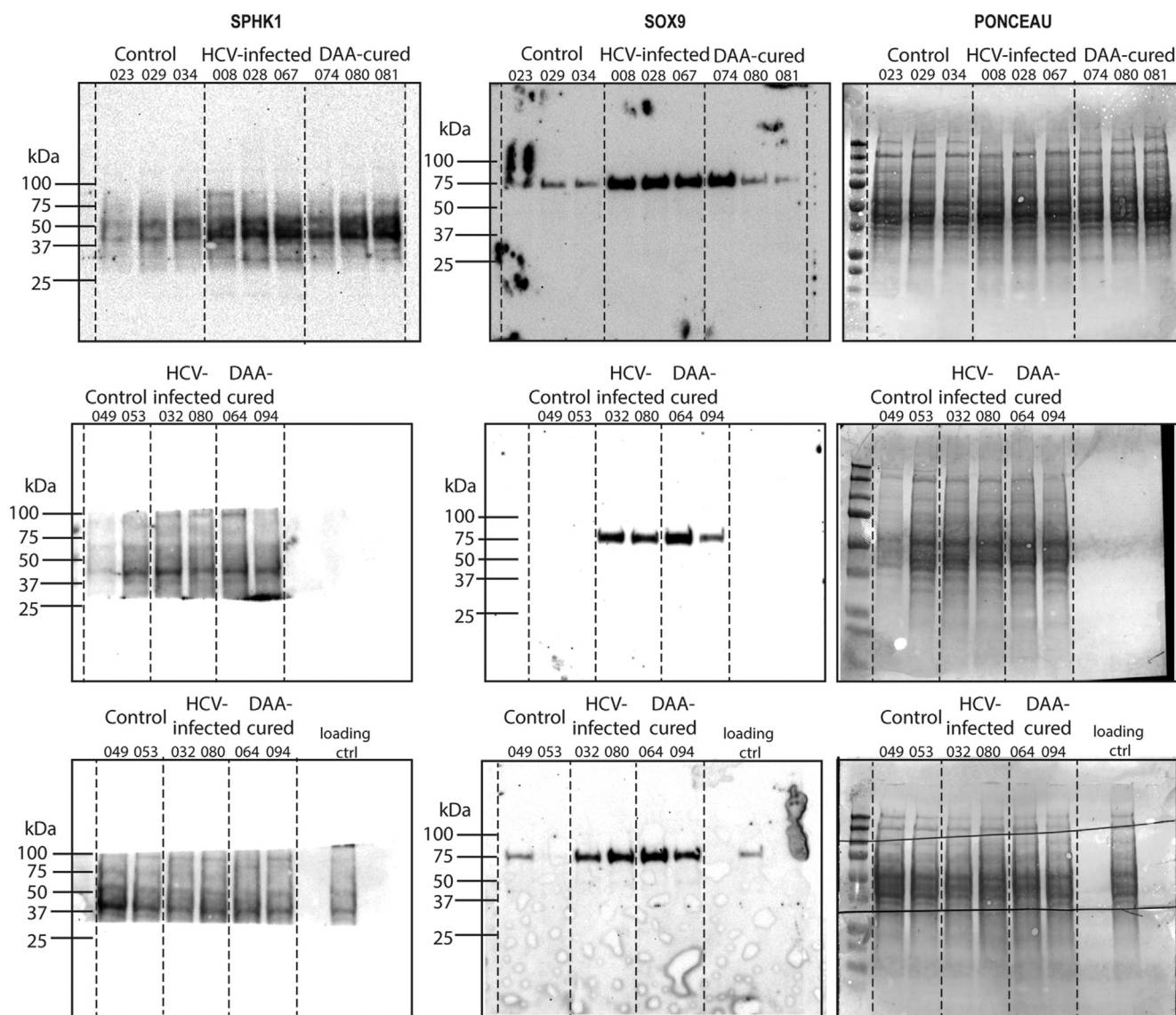
**Supplementary Figure 1.** Persistent H3K27ac modifications in the livers of DAA- and IFN-cured HCV-infected patients. (A) Unsupervised clustering of normalized read counts in ChIP-Seq peaks of genes linked with significant ( $P < .05$ ) H3K27ac modifications in HCV-infected ( $n = 18$ ), DAA-cured ( $n = 8$ ), and IFN-cured ( $n = 13$ ) vs noninfected control ( $n = 6$ ) patients. (B) Persistent H3K27ac modifications among DAA-cured and IFN-cured patients correlate (see Spearman rank correlation coefficients  $r$  and  $P$  values) with H3K27ac modifications among IFN-cured patients.



**Supplementary Figure 2.** Differential epigenetic modifications on immune-related gene signature among HCV-infected, NASH, and HBV-infected liver samples. To analyze the role of epigenetic changes in the disease immune responses, we extracted immune-related genes from MSigDB and performed a restricted correlation study of genes with H3K27ac modifications among NASH, HBV-infected, and HCV-infected patients. Common H3K27ac modifications were analyzed and Spearman rank correlation coefficients and *P* values are shown.

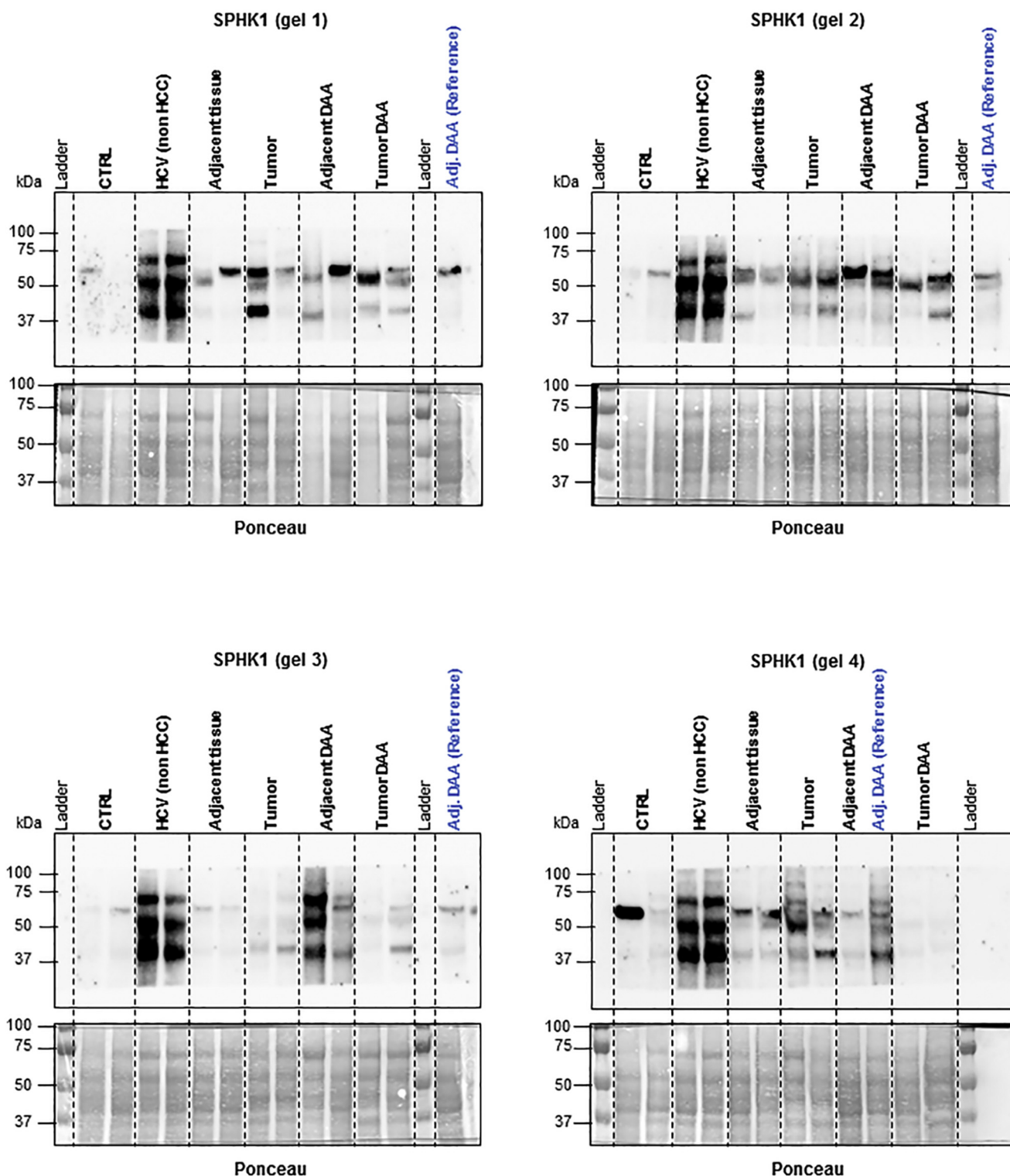


**Supplementary Figure 3.** Hallmark pathway analysis of the 35 genes enriched for H3K27ac modifications and overexpressed in infected and cured human (*n* = 32) and mice (*n* = 15) samples. The 38 genes harboring significant H3K27ac changes from the livers of HCV-infected and DAA-cured patients and mice were subjected to GSEA using hallmark gene sets from the MSigDB Molecular Signatures Database.

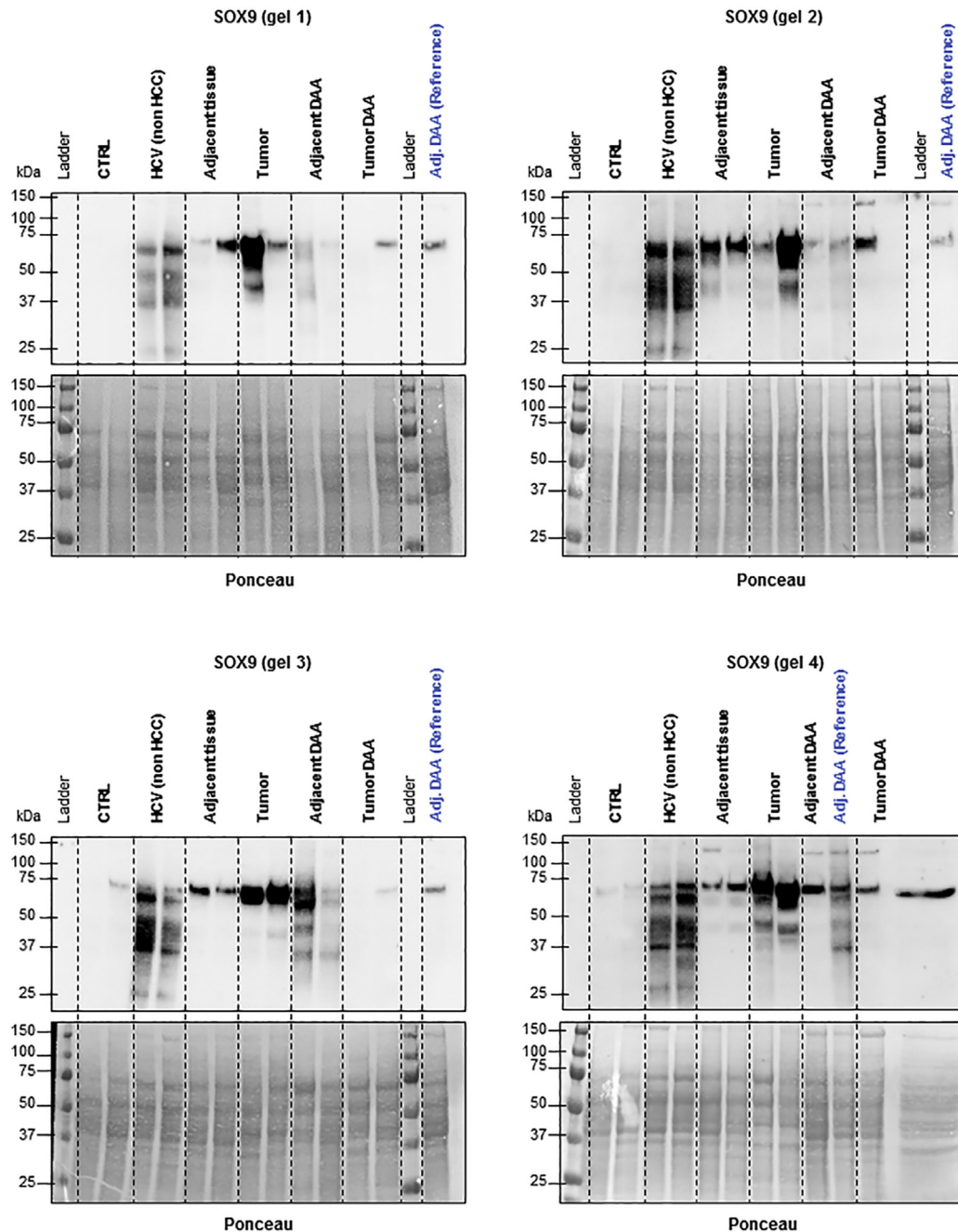


**Supplementary Figure 4.** Full-length immunoblots of SPHK1 and SOX9 protein levels in the livers of control, HCV-infected, and DAA-cured humanized mice. Full-length blots corresponding to representative blots shown in [Figure 6D](#) are shown. Reducing 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis of mouse liver lysates was performed as described in the Methods section.





**Supplementary Figure 5.** Full-length immunoblots of SPHK1 protein level in the livers of control, HCV-infected, and DAA-cured patients (HCC and adjacent tissue). For patients with HCC, SPHK1 was detected in tumor and surrounding tissues (adjacent tissue). A reference sample was loaded on each gel for data normalization. Full-length blots of SPHK1 corresponding to representative blots shown in [Figure 6E](#). Reducing 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis of liver biopsy lysates was performed as described in the Methods section. Multiple weight products visible on the blot could be a result of post-translational protein modifications including glycosylation, phosphorylation, and/or ubiquitination. CTRL, control.



**Supplementary Figure 6.** Full-length immunoblots of SOX9 protein levels in the livers of control, HCV-infected and DAA-cured patients (HCC or adjacent tissue). For patients with HCC, SOX9 was detected in tumor and surrounding tissues (adjacent tissue). A reference sample was loaded on each gel for data normalization. Full-length blots of SOX9 corresponding to representative blots shown in [Figure 6E](#). Reducing 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis of liver biopsy lysates was performed as described in the Methods section. Multiple weight products visible on the blot could be a result of post-translational protein modifications including glycosylation, phosphorylation, and/or ubiquitination. CTRL, control.

---

7.2 **Felli E\*, Saviano A\***, Tripon S, Baumert T, Pessaux P, Baumert TF. Letter to the Editor: Abdominal Surgery in Idiopathic Noncirrhotic Portal Hypertension: Is Preemptive TIPS Reducing Postoperative Complications? *Hepatology* 2019, Oct 9. doi: 10.1002/hep.30985. [Epub ahead of print]. **\*These authors contributed equally to this work**

Article type : Correspondence

## **Letter to the Editor: Abdominal Surgery in Idiopathic Noncirrhotic Portal Hypertension: Is Preemptive TIPS Reducing Postoperative Complications?**

Emanuele Felli,<sup>1,2\*</sup> Antonio Saviano,<sup>1-3\*</sup> Simona Tripon,<sup>1,2</sup> Thomas F. Baumert,<sup>1-3</sup> and Patrick Pessaux<sup>1-3</sup>

<sup>1</sup>Department of General, Digestive, and Endocrine Surgery, Strasbourg University Hospitals, Institut Hospitalo-Universitaire, Strasbourg, France; <sup>2</sup>Institut of Viral and Liver Disease, Inserm U1110, Strasbourg, France; <sup>3</sup>University of Strasbourg, Strasbourg, France.

\*These authors contributed equally to this work.

Correspondence should be addressed to:

Emanuele Felli, M.D.,  
Pôle Hépatodigestif, Nouvel Hôpital Civil  
1, Place de l'Hôpital, 67091 Strasbourg cedex, France

Antonio Saviano, M.D.,  
Institut de Recherche sur les Maladies Virales et Hépatiques, Inserm U1110  
3, rue de Koeberlé, 67000 Strasbourg, France  
+33 (0) 368 85 3727

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/HEP.30985](#)

This article is protected by copyright. All rights reserved



E-mail: emanuele.felli@chru-strasbourg.fr; saviano@unistra.fr

Abbreviations: INCPH, idiopathic noncirrhotic portal hypertension; TIPS, transjugular intrahepatic portosystemic shunt.

To the Editor

We read with great interest the article by Elkrief et al. reporting long-term outcomes of abdominal surgery in patients with idiopathic noncirrhotic portal hypertension (INCPH)<sup>(1)</sup>. In a subgroup analysis, the authors compared the outcome of patients who had (n = 33) or did not have (n = 10) a preemptive transjugular intrahepatic portosystemic shunt (TIPS), showing that TIPS before surgery had no significant impact on postoperative outcomes. This retrospective subgroup analysis had some limitations. The comparison of the clinical characteristics between the groups was performed on the data after TIPS placement, and the small sample size reduced the statistical power of the analysis. These data are still preliminary to draw any firm conclusion—as already discussed by the authors—but could potentially dissuade clinicians in using preemptive TIPS in this clinical context.

INCPH is a rare disease, and it is unlikely that survival analysis of larger cohorts will be available in the immediate future. To help readers and to highlight the magnitude of the differences between the two groups, we reanalyzed the clinical and postoperative data reported by Elkrief et al. and calculated the effect size using logit computation of standardized mean differences. As shown in Table 1, the two groups showed large differences, not only in beta-blocker use but also in history of variceal bleeding, portal vein thrombosis, and portosystemic collaterals (higher in the TIPS group). Importantly, large differences were found in the occurrence of grade  $\geq 3$  postoperative complications within 1 month after surgery (lower in the TIPS group). Thus, it cannot be excluded that patients who underwent decompression had more severe portal hypertension before TIPS positioning and that preemptive TIPS reduced the occurrence of severe postoperative complications and/or allowed the achievement of long-term outcomes comparable to patients with less severe complications. The results need to be confirmed on larger, ideally prospective, cohorts and should not discourage the use of this

technique in the care of preoperative patients. Moreover, preemptive TIPS could also be an important measure to increase surgical eligibility for complex operations by diminishing intraoperative bleeding and allowing a technically less demanding dissection.<sup>(2)</sup> In patients undergoing splenectomy, a pre-emptive calibrated TIPSS could reduce the incidence of portal vein thrombosis (postoperative rate of 50% in Elkrief et al.). Hence, even though conclusive data on the postoperative benefit of preemptive TIPS in these patients are still lacking, a potential benefit in the surgical feasibility and severe postoperative complications should be considered in the preoperative workup.

**Acknowledgment:** This work was supported by ARC grants (TheraHCC, TheraHCC2.0 IHUARC IHU201301187, and IHUARC2019 IHU201901299).

## REFERENCES

- 1) Elkrief L, Ferrusquia-Acosta J, Payancé A, Moga L, Tellez L, Praktiknjo M, et al. Abdominal surgery in patients with idiopathic noncirrhotic portal hypertension: a multicenter retrospective study. HEPATOLOGY 2019 Mar 29. doi: 0.1002/hep.30628.
- 2) Lahat E, Lim C, Bhangui P, Fuentes L, Osseis M, Moussallem T et al. Transjugular intrahepatic portosystemic shunt as a bridge to non-hepatic surgery in cirrhotic patients with severe portal hypertension: a systematic review. HPB (Oxford). 2018 Feb;20(2):101-109.

| Variable   | Effect Size<br>(Patient With Portal Decompression<br>vs. Patient Without) |
|--|---|
| <b>Clinical features</b>                                       |   |
| Male   | 0.3822  |
| Age  | 0.0204  |
| Age-adjusted Charlson Comorbidity Index                        | 0.0669  |
| ASA score  | 0.0579  |
| At least one extrahepatic comorbidity<br>associated with INCPH | -0.2375   |
| History of ascites   | 0.085   |
| Ascites at surgery   |   |
| Absent   | -0.085  |
| Controlled with diuretics                                      | 0.1611  |
| Clinically detected  | -0.1192   |
| <b>Endoscopic data</b>   |   |
| Previous variceal bleeding                                     | <b>0.8493</b>   |
| <b>Treatments</b>  |   |
| Anticoagulation therapy  | 0.1684  |
| Antiplatelet agents  | -0.3822   |
| Diuretic agents  | 0.1586  |
| Beta blockers  | <b>-1.2448</b>  |
| <b>Imaging data</b>  |   |
| Portal vein thrombosis   | <b>0.8023</b>   |
| Portosystemic collaterals at imaging                           | <b>0.7977</b>   |

---

**Laboratory data**

|                  |         |
|------------------|---------|
| Platelets        | -0.1287 |
| Hemoglobin       | -0.0096 |
| INR              | 0.1019  |
| Serum bilirubin  | 0.2177  |
| Serum creatinine | -0.0273 |
| Serum albumin    | -0.0664 |
| MELD score       | 0.056   |

**Surgical data**

|                      |         |
|----------------------|---------|
| Major surgery        | 0.2357  |
| Minor surgery        | -0.2357 |
| Emergency procedures | 0.1611  |

**Postoperative outcomes**

|   |                |
|---|----------------|
| Occurrence of $\geq 1$ grade $\geq 3$ postoperative complication within 1 month after surgery | <b>-0.5832</b> |
| Occurrence of $\geq 1$ portal hypertension-related complication within 3 months after surgery | -0.2296        |
| Death within 6 months after surgery   | 0.0581         |
| Unfavorable outcome   | -0.0367        |

---

Effect size values <0.1 indicated very small differences; between 0.1 and 0.3 indicated small differences, between 0.3 and 0.5 indicated moderate differences and >0.5 indicated large differences (in bold). Effect size for continuous variables were calculated after estimation to mean and SD and log transformation (Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014 Dec 19;14:135.)

Abbreviations: ASA, American Society of Anesthesiologists : INR, international normalized ratio; MELD, Model for End-Stage Liver Disease.

**TABLE 1. Effect Size for the Comparison of the Patients With/Without Portal Decompression Before Surgery**

---

7.3 Wakabayashi T, Ouhmich F, Gonzalez-Cabrera C, Felli E, **Saviano A**, Agnus V, Savadjiev P, Baumert TF, Pessaux P, Marescaux J, Gallix B. Radiomics in hepatocellular carcinoma: a quantitative review. *Hepatol Int.* 2019 Aug 31.



# Radiomics in hepatocellular carcinoma: a quantitative review

Taiga Wakabayashi<sup>1</sup> · Farid Ouhmich<sup>2</sup> · Cristians Gonzalez-Cabrera<sup>2</sup> · Emanuele Felli<sup>1,2,3,4,5</sup> · Antonio Saviano<sup>2,4,5</sup> · Vincent Agnus<sup>2</sup> · Peter Savadjiev<sup>6</sup> · Thomas F. Baumert<sup>2,4,5</sup> · Patrick Pessaux<sup>1,2,3,4,5</sup> · Jacques Marescaux<sup>1,2,3</sup> · Benoit Gallix<sup>2,6</sup>

Received: 29 March 2019 / Accepted: 26 July 2019

© Asian Pacific Association for the Study of the Liver 2019

## Abstract

Radiomics is an emerging field which extracts quantitative radiology data from medical images and explores their correlation with clinical outcomes in a non-invasive manner. This review aims to assess whether radiomics is a useful and reproducible method for clinical management of hepatocellular carcinoma (HCC) by reviewing the strengths and weaknesses of current radiomics literature pertaining specifically to HCC. From an initial set of 48 articles recovered through database searches, 23 articles were retained to be included in this review after full screening. Among these 23 studies, 7 used a radiomics approach in magnetic resonance imaging (MRI). Only two studies applied radiomics to positron emission tomography–computed tomography (PET–CT). In the remaining 14 articles, a radiomics analysis was performed on computed tomography (CT). Eight studies dealt with the relationship between biological signatures and imaging findings, and can be classified as radiogenomic studies. For each study included in our review, we computed a Radiomics Quality Score (RQS) as proposed by Lambin et al. We found that the RQS (mean  $\pm$  standard deviation) was  $8.35 \pm 5.38$  (out of a possible maximum value of 36). Although these scores are fairly low, and radiomics has not yet reached clinical utility in HCC, it is important to underscore the fact that these early studies pave the way for the radiomics field with a focus on HCC. Radiomics is still a very young field, and is far from being mature, but it remains a very promising technology for the future for developing adequate personalized treatment as a non-invasive approach, for complementing or replacing tumor biopsies, as well as for developing novel prognostic biomarkers in HCC patients.

**Keywords** Radiomics · Radiogenomics · Hepatocellular carcinoma · Tumor heterogeneity

## Introduction

Hepatocellular carcinoma (HCC) is the most common liver cancer. HCC mostly occurs in patients with chronic liver disease such as cirrhosis or severe fibrosis. Its major causes are chronic liver disease due to chronic hepatitis B and C virus infection or metabolic liver disease, such as non-alcoholic steatohepatitis or alcoholic liver disease. HCC is poorly symptomatic at the early stages of its development, and often becomes symptomatic only at an advanced stage when curative treatments are no longer possible. Therefore, the prognosis of HCC remains unsatisfactory [1].

Recently, tumor heterogeneity in terms of biological and genomic characteristics has become a topic of interest in cancer research [2]. Tumor heterogeneity can be demonstrated not only within primary cancers and various metastases (inter-tumor heterogeneity), but also within the same tumor (intra-tumor heterogeneity). Numerous publications

✉ Benoit Gallix  
benoit.gallix@ihu-strasbourg.eu

<sup>1</sup> Institut de Recherche Contre les Cancers de l'Appareil Digestif (IRCAD), Strasbourg, France

<sup>2</sup> Institut hospitalo-universitaire (IHU), Institute for Minimally Invasive Hybrid Image-Guided Surgery, Université de Strasbourg, 1 Place de l'Hôpital, 67000 Strasbourg, France

<sup>3</sup> General, Digestive, and Endocrine Surgery, Nouvel Hôpital Civil, Université de Strasbourg, Strasbourg, France

<sup>4</sup> Inserm, U1110, Institut de Recherche sur les Maladies Virales et Hépatiques, Université de Strasbourg, Strasbourg, France

<sup>5</sup> Pôle Hépatodigestif, Hôpitaux Universitaires, Strasbourg, France

<sup>6</sup> Department of Diagnostic Radiology, McGill University, Montreal, Canada

have shown that HCCs are extremely heterogeneous both in terms of their genotype and phenotype [3, 4]. Thus, not only can different patients develop very different types of cancer, but tumors in the same patient can also be heterogeneous. Patient prognosis depends strongly on this phenotypic expression, which could be evaluated, for example, by analyzing pathological characteristics, such as the histological grade of the tumor [5] and microscopic vascular invasion [6]. Many staging systems including clinical, biological, and imaging data have been developed such as the Barcelona Clinic Liver Cancer staging system [7], the Cancer of the Liver Italian Program [8, 9], and the Okuda criteria [10]. However, beyond the size and number of lesions, none of these scoring systems include information on the tumor phenotype that affects patient survival in a significant way [11]. Diagnostic and therapeutic trends in liver cancer are changing; they now tend to be determined by significant biological and genomic tumor characteristics.

### Tumor characterization via medical imaging

Among all techniques for interrogating tumor phenotype and heterogeneity, medical imaging provides several advantages [11, 12]. By allowing an evaluation of tumors as a whole, in a minimally invasive and reproducible manner, imaging is complementary to biopsies, which only provide samples that are not always representative of tumor heterogeneity [13]. Since biopsies merely aim to sample a small portion of the tumor and it is difficult to repeat pathological assessments, they provide limited information regarding tumor heterogeneity. Conversely, medical imaging methods such as computed tomography (CT), positron emission tomography (PET) or magnetic resonance imaging (MRI) can capture a tumor in its entire 3D extent with features that reflect tumor heterogeneity.

Such cross-sectional imaging techniques have become essential tools for modern oncology management [14–16]. Protocols for image acquisition based on these modalities have reached such a degree of sophistication that to make a therapeutic decision, tissue biopsy is often unnecessary when the diagnostic criteria for HCC are all met [17–19]. However, the methods currently used to assess the prognosis of patients with HCC based on the acquired images remain very rudimentary and are simply based on size, number of tumors, and vascular invasion as subjectively analyzed by the radiologist [7–10, 20–22]. As we recognize the need to go beyond tumor size and number, given the sophistication of the acquired imaging signal, advanced image analysis tools are now required to establish biomarkers from the complex signal that can be extracted from the images [22].

### Going beyond size: semantic descriptors of tumor appearance

To improve image-based tumor characterization, one possible approach is based on a qualitative or semi-quantitative analysis as performed by an expert radiologist, using standardized reading scores. Examples of characteristics generally described for HCC by radiologists include the presence of arteries in the tumor, a peri-tumoral halo or the tumor's apparent heterogeneity. Specifically, HCC could be encapsulated, well limited, or homogeneously hypervascularized after contrast injection or, on the contrary, poorly limited with vascular invasion, heterogeneously enhanced, and with a larger area of necrosis. This type of image analysis is referred to as a “semantic” analysis of lesions, where images are evaluated by one or more trained radiologists on the basis of semantic descriptors of the lesion(s) that are part of the established radiologist's lexicon. Another example of a semantic approach to tumor classification is the LI-RADS classification, which provides a standardized radiological lexicon built by consensus among expert radiologists [23].

The semantic analysis approach is interesting because it often provides a pathophysiological explanation for the image descriptions. The process of quantifying visual semantic characteristics unfortunately remains quite subjective and difficult to reproduce. Its implementation also poses practical problems because this process is very time consuming and cannot easily be used with large populations or integrated into clinical practice. Furthermore, its low inter- and intra-observer reproducibility makes this analysis difficult to standardize.

### Going beyond size: quantitative descriptors of tumor appearance

Another approach to image-based tumor characterization is based on quantitative image descriptors. This type of approach is known as radiomics, and it aims to quantify the morphological appearance of the tumor, i.e., its imaging phenotype, using mathematically defined quantitative features [24, 25]. This type of quantitative information cannot be easily assessed by a radiologist, but can be computed with specialized computer algorithms. Radiomics was popularized by Lambin et al. [26] in 2012, and since then has been extensively used as a methodology to assess tumor heterogeneity, to establish a correlation with clinical or biological information [27].



## Radiomics

The radiomics analysis pipeline consists of three main steps: (1) tumor segmentation, (2) computation of radiomics features within the segmented tumoral region and (3) feature selection, model building, and classification. The details of every stage of the radiomics pipeline, their implementation details, and limitations have been discussed at length elsewhere [28–31]. Here, we briefly mention a few details that are relevant to the findings of our review.

### Segmentation

Radiomics requires a tumor region to be segmented in order to define the image region where quantitative descriptors are to be computed. Automatic segmentation has been a long-standing objective in the computer vision and machine learning fields [32], but remains difficult to achieve. This is why most radiomics studies still rely on manual tumor segmentation. Unfortunately, the use of manual segmentation not only makes measurements long and tedious, but also hinders measurement reproducibility. It should be noted that in studies using semantic analysis, segmentation is not necessary [33–40].

### Radiomics features

Many different quantitative descriptors (features) have been proposed for radiomics [28–31]. The studies included in the present review typically use first-order, second-order, and higher-order mathematical descriptors such as grayscale matrix analysis (co-occurrence) which takes into account the relationships between neighboring pixels. A filtering step—using for example Gauss Laplacian filters—is usually performed prior to any analysis to reduce noise and improve performance [41–46].

### Feature selection, model building, classification

One difficulty of radiomics is that it can calculate thousands of parameters for a single image. If the number of parameters is very high and the population is small—a few dozen patients—there is a significant risk of overfitting. This means that, in practice, there will almost always be parameters which are statistically correlated with the patient's condition. To limit this risk, the number of parameters must be significantly reduced before building the statistical model and, if possible, the model needs to

be tested on an independent imaging dataset, obtained for instance at a different institution [32].

### Limitations

The limitations of this method have been thoroughly analyzed in previous review articles about radiomics [28–32, 47–50]. One of the main limitations is the lack of standardization of image acquisition (such as slice thickness, choice of MRI sequences, or timing after contrast injection), which could add a strong bias to the post-imaging workflow.

## Radiogenomics

Radiogenomics refers to the study of correlations between genome and molecular measurements on one hand, and radiological measurements [either quantitative or qualitative (semantic) features] on the other [51–53]. Radiomics and radiogenomics have the same objective, which is to transform radiological images into objective measurements representative of tumor heterogeneity.

### Purpose

In recent years, we have seen the publication of numerous studies on radiomics with the objective of improving the diagnosis and stratification of patients with primary liver cancer [54]. The results of these studies are sometimes contradictory and complex to reproduce. Unfortunately, many of the published works show significant methodological weaknesses which have limited their impact in clinical practice. Therefore, there is a need to clarify the performance of radiomics as a prognostic and stratification tool for HCCs. We have, thus, conducted this systematic review to assess whether radiomics is a useful and reproducible method for HCC clinical management in terms of diagnosis, prognosis, and estimation of treatment response by reviewing the advantages and qualities of the studies included.

## Review strategy

This review was conducted for all studies, published between January 1, 2007 and December 23, 2018 following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [55]. We used the following search strategy on PubMed and Embase: ((hepatocellular carcinoma [Title]) and (radiomics [Title] OR radiogenomics [Title] OR omics [Title])), and a combination of associated terms from the controlled MeSH vocabulary. The final search was carried out on 23 December 2018. Inclusion criteria were (1) human studies, (2) English language studies, (3) full-text articles, and (4) studies reporting on

semantic features or radiomics analyses for HCC. Exclusion criteria included (1) animal/experimental studies, (2) abstracts, reviews, and case reports, (3) only ultrasound-related studies, and (4) no investigation on clinical outcome. The existing review articles were analyzed to look for possible additional references. Every abstract was reviewed for initial selection, and then all chosen articles were fully downloaded. Two authors (T.W, F.O), and an independent third one (B.G) when consensus was needed, individually assessed each manuscript to eliminate those which failed to meet the inclusion criteria. In accordance with the fore-mentioned search strategy and criteria, we found 48 articles and excluded the 7 review articles, 7 articles without clinical outcomes, 7 articles focusing on other etiology, 2 ultrasound-related studies, 1 animal study, and 1 article without imaging analysis. Finally, we included 23 of them in this review after a full screening (Fig. 1). The articles are summarized in Table 1.

The 23 studies included in our review use either semantic or quantitative features. Table 2 describes the types of features used in the quantitative and semantic categories. Quantitative features are computed via specialized software and are classified as first-order, second-order, and high-order descriptors, and morphological features. Semantic features are visually interpreted by radiologists and are defined as eight features in our review: two-tract predictor of venous invasion, non-smooth tumor margin, peritumoral enhancement, tumor size, tumor liver difference, PET–CT positivity, infiltrative pattern, and mosaic appearance.

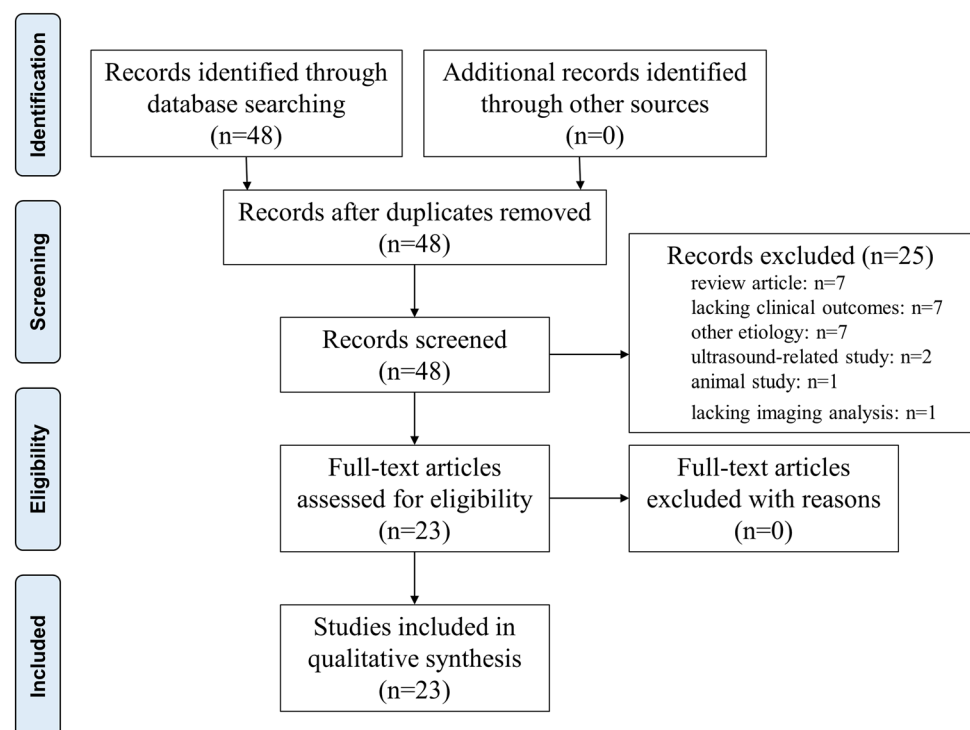
For all articles, we analyzed separately five phases within the radiomics workflow: (1) data inclusion and selection criteria, (2) description and standardization of imaging acquisitions, (3) feature extraction, (4) exploratory analysis, and (5) modeling [24]. For each study, the following data were systematically recorded: first author, year of publication, type of study (retrospective or prospective), number of centers (single or multicentric), objective of the study (tumor detection, tumor characterization, prognosis, response to treatment), type of imaging modality (MRI, CT, PET), technical parameters (slice thickness, magnetic field for MRI scanner, contrast media injection), main radiomics investigated features, presence of genomic consideration, number of patients used to build the model, and presence of a validation cohort.

## Findings

### Variations in imaging modalities and protocols

Seven out of the 23 studies used a radiomics approach on MRI images [33, 34, 56–60]. In one study, both CT and MRI were investigated [34]. MRI provides a very rich signal which can provide accurate information about tumors. However, MRI acquisitions are very difficult to standardize, with numerous acquisition parameters and many variations between manufacturers. MRI is also sensitive to many artifacts which complicate the reproducibility of measurements such as motion artifacts—cardiac or respiratory—due to

**Fig. 1** Study selection



**Table 1** Summary of radiomic analyses from 23 studies

| Author               | Purposes  | Treatment           | Modality (slice thickness)       | #Patients | Mean tumor size   | Sig-nificant features | % RQS score (points) | Conclusion  |
|----------------------|---|---------------------|----------------------------------|-----------|---|-----------------------|----------------------|---|
| Wu et al. [30]       | Prediction of a pathological grading  | –                   | T1- and T2-weighted MRI (7 mm)   | 170       | –   | Q                     | 42 (15)              | A computed radiomics signature itself or combined with clinical factors could help to classify the patients into high-grade or low-grade HCC cases                                |
| Zhou et al. [31]     | Prediction of a pathological grading  | –                   | Gd-DTPA MRI (4.4 mm)             | 46        | Low-grade: $3.8 \pm 2.1$ cm, High-grade: $6.3 \pm 2.9$ cm | Q                     | 22 (8)               | Two extracted features were sufficient to classify the histopathological HCC grade  |
| Miura et al. [32]    | Comparison of clinicopathological properties of high HCC                                | Hepatectomy or TACE | EOB-MRI                          | 77        | High-HCC: $3.6 \pm 1.2$ cm, Low-HCC: $5.3 \pm 4.9$ cm     | Q                     | 17 (6)               | High HCC differs from low HCC in PIVKA-II level, histopathological differentiation, and expression of SLC01B3   |
| Kim et al. [33]      | Prediction of MVI   | Hepatectomy or LT   | Contrast MRI (2–5.5 mm)          | 104       | 3.6 cm  | S                     | 14 (5)               | “Peritumoral enhancement” on MRI may be correlated with MVI during surgical planning for patients with HCC  |
| Hectors et al. [34]  | Quantification of tumor heterogeneity, prediction of histopathology and gene expression | –                   | mpMRI                            | 32        | $4.4 \pm 3.3$ cm  | Q                     | 39 (14)              | First-order statistical features from mpMRI showed high intra-tumoral and inter-tumoral heterogeneity on the HCC lesions. This heterogeneity was also remarkable at genomic level |
| Starmans et al. [35] | Diagnosis   | –                   | T2-weighted (Fat–Sat) MRI (7 mm) | 119       | 3 cm  | Q                     | 0 (0)                | Patient-based (age and gender) and textural characteristics were used to differentiate between benign and malignant tumors  |

Table 1 (continued)

| Author               | Purposes                                    | Treatment   | Modality (slice thickness)    | #Patients                  | Mean tumor size  | Sig-nificant features | % RQS score (points) | Conclusion   |
|----------------------|---|-------------|-------------------------------|----------------------------|--|-----------------------|----------------------|--|
| Renzulli et al. [36] | Prediction of MVI based on TTPVI            | Hepatectomy | Contrast CT (2.5 mm), EOB-MRI |                            | 125 3.3 cm (1.8–5.2 cm)  | S                     | 8 (3)                | Tumor size, “non-smooth tumor margins”, “peritumoral enhancement”, and “TTPVI” were correlated with the presence of MVI in HCC   |
| Blanc et al. [37]    | OS and PFS                                  | TARE        | PET/CT                        |                            | 47 6 cm (4.3–9 cm)   | Q                     | 19 (7)               | Radiomics signature computed with textural features was highly correlated with survival  |
| Park et al. [38]     | OS and DFS                                  | Hepatectomy | PET/CT                        |                            | 92 2.5 cm  | S                     | 17 (6)               | High correlation was found between PET positivity and survival. Additionally, PET positivity may also lead to determine the resection margin to improve survival   |
| Kuo et al. [39]      | Treatment sensitivity and prediction of MVI | –           | Contrast CT (2.5 mm)          |                            | 30 –   | S                     | 19 (7)               | “Tumor margin” showed a strong correlation with MVI, TNM, and the expression of a drug response gene. “Internal arteries” also showed a correlation with MVI   |
| Peng et al. [40]     | Prediction of MVI                           | –           | Contrast CT (5 mm)            | Training: 184<br>Test: 120 | Training: MVI (+) 6.3 cm, MVI (–) 5.7 cm, Validation: MVI (+) 6.4 cm, MVI (–) 4.9 cm | S, Q                  | 47 (17)              | Radiological features and a radiomics signature computed with first-order statistical features showed a correlation with MVI   |
| Segal et al. [41]    | OS and Prediction of MVI genes              | Hepatectomy | Contrast CT                   | Training: 30<br>Test: 32   | –  | S                     | 42 (15)              | “Internal arteries” was found as a key imaging trait to predict OS and MVI in combination with “hypodense halo”. Those features were also correlated with the expression of genes involved in the development of HCC lesions |

Table 1 (continued)

| Author               | Purposes                                   | Treatment           | Modality (slice thickness) | #Patients                  | Mean tumor size                                 | Sig-nificant features | % RQS score (points) | Conclusion   |
|----------------------|--|---------------------|----------------------------|----------------------------|---|-----------------------|----------------------|--|
| Banerjee et al. [42] | OS- and RFS-based on RVI                   | Hepatectomy or LT   | Contrast CT (2.5–3 mm)     | 157                        | 2.8 cm (1.8–4.5 cm)                             | S                     | 53 (19)              | RVI computed with three different imaging traits was correlated with MVI   |
| Taouli et al. [43]   | Prediction of MVI and aggressive phenotype | –                   | Contrast CT (5 mm)         | 38                         | 5.7 ± 3.2 cm                                    | S, Q                  | 19 (7)               | Correlation was found between some imaging traits and the aggressive profile of the tumors   |
| Zheng et al. [44]    | Recurrence prediction and OS               | Hepatectomy         | Contrast CT                | Training: 212<br>Test: 107 | –   | Q                     | 47 (17)              | A radiomics score computed with textural features was sufficient to predict post-operative recurrence and survival in patients with solitary HCC.  |
| Akai et al. [45]     | OS and DFS                                 | Hepatectomy         | Contrast CT (5 mm)         | 127                        | 3.7 cm (2.4–7 cm)                               | Q                     | 25 (9)               | First-order statistical features were sufficient to predict post-operative survival  |
| Zhou et al. [46]     | Recurrence prediction                      | Hepatectomy         | Contrast CT (1.25 mm)      | 215                        | –   | Q                     | 25 (9)               | A radiomics signature using first-order statistical features combined with clinical factors was a good predictor of early recurrence after surgery |
| Chen et al. [48]     | OS and DFS                                 | Hepatectomy         | Contrast CT (1.25 mm)      | 61                         | –   | Q                     | 17 (6)               | Tumor prognosis could be predicted using Gabor and Wavelet filtration responses  |
| Li et al. [49]       | OS and treatment sensitivity               | Hepatectomy or TACE | Contrast CT (1.25 mm)      | 130                        | 8.0 cm (5.1–18.7 cm)                            | Q                     | 19 (7)               | Wavelet features were correlated with post-operative survival on HCC patients, suggestive of the suitable treatment choice                         |
| Raman et al. [50]    | Diagnosis                                  | –                   | Contrast CT (3 mm)         | 80                         | Adenoma 7 ± 3 cm,<br>FNH 6 ± 3 cm, HCC 8 ± 3 cm | Q                     | 3 (1)                | A model created using first-order statistical features could differentiate three types of common hypervascular lesions                             |

Table 1 (continued)

| Author            | Purposes                                 | Treatment           | Modality (slice thickness) | #Patients | Mean tumor size | Sig-nificant features | % RQS score (points) | Conclusion   |
|-------------------|--|---------------------|----------------------------|-----------|-----------------|-----------------------|----------------------|--|
| Xia et al. [51]   | OS with interpretable biological meaning | Hepatectomy or LT   | Contrast CT                | 37        | –               | Q                     | 22 (8)               | The volume of transition between tumor and liver, and the heterogeneity of the lesion were correlated with survival. Those two features associated with six others, were correlated with prognostic genes expression |
| Bakr et al. [52]  | Prediction of MVI                        | –                   | Contrast CT ( $\leq 3$ mm) | 28        | 7.4 cm          | Q                     | 3 (1)                | Textural features computed with single-phased or combined-phased images were correlated with MVI   |
| Cozzi et al. [53] | OS and local control                     | Radiotherapy (VMAT) | Non-contrast CT (3 mm)     | 138       | –               | Q                     | 14 (5)               | Survival could be predicted using a radiomics signature made by a single shape-based feature   |

CT computed tomography, DFS disease-free survival, EOB-MRI ethoxybenzyl-diethylenetriamine pentaacetic acid, FNH focal nodular hyperplasia, HCC hepatocellular carcinoma, LT liver transplantation, mpMRI multiparametric magnetic resonance imaging, MRI magnetic resonance imaging, MVI microvascular invasion, OS overall survival, PET positron emission tomography, PET/CT computed tomography integrated with positron emission tomography, PIVKA-II protein induced by vitamin K absence-II, Q quantitative features, RQS radiomics quality score, RVI radiogenomic venous invasion, sec seconds, S semantic features, TACE transarterial chemoembolization, TTPVI two-trait predictor of vascular invasion, TARE transarterial radioembolization, VMAT volumetric modulated arc therapy

**Table 2** Summary of significant extracted features in 23 studies

| Quantitative features    | Number of studies | Semantic features                      | Number of studies |
|--------------------------|-------------------|--|-------------------|
| First-order descriptors  | 12                | Two-trait predictor of venous invasion |                   |
| Second-order descriptors | 7                 | Internal arteries                      | 6                 |
| High-order descriptors   | 3                 | Hypo-attenuating halos                 | 4                 |
| Morphological feature    | 1                 | Non-smooth tumor margin                | 3                 |
|                          |                   | Peritumoral enhancement                | 2                 |
|                          |                   | Tumor size                             | 2                 |
|                          |                   | Tumor liver difference (estimated)     | 1                 |
|                          |                   | PET/CT positivity                      | 1                 |
|                          |                   | Infiltrative pattern                   | 1                 |
|                          |                   | Mosaic appearance                      | 1                 |

First-order descriptors comprise shape (compactness or sphericity), skewness, kurtosis, mean, energy, median, entropy, peak, standard deviation, intensity ratio between tumor and liver, enhancement ratio, and tumor liver difference (computed). Second-order descriptors comprise gray level matrices, cluster prominence, strength, and textural features variance. Morphological feature comprises tumor margin volume

*PET/CT* computed tomography integrated with positron emission tomography

long acquisition times or field homogeneity with image and signal distortion consequences.

Only two studies performed a radiomics analysis using PET–CT data [35, 61]. Blanc-Durand et al. [61] reported that a radiomics signature computed on whole-liver PET 18F-FDG imaging performed before transarterial radioembolization using Yttrium-90 predicted progression-free survival (PFS) and overall survival (OS) in patients with advanced HCC. This study is unique, as it uses an integrative whole-liver approach and underlines the importance of including not only tumor lesions, but also adjacent liver parenchyma to explore the tumor environment. In all the other articles, a radiomics analysis was carried out on CT images. An iodine contrast agent was used in all studies [34, 37–46, 62, 63] except for one study [64]. Four studies dealt with quantitative features from contrast-enhanced multiple-phase CT images [43–44, 63] and all six studies dealt with only semi-quantitative (semantic) features from contrast-enhanced multiple-phase CT images [34, 37–40]. The other four studies focused on quantitative characteristics from contrast-enhanced single-phase CT images (arterial phase in two [41, 46], portal phase in two [45, 62]).

### Clinical utility of radiomic analysis in HCC

Eight studies dealt with the relationship between the biological variables and imaging findings. For these studies, we used the terminology of radiogenomics—which is used often in the literature—although most refer to microscopic vascular invasion (MVI), which is not a genomic variable [34, 36, 39–40, 58, 59, 62]. Four of these articles considered MVI in their studies [33, 38–39]. In fact, MVI is the most frequent feature required to investigate the correlation with pathological characterization in our review. Among all the

included articles, eight studies focused on the correlation between radiomic features and MVI [33, 34, 37–40, 63]. Bakr et al. [63] demonstrated that quantitative features which capture the lesion texture, intensities, and shape extracted from triphasic CT images had a better accuracy in MVI prediction, compared to two previously reported signatures based on semantic features, radiogenomic venous invasion [38, 39], and TTPVI [34].

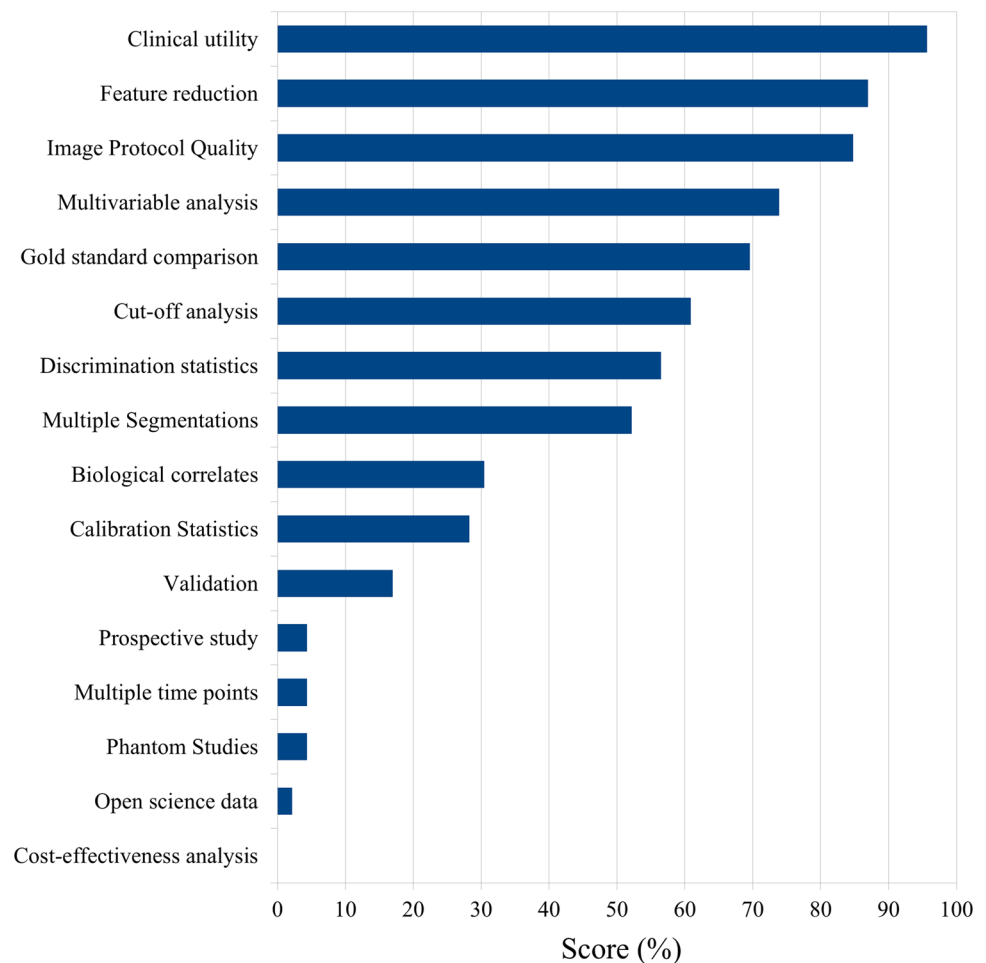
The percentage of studies performed for tumor characterization was 61% (14/23) [34–40, 46, 56–60, 63]. Raman et al. [46] described a model that distinguished successfully different lesion types (focal nodular hyperplasia, hepatic adenomas, and HCC) and normal liver tissue with high predicted classification performance accuracy, as compared to two human readers.

Twelve (52%) out of 23 studies were conducted to aid with prognosis [35, 38, 39, 42–45, 58, 61, 62, 64]. Cozzi et al. [64] have described a radiomics method to predict tumor response and OS for patients treated with arc-based radiotherapy. The other 11 studies were related to tumor prognosis after surgical treatment. Zheng et al. [41] demonstrated that a radiomics score measured on baseline CT was a prognostic factor of the outcome in patients who underwent liver resection for HCC. They concluded that this score might be complementary to the current staging system and help to stratify individualized treatments for solitary HCC patients.

### Quality assessment of radiomic studies for HCC

To assess the quality of the included studies, we used the Radiometric Quality Score (RQS) as published by Lambin et al. [24]. The RQS—which evaluates 16 key components of the radiomics workflow—is a tool which analyzes the

**Fig. 2** Completing rate of each query item in radiomics quality score for 23 studies



quality of a radiomics study. It assigns points according to 16 different criteria, for a maximum score of 36. In our work, the RQS score was evaluated by two authors (T.W. and F.O.) first separately, and then by consensus.

The results of the quality evaluations according to the RQS criteria are presented in Table 1 and the RQS scores are detailed by criteria in Fig. 2. The RQS (mean  $\pm$  standard deviation) was  $8.35 \pm 5.38$  (representing 23% of the possible maximum value of 36). All but one study were scored below 18 (50%) due to a lack of external validation and/or to retrospective design as shown in Table 1. The main three reasons for entirely insufficient scores in the reviewed articles are the lack of prospective design except in one study [59], the lack of validation except in four studies [37, 38, 41, 56], and the lack of open-access scientific data resources except in two studies [38, 41]. Additionally, no studies have attempted to analyze the cost effectiveness of a radiomics approach applied to a specific clinical situation.

The prospective nature of the study is a major component of the RQS, representing almost 20% of the total score (7 points out of 36). A prospective study ensures that included patients could undergo a consistent imaging

protocol, which would provide results that are more reliable as compared to a retrospective study. From all the studies included in the present review, all but one study [59] were retrospective evaluations.

Next, our analysis shows that existing radiomics studies in HCC have involved only relatively small patient populations, with an average number of patients of 110, with half of the studies including fewer than 100 cases. In addition, there is a general lack of validation in an independent population. Most of the studies trained the algorithm and evaluated its performance in only a small group of patients, risking overfitting. In the RQS, the type of validation performed accounts for as much as 10/36 points (nearly 28%), with the highest score being given to validation on independent datasets. In this regard, one option is the use of open-access scientific data. If such data are available for a radiomics study, researchers will be able to use the data set for a validation, reproduction, or replication with various data sets [24]. However, the lack of open-access data for HCC is a considerable limitation and results in a reduction by as many as 4 points of the total RQS.



Finally, only two articles used semi-automatic segmentation [60, 61], while all the other studies used manual tumor delineation. This could be a limitation to reproducibility and a barrier to the deployment of the method because as noted earlier, manual segmentation is very operator dependent and time consuming.

### Semantic analysis of HCC

Eight articles in our review used semantic analysis, with potentially interesting results allowing the standardization of vascular invasion criteria [34–40]. One study examined the correlation between a quantitative and a semantic characteristic in an attempt to reduce variability between the observers [37].

### Discussion

The results of our analysis showed that the overall quality of the studies evaluated is low or moderate with an average RQS score of 8.35 or less than 25%. This underlines the fact that radiomics is a very recent technique that has not yet reached maturity, but also that this method is complex and that its standardization is not easy to implement.

This review demonstrates the importance of being cautiously optimistic about radiomic signatures. This new field of research has led to an accumulation of experimental and analytical work, most often thorough retrospective studies. However, the consolidation and standardization of experimental methods have not been standardized or validated. This review shows that the published radiomics work on HCC adds little to scientific knowledge, and is currently not useful in clinical decision making. However, radiomics is a very young field, far from being mature, and has many subtleties that researchers are just learning to manage. In any case, the automated calculation of oncology biomarkers based on data acquired through medical imaging remains a necessity and is a matter of urgency. Radiomics in its current conventional form is probably only one step in the development of reliable computational image biomarkers that will probably need to be specific to a particular organ and tumor type. Although the results of our review article are somewhat disappointing regarding HCC, it is important to note that these published studies pave the way of the field of radiomics with a focus on HCC. Also, they demonstrate that radiomics is a topic of current interest for the management of HCC.

For radiomics to be a promising option for personalized medicine, it becomes clear that the methods of analysis should be standardized and automated. Radiomics is particularly interesting in the case of HCC because this tumor has an extremely varied phenotype, depending not only on

the tumor type but also on the underlying liver disease, and this phenotype has a direct impact on the evolution of cancer. New prospective studies integrating the potential clinical impact are, therefore, needed. The standardization of image acquisition methods and injection protocols is also essential to obtain more relevant results.

### Future perspectives in radiomics approaches for HCC

It is possible to yield additional accuracy with a standardization of CT scan, MRI, and PET protocols. Further developments may also potentially require high-resolution imaging modalities. As a consequence, radiomics features may become promising diagnostic and prognostic factors, in particular with a carefully conducted validation. However, it will always be preferable for radiomics studies to be conducted on large patient populations, ideally collected prospectively from multiple institutions.

One of the most important challenges radiomics has to face is the segmentation step. An ideal segmentation will define correctly the target region in the image with high reproducibility and at a low cost. However, this ideal scenario is far from being achieved. Currently used manual drawings of the tumor region lead to a high rate of disagreement among interpreters, missing crucial information because of tumor heterogeneity [65]. The increasing number of publicly available liver image datasets and the development of machine and deep learning can help in automating liver and lesion segmentation [66, 67].

While radiomics in HCC is in the early stages of development, recent work in biology has shown that variations in phenotype, such as those potentially observed through imaging, are at least as important as tumor genetics. In this context, the search for imaging biomarkers able to quantify variations in tumor phenotype remains a promising avenue for research. These new biomarkers will have to be built in concordance with the latest discoveries in HCC biology, in order to attempt capturing the changes that occur specifically at the interface between the tumor and the liver, in terms of immune and inflammatory reaction, as well as tumor heterogeneity. A better quality of radiomics analyses can be achieved using the entire tumor [68] plus the peritumoral environment with a three-dimensional analysis. An analysis of the whole liver and factors affecting its structure, its baseline signal, and vascularization should also be associated with the tumor analysis [61]. This is needed because liver cancer is not an isolated cancer, but occurs, in most cases, in a pre-existing chronic liver disease. To do this, we must develop computer analysis tools specific to the tumor under investigation, while also integrating the adjacent hepatic tissue into the analysis. Furthermore, there is a need to use more complex image analysis methods—including artificial

intelligence—that are more specific than the simple accumulation of a large number of very generic and non-specific features used in “classical” radiomics.

### Utilization of deep learning in radiomic analysis

In the case of conventional radiomics, the features mined by the discovery algorithm are designed by experts in medical image processing. However, a new class of artificial intelligence method known as deep learning may replace this approach [69]. Deep learning radiomics automatically identify—without human intervention—the best characteristics for a specific task [70–72] without the need for tumor segmentation. However, regardless of the image analysis method used, it is essential to create public image databases of patients with chronic liver diseases, with or without cancer, and make them accessible to researchers. This will make it possible to improve patient prognosis and to anticipate response to therapy for patient stratification. Unfortunately, to our knowledge, there is only one open access database fulfilling those criteria for the liver [73].

### Summary

In summary, radiomics is at its very early stages in HCC and many challenges need to be addressed. Nevertheless, recent pilot studies using radiomics in patients with HCC have shown their potential. For diagnosis, radiomics may help to characterize pathological and molecular liver lesions. For prognosis, image features could be independent prognostic factors, as they can be associated with tumor biological characteristics. By estimating treatment response, radiomics analysis may also help to pave the way for personalized medicine. Additionally, there is a need for prospective evaluations to allow for potential clinical applications. As shown in other cancer entities, radiomics may be an appropriate option for personalized treatment, as a non-invasive approach which can complement or replace tumor biopsy, and which can also be used to develop novel prognostic biomarkers in HCC patients.

**Acknowledgements** The authors acknowledge the support of ARC, Paris and Institut hospitalo-universitaire, Strasbourg (TheraHCC IHUARC IHU201301187), as well as the European Union (ERC-AdG-2014-671,231-HEPCIR, H2020-667273-HEPCAR). In addition, the authors are grateful to Camille Goustiaux, Christopher Burel, and Guy Temporal for their assistance in proofreading the manuscript.

**Author contributions** TW and BG designed the research; TW and FO extracted the data; TW, PS and BG wrote the paper; CG, EF, AS, VA, TFB, PP, and BG edited the paper; JM supervised the paper; All authors read and approved the final manuscript.

### Compliance with ethical standards

**Conflict of interest** Thomas F. Baumert, Patrick Pessaux, Jacques Marescaux, and Benoit Gallix have received research grants from ARC, Paris and Institut hospitalo-universitaire, Strasbourg (TheraHCC IHUARC IHU201301187). Antonio Saviano and Thomas F. Baumert have received research grants from the European Union (ERC-AdG-2014-671231-HEPCIR, H2020-667273-HEPCAR). Taiga Wakabayashi, Farid Ouhmich, Cristians Gonzalez-Cabrera, Emanuele Felli, Vincent Agnus, and Peter Savadjiev declare that they have no conflict of interest.

### References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
2. Hiley C, de Bruin EC, McGranahan N, Swanton C. Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine. *Genome Biol* 2014;15:453.
3. Lin DC, Mayakonda A, Dinh HQ, Huang P, Lin L, Liu X, Ding LW, et al. Genomic and Epigenomic Heterogeneity of Hepatocellular Carcinoma. *Cancer Res* 2017;77:2255–2265.
4. Lu LC, Hsu CH, Hsu C, Cheng AL. Tumor heterogeneity in hepatocellular carcinoma: facing the challenges. *Liver Cancer* 2016;5:128–138.
5. Martins-Filho SN, Paiva C, Azevedo RS, Alves VAF. Histological grading of hepatocellular carcinoma—a systematic review of literature. *Front Med (Lausanne)* 2017;4:193.
6. Mazzaferro V, Llovet JM, Miceli R, Bhoori S, Schiavo M, Mariani L, Camerini T, et al. Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. *Lancet Oncol* 2009;10:35–43.
7. Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis* 1999;19:329–338.
8. Prospective validation of the CLIP score: a new prognostic system for patients with cirrhosis and hepatocellular carcinoma. The Cancer of the Liver Italian Program (CLIP) Investigators. *Hepatology* 2000;31:840–845.
9. Farinati F, Rinaldi M, Gianni S, Naccarato R. How should patients with hepatocellular carcinoma be staged? Validation of a new prognostic system. *Cancer* 2000;89:2266–2273.
10. Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, Hasegawa H, Nakajima Y, et al. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. *Cancer* 1985;56:918–928.
11. Bruix J, Gores GJ, Mazzaferro V. Hepatocellular carcinoma: clinical frontiers and perspectives. *Gut* 2014;63:844–855.
12. Sadot E, Simpson AL, Do RK, Gonen M, Shia J, Allen PJ, D’Angelica MI, et al. Cholangiocarcinoma: correlation between molecular profiling and imaging phenotypes. *PLoS One* 2015;10:e0132953.
13. Sherman M, Bruix J. Biopsy for liver cancer: how to balance research needs with evidence-based clinical practice. *Hepatology* 2015;61:433–436.
14. Hricak H. Oncologic imaging: a guiding hand of personalized cancer care. *Radiology* 2011;259:633–640.
15. Sharma B, Martin A, Stanway S, Johnston SR, Constantinidou A. Imaging in oncology—over a century of advances. *Nat Rev Clin Oncol* 2012;9:728–737.

16. Tirkes T, Hollar MA, Tann M, Kohli MD, Akisik F, Sandrasegaran K. Response criteria in oncologic imaging: review of traditional and new criteria. *Radiographics* 2013;33:1323–1341.
17. Elsayes KM, Hooker JC, Agrons MM, Kielar AZ, Tang A, Fowler KJ, Chernyak V, et al. 2017 Version of LI-RADS for CT and MR imaging: an update. *Radiographics* 2017;37:1994–2017.
18. An C, Rakhmonova G, Choi JY, Kim MJ. Liver imaging reporting and data system (LI-RADS) version 2014: understanding and application of the diagnostic algorithm. *Clin Mol Hepatol* 2016;22:296–307.
19. European Association for the Study of the Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2018;69:182–236.
20. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;47:207–214.
21. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000;92:205–216.
22. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis* 2010;30:52–60.
23. Tang A, Bashir MR, Corwin MT, Cruite I, Dietrich CF, Do RKG, Ehman EC, et al. Evidence supporting LI-RADS major features for CT- and MR imaging-based diagnosis of hepatocellular carcinoma: a systematic review. *Radiology* 2018;286:29–48.
24. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14:749–762.
25. Cassinotto C, Dohan A, Zogopoulos G, Chiche L, Laurent C, Sa-Cunha A, Cuggia A, et al. Pancreatic adenocarcinoma: a simple CT score for predicting margin-positive resection in patients with resectable disease. *Eur J Radiol* 2017;95:33–38.
26. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48:441–446.
27. Lee G, Lee HY, Ko ES, Jeong WK. Radiomics and imaging genomics in precision medicine. *Precis Future Med* 2017;1:10–31.
28. Aerts HJ. The potential of radiomic-based phenotyping in precision medicine. A review. *JAMA Oncol* 2016;2:1636–1642.
29. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;5:4006.
30. Larue RT, Defraene G, De Ruyscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br J Radiol* 2017;90:20160665.
31. O'Connor JP, Rose CJ, Waterton JC, Carano RA, Parker GJ, Jackson A. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. *Clin Cancer Res* 2015;21:249–257.
32. Savadjiev P, Chong J, Dohan A, Vakalopoulou M, Reinhold C, Paragios N, Gallix B. Demystification of AI-driven medical image interpretation: past, present and future. *Eur Radiol* 2019;29(3):1616–1624.
33. Kim KA, Kim MJ, Jeon HM, Kim KS, Choi JS, Ahn SH, Cha SJ, et al. Prediction of microvascular invasion of hepatocellular carcinoma: usefulness of peritumoral hypointensity seen on gadoxetate disodium-enhanced hepatobiliary phase images. *J Magn Reson Imaging* 2012;35:629–634.
34. Renzulli M, Brocchi S, Cucchetti A, Mazzotti F, Mosconi C, Sportoletti C, Brandi G, et al. Can current preoperative imaging be used to detect microvascular invasion of hepatocellular carcinoma? *Radiology* 2016;279:432–442.
35. Park JH, Kim DH, Kim SH, Kim MY, Baik SK, Hong IS. The clinical implications of liver resection margin size in patients with hepatocellular carcinoma in terms of positron emission tomography positivity. *World J Surg* 2018;42:1514–1522.
36. Kuo MD, Gollub J, Sirlin CB, Ooi C, Chen X. Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma. *J Vasc Interv Radiol* 2007;18:821–831.
37. Peng J, Zhang J, Zhang Q, Xu Y, Zhou J, Liu L. A radiomics nomogram for preoperative prediction of microvascular invasion risk in hepatitis B virus-related hepatocellular carcinoma. *Diagn Interv Radiol* 2018;24:121–127.
38. Segal E, Sirlin CB, Ooi C, Adler AS, Gollub J, Chen X, Chan BK, et al. Decoding global gene expression programs in liver cancer by noninvasive imaging. *Nat Biotechnol* 2007;25:675–680.
39. Banerjee S, Wang DS, Kim HJ, Sirlin CB, Chan MG, Korn RL, Rutman AM, et al. A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma. *Hepatology* 2015;62:792–800.
40. Taouli B, Hoshida Y, Kakite S, Chen X, Tan PS, Sun X, Kihira S, et al. Imaging-based surrogate markers of transcriptome subclasses and signatures in hepatocellular carcinoma: preliminary results. *Eur Radiol* 2017;27:4472–4481.
41. Zheng BH, Liu LZ, Zhang ZZ, Shi JY, Dong LQ, Tian LY, Ding ZB, et al. Radiomics score: a potential prognostic imaging feature for postoperative survival of solitary HCC patients. *BMC Cancer* 2018;18:1148.
42. Akai H, Yasaka K, Kunitatsu A, Nojima M, Kokudo T, Kokudo N, Hasegawa K, et al. Predicting prognosis of resected hepatocellular carcinoma by radiomics analysis with random survival forest. *Diagn Interv Imaging* 2018;99:643–651.
43. Zhou Y, He L, Huang Y, Chen S, Wu P, Ye W, Liu Z, et al. CT-based radiomics signature: a potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma. *Abdom Radiol (NY)* 2017;42:1695–1704.
44. Chen S, Zhu Y, Liu Z, Liang C. Texture analysis of baseline multiphasic hepatic computed tomography images for the prognosis of single hepatocellular carcinoma after hepatectomy: a retrospective pilot study. *Eur J Radiol* 2017;90:198–204.
45. Li M, Fu S, Zhu Y, Liu Z, Chen S, Lu L, Liang C. Computed tomography texture analysis to facilitate therapeutic decision making in hepatocellular carcinoma. *Oncotarget* 2016;7:13248–13259.
46. Raman SP, Schroeder JL, Huang P, Chen Y, Coquia SF, Kawamoto S, Fishman EK. Preliminary data using computed tomography texture analysis for the classification of hypervascular liver lesions: generation of a predictive model on the basis of quantitative spatial frequency measurements—a work in progress. *J Comput Assist Tomogr* 2015;39:383–395.
47. Savadjiev P, Chong J, Dohan A, Agnus V, Forghani R, Reinhold C, Gallix B. Image-based biomarkers for solid tumor quantification. *Eur Radiol* 2019. doi:10.1007/s00330-019-06169-w
48. Scrivener M, de Jong EEC, van Timmeren JE, Pieters T, Ghaye B, Geets X. Radiomics applied to lung cancer: a review. *Transl Cancer Res* 2016;5:398–409.
49. Valdora F, Houssami N, Rossi F, Calabrese M, Tagliafico AS. Rapid review: radiomics and breast cancer. *Breast Cancer Res Treat* 2018;169:217–229.
50. Grossmann P, Gutman DA, Dunn WD, Jr., Holder CA, Aerts HJ. Imaging-genomics reveals driving pathways of MRI derived volumetric tumor phenotype features in Glioblastoma. *BMC Cancer* 2016;16:611.
51. Bai HX, Lee AM, Yang L, Zhang P, Davatzikos C, Maris JM, Diskin SJ. Imaging genomics in cancer research: limitations and promises. *Br J Radiol* 2016;89:20151030.

52. Pinker K, Shitano F, Sala E, Do RK, Young RJ, Wibmer AG, Hricak H, et al. Background, current role, and potential applications of radiogenomics. *J Magn Reson Imaging* 2018;47:604–620.
53. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016;278:563–577.
54. Jeong WK, Jamshidi N, Felker ER, Raman SS, Lu DS. Radiomics and radiogenomics of primary liver cancers. *Clin Mol Hepatol* 2019;25(1):21–29.
55. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6(7):e1000097.
56. Wu M, Tan H, Gao F, Hai J, Ning P, Chen J, Zhu S, et al. Predicting the grade of hepatocellular carcinoma based on non-contrast-enhanced MRI radiomics signature. *Eur Radiol* 2019;29(6):2802–2811.
57. Zhou W, Zhang L, Wang K, Chen S, Wang G, Liu Z, Liang C. Malignancy characterization of hepatocellular carcinomas based on texture analysis of contrast-enhanced MR images. *J Magn Reson Imaging* 2017;45:1476–1484.
58. Miura T, Ban D, Tanaka S, Mogushi K, Kudo A, Matsumura S, Mitsunori Y, et al. Distinct clinicopathological phenotype of hepatocellular carcinoma with ethoxybenzyl-magnetic resonance imaging hyperintensity: association with gene expression signature. *Am J Surg* 2015;210:561–569.
59. Hectors SJ, Wagner M, Bane O, Besa C, Lewis S, Remark R, Chen N, et al. Quantification of hepatocellular carcinoma heterogeneity with multiparametric magnetic resonance imaging. *Sci Rep* 2017;7:2452.
60. Starmans MPA, Miclea RL, van der Voort SR, Niessen WJ, Thomeer MG, Klein S: Classification of malignant and benign liver tumors using a radiomics approach. In: Angelini ED, Landman BA, eds. *Medical imaging 2018: image processing*, vol 10574. Bellingham: Spie-Int Soc Optical Engineering, 2018 (**Epub ahead of print**).
61. Blanc-Durand P, Van Der Gucht A, Jreige M, Nicod-Lalonde M, Silva-Monteiro M, Prior JO, Denys A, et al. Signature of survival: a (18)F-FDG PET based whole-liver radiomic analysis predicts survival after (90)Y-TARE for hepatocellular carcinoma. *Oncotarget* 2018;9:4549–4558.
62. Xia W, Chen Y, Zhang R, Yan Z, Zhou X, Zhang B, Gao X. Radiogenomics of hepatocellular carcinoma: multiregion analysis-based identification of prognostic imaging biomarkers by integrating gene data-a preliminary study. *Phys Med Biol* 2018;63:035044.
63. Bakr S, Echegaray S, Shah R, Kamaya A, Louie J, Napel S, Kothary N, et al. Noninvasive radiomics signature based on quantitative analysis of computed tomography images as a surrogate for microvascular invasion in hepatocellular carcinoma: a pilot study. *J Med Imaging (Bellingham)* 2017;4:041303.
64. Cozzi L, Dinapoli N, Fogliata A, Hsu WC, Reggiori G, Lobjalo F, Kirienko M, et al. Radiomics based analysis to predict local control and survival in hepatocellular carcinoma patients treated with volumetric modulated arc therapy. *BMC Cancer* 2017;17:829.
65. Echegaray S, Gevaert O, Shah R, Kamaya A, Louie J, Kothary N, Napel S. Core samples for radiomics features that are insensitive to tumor segmentation: method and pilot study using CT images of hepatocellular carcinoma. *J Med Imaging (Bellingham)* 2015;2:041011.
66. Parekh VS, Jacobs MA. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. *NPJ Breast Cancer* 2017;3:43.
67. Papp L, Potsch N, Grahovac M, Schmidbauer V, Woehrer A, Preusser M, Mitterhauser M, et al. Glioma survival prediction with combined analysis of in vivo (11)C-MET PET features, ex vivo features, and patient features by supervised machine learning. *J Nucl Med* 2018;59:892–899.
68. Ng F, Ganeshan B, Kozarski R, Miles KA, Goh V. Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. *Radiology* 2013;266:177–184.
69. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436–444.
70. Chartrand G, Cheng PM, Vorontsov E, Drozdal M, Turcotte S, Pal CJ, Kadoury S, et al. Deep learning: a primer for radiologists. *Radiographics* 2017;37:2113–2131.
71. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak J, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60–88.
72. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer* 2018;18:500–510.
73. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, et al. The cancer imaging archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 2013;26:1045–1057.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

---

7.4 Turon-Lagot V, **Saviano A**, Schuster C, Verrier ÉR. Hepatitis D virus: viral cycle and new therapeutic approaches. *Virologie*. 2019 Jun 1;23(3):149-159. French.

# Virus de l'hépatite D : cycle viral et nouvelles stratégies thérapeutiques

Vincent Turon-Lagot

Antonio Saviano

Catherine Schuster

Éloi R. Verrier

Université de Strasbourg, Inserm,  
Institut de recherche sur les maladies  
virales et hépatiques UMR\_S1110,  
F-67000 Strasbourg, France

**Résumé.** De récentes estimations suggèrent qu'environ 70 millions de personnes sont infectées par le virus de l'hépatite D (ou delta, HDV). HDV est un petit virus satellite du virus de l'hépatite B (HBV) capable d'achever son cycle viral uniquement en présence de ce dernier. L'hépatite D est la forme la plus sévère d'hépatite virale chronique. Elle est responsable d'une aggravation et d'une accélération de la progression de la maladie hépatique, en comparaison des patients mono-infectés par le HBV. Les traitements actuels, basés sur l'interféron pégylé sont peu efficaces et ne permettent que rarement l'élimination définitive du virus. L'absence d'un modèle d'étude simple a longtemps enfreint la compréhension des interactions HDV-hépatocytes, et notamment l'identification de facteurs hépatiques impliqués dans le cycle viral. Ces facteurs sont des cibles d'intérêt pour le développement de nouvelles stratégies thérapeutiques dont certaines sont en cours d'essai clinique. Cette revue résume les connaissances actuelles de la virologie moléculaire du HDV et fait le point sur les nouvelles solutions thérapeutiques en cours de développement.

**Mots clés :** HDV, virus hépatiques, stratégies antivirales

**Abstract.** An estimated 70 million people are chronically infected with hepatitis D (delta) virus (HDV) worldwide. HDV is a small satellite virus of hepatitis B virus (HBV) requiring HBV for the completion of its cycle. Hepatitis D is the most severe form of chronic viral hepatitis. It is responsible for an acceleration and an aggravation of chronic liver disease compared to HBV mono-infected patients. Current treatments based on pegylated interferon rarely allow viral clearance in chronically infected patients. For long time, the absence of easy-to-use models has limited the knowledge on virus-host interactions. Notably, hepatocyte host factors involved in the viral life cycle remain largely unknown. These host factors are potential therapeutic targets for novel treatment strategies, including molecules currently evaluated in clinical trials. This review summarizes our knowledge on HDV molecular virology and innovative therapeutic strategies targeting hepatocyte host factors.

**Key words:** HDV, hepatotropic viruses, antiviral strategies

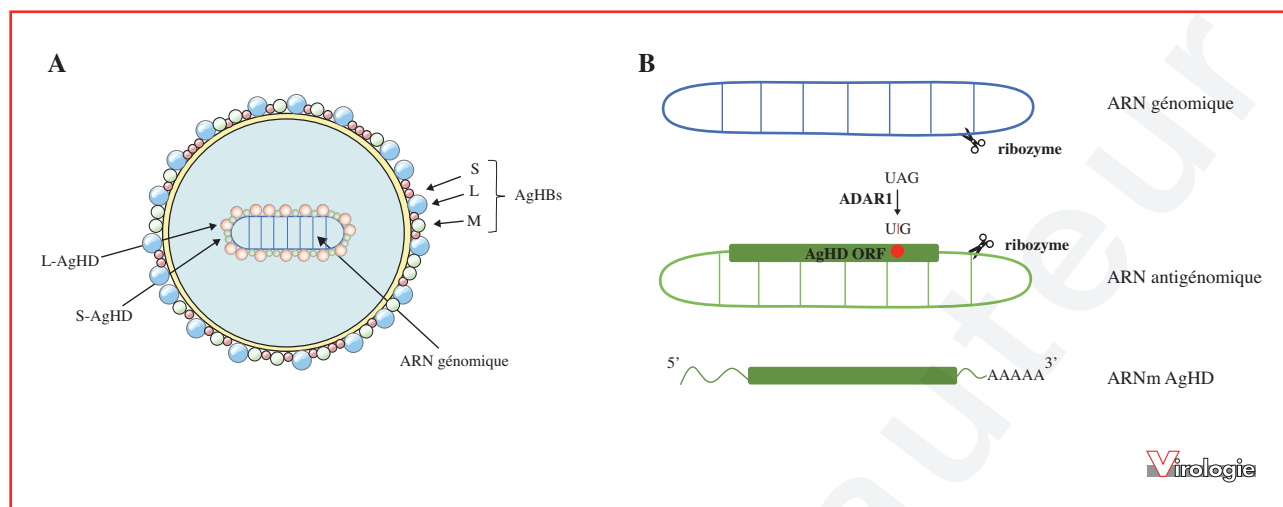
## Introduction : épidémiologie

L'hépatite D ou delta est considérée comme la forme la plus sévère d'hépatite virale chronique. Identifié initialement chez des patients atteints d'hépatite B chronique (HBC) comme un nouvel antigène du virus de l'hépatite B (HBV pour *hepatitis B virus*) [1], l'agent pathogène responsable est un petit virus à ARN infectant exclusivement les

hépatocytes et nécessitant la présence de HBV pour achever son cycle répliatif [2, 3]. Identifié aujourd'hui comme le plus petit virus capable d'infecter les animaux, le virus de l'hépatite D (HDV pour *hepatitis D virus*) est responsable d'une accélération et d'une aggravation de la progression de la maladie hépatique chez les patients souffrant d'HBC [4], déjà première cause mondiale de carcinome hépatocellulaire [5]. Initialement, il était estimé que l'hépatite D touchait environ 5 % des patients chroniquement infectés par le HBV, soit entre 15 et 20 millions de patients dans le monde [4]. Une récente

**Correspondance :** E.R. Verrier  
<e.verrier@unistra.fr>





**Figure 1. Structure du virus de l'hépatite D (HDV).** (A) Représentation schématique de la particule virale du HDV. Le virion est composé d'une enveloppe lipidique dans laquelle sont enchâssées les trois formes des protéines d'enveloppe du virus de l'hépatite B (HBV) qui forment l'AgHBs (S, M et L). Le génome du HDV est composé d'un ARN simple brin circulaire à polarité négative entouré par les deux formes de l'antigène delta (L-AgHD et S-AgHD) formant le complexe ribonucléoprotéique. (B) Le génome du HDV est fortement apparié (> 70 % d'appariement) induisant une structure en bâtonnet. Il contient une séquence ribozyme qui catalyse la fragmentation des différentes unités de génome produites au cours de la réplication en cercles roulants. La transcription de l'ARN génomique permet la synthèse de deux ARNs différents : l'ARN antigénomique et l'ARNm AgHD. L'ARN antigénomique contient également une séquence ribozyme et permet la néosynthèse d'ARN génomique. Il contient également un site d'édition. L'enzyme ADAR1 catalyse la modification d'une adénosine (A) du codon stop de l'ORF S-AgHD en inosine (I). Cette édition de l'ARN induit la production de la seconde forme de AgHD, plus longue de 19 acides aminés.

méta-analyse présente une prévalence mondiale de 0,98 % inégalement répartie à travers le globe, correspondant à une prévalence de 10,58 % chez les patients HBV [6]. Même si les deux chiffres sont difficilement conciliables (de 62 millions à 27 millions de patients en fonction de l'approche), cette étude suggère un nombre de patients infectés par le HDV dans le monde bien plus important que les études précédentes. Certaines régions telles le bassin méditerranéen (27,8 % des patients atteints d'hépatite) [7], l'Afrique du Nord (20,7 % des patients atteints d'hépatite) [8] et l'Afrique centrale (38 % des patients atteints d'hépatite) [9] sont particulièrement touchées. En Europe et aux États-Unis, une majorité de patients infectés par le HDV sont des usagers de drogue injectable [6]. En dépit de la généralisation de la vaccination anti-HBV, la prévalence du HDV augmente dans les pays développés, ce rebond étant notamment dû à l'immigration depuis les régions endémiques mais également au manque de vigilance des populations à risque à l'égard des virus hépatiques [10].

À l'heure actuelle, aucun traitement ne permet efficacement d'éliminer le virus [4]. La recherche de nouvelles solutions thérapeutiques a longtemps été freinée par le manque de connaissances des interactions HDV-hépatocytes, en grande partie dû à l'absence de modèle d'étude simple. Des découvertes récentes sont à l'origine d'avancées significatives dans la compréhension du cycle viral, notamment dans

l'identification de nouveaux facteurs hépatiques impliqués dans le cycle viral, donnant lieu à l'émergence de nouvelles stratégies de traitement, dont certaines sont actuellement testées en essai clinique.

## Virologie moléculaire du HDV

### Structure des virions

Le HDV est un petit virus enveloppé satellite du HBV d'environ 35 nm de diamètre (*figure 1*) [2]. Il est caractérisé par la présence d'une enveloppe comportant les antigènes de surface du HBV (AgHBs), indispensables à la formation des particules virales, qui contient le génome viral circulaire associé aux deux formes de l'antigène delta (AgHD) formant un complexe ribonucléoprotéique (RNP) [2, 11]. Cet ARN circulaire de polarité négative permet l'expression d'ARNm codant pour les formes, courte (S-AgHD) et longue (L-AgHD) de l'antigène delta, toutes deux impliquées dans le cycle réplcatif et la formation de la RNP virale [4]. Cette structure très particulière lui confère une place à part dans la classification des virus, seul représentant du genre non classé des Deltavirus. Le HDV est décliné en 8 génotypes hétérogènes inégalement répartis sur l'ensemble du globe [12]. Si le génotype 1 est retrouvé dans le monde entier, les génotypes 5 et 8 prédo-

minent en Afrique. Les génotypes 2 et 4, associés à une maladie hépatique moins agressive [13] sont retrouvés en Extrême-Orient. Enfin, le génotype 3, associé à un risque accru d'insuffisance hépatique est principalement présent en Amérique du Sud [12, 13].

### *Origine du HDV : évolution depuis un viroïde de plante ou d'un ARN cellulaire ?*

L'origine du HDV n'est à ce jour pas connue mais deux hypothèses majeures coexistent depuis une vingtaine d'années et proposent que HDV dérive soit de viroïdes de plantes soit d'un ARNm cellulaire.

Parmi les caractéristiques du génome du HDV, sa structure circulaire et son fort taux d'appariement (*figure 1B*) le rapprochent fortement des génomes de viroïdes de plante, petits ARN circulaires simple brin à haut potentiel réplcatif, de taille cependant bien inférieure au génome du HDV, comprise entre 250 et 400 nucléotides [14-16]. Une autre différence notable est la présence sur l'antigénome du HDV d'un cadre ouvert de lecture codant pour l'antigène delta, alors que les génomes de viroïdes (à l'exception du scRYMV [17]) ne codent pour aucune protéine. En revanche, le mécanisme de réplcation du génome, intégralement tributaire des ARN polymérases ADN dépendantes eucaryotes, est relativement similaire au mécanisme retrouvé chez certains viroïdes, comme les Avsunviroïdes [16]. Contrairement au HDV, les viroïdes de plantes ne s'assemblent pas en virions et se transmettent d'une plante à une autre par les graines ou des blessures, et aux cellules voisines au sein d'une plante infectée *via* les plasmodesmes [18]. Jusqu'à ce jour, aucune transmission active de cellule à cellule n'a été observée pour le HDV. Malgré tout, ces similitudes suggèrent que le HDV et les viroïdes pourraient descendre d'un même ancêtre commun, ou que l'un soit le précurseur de l'autre. L'hypothèse que le HDV provienne d'un viroïde ayant acquis une séquence codante lors de son évolution a été proposée en 1996 après la découverte dans le génome humain d'un gène codant pour une protéine interagissant avec l'antigène du virus (DIPA pour *delta-interacting protein A*) et présentant 60 % de similarité avec la séquence de la protéine virale [19]. De cette découverte, Brazas et Ganem ont émis l'hypothèse que le HDV dérive d'un ARN semblable à un viroïde ayant capturé la séquence codant pour la protéine DIPA à partir de l'ARNm cellulaire. D'autres travaux suggèrent que le HDV pourrait dériver d'un ARN cellulaire. Il y a une dizaine d'années, une étude démontra la présence d'une activité ribozyme dans la séquence de l'ARNm codant pour la *cytoplasmic polyadenylation element-binding protein 3* (CPEB3) [20]. Ce ribozyme se trouve dans un intron de l'ARNm de CPEB3 et s'apparente au ribozyme du HDV par sa structure et son activité biochimique. Le fait que ce

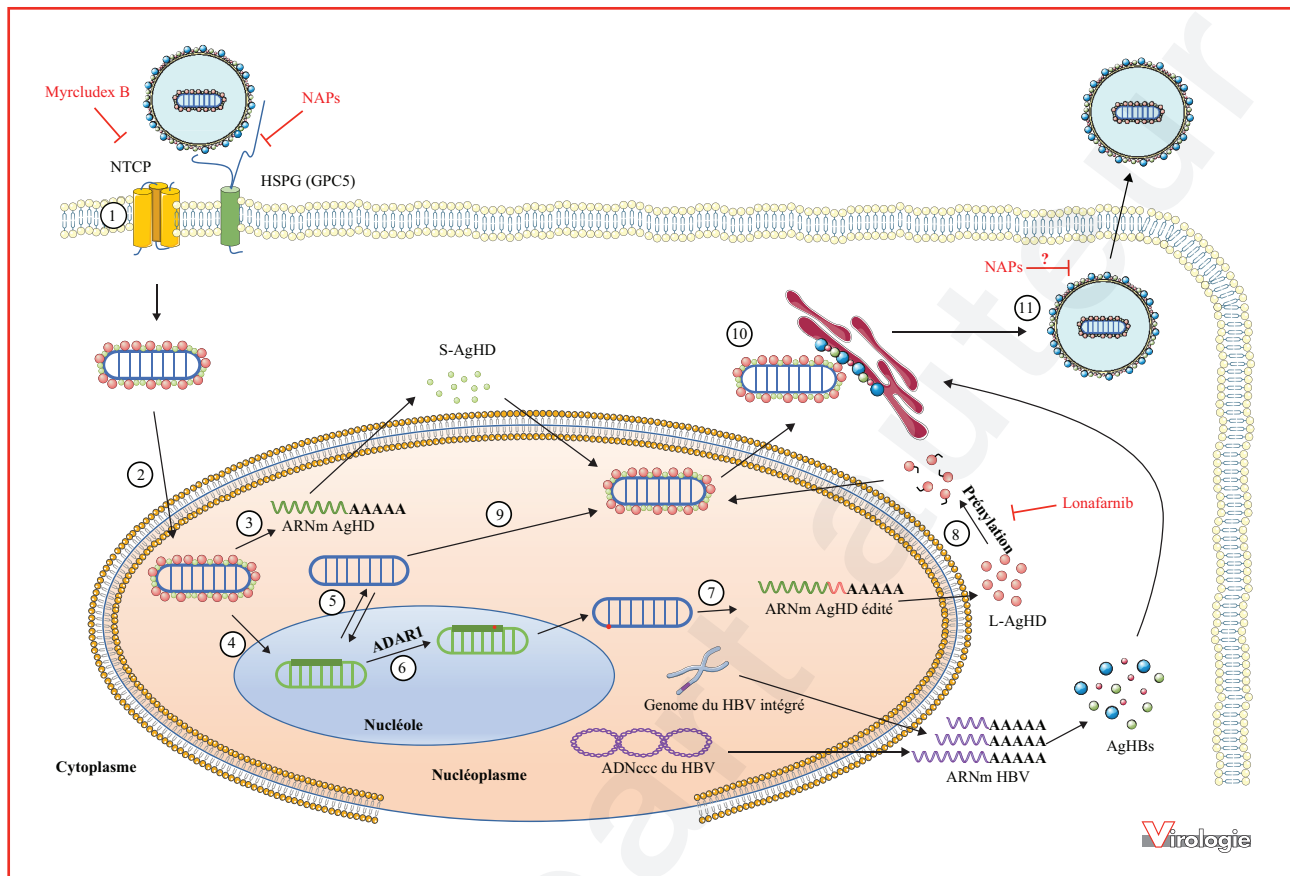
ribozyme ne soit présent que chez les mammifères a conduit les auteurs à émettre l'hypothèse que le génome du HDV pourrait dériver du transcriptome humain. L'un des arguments repose sur la mise en évidence des ARNs cellulaires circulaires (cARN) décrits dès le début des années 90 [21]. Ces cARN sont abondants et interviennent dans différents processus de régulation, comme la prolifération cellulaire et la progression cancéreuse, et interagissent notamment avec des microARN. Chaque cARN peut interagir avec plusieurs copies d'un même microARN et jouer ainsi un rôle d'« éponge » à microARN [22-24]. Toutefois, les cARN actuellement décrits ne comportent ni la séquence pouvant coder pour AgHD, ni l'activité ribozyme. De plus, la très récente description d'un virus partageant de nombreuses similarités avec le HDV chez les oiseaux, mais ne semblant pas dépendre d'un hépadnavirus pour l'achèvement de son cycle viral infirme l'hypothèse d'un HDV exclusivement humain [25].

### *Cycle viral du HDV*

Les virions HDV et HBV partagent les mêmes protéines d'enveloppe ; aussi, le mécanisme d'entrée dans les hépatocytes est supposé identique pour les deux virus. Le cycle viral est initié par l'attachement de la particule virale à la surface des hépatocytes *via* son interaction avec les chaînes de sucres des protéoglycanes à héparane sulfate (HSPG) [26], dont GPC5 [27] (*figure 2*). Il se lie ensuite de manière spécifique à son récepteur hépatocytaire, le transporteur d'acides biliaires NTCP (*sodium-taurocholate co-transporting polypeptide*), induisant l'entrée du virus dans la cellule [28]. Les 75 résidus myristoylés en N-terminal du domaine pré-S1 de la grande protéine d'enveloppe du HBV interagissent avec NTCP, potentiellement au niveau du site de liaison des acides biliaires [29, 30]. Un autre domaine du transporteur est impliqué dans l'entrée du virus mais sans interaction directe avec l'enveloppe virale, sa fonction dans l'entrée restant inconnue [28]. La liaison à NTCP induit l'endocytose de la particule virale selon un mécanisme encore inconnu. Le transport de la RNP du HDV du cytoplasme vers le noyau est également peu documenté, mais un signal NLS (signal de localisation nucléaire) a été identifié et caractérisé dans les deux formes d'AgHD [31].

Le génome du HDV ne codant pour aucune protéine non structurale, la réplcation virale et la transcription de l'ARNm sont entièrement dépendantes des polymérases cellulaires. Le génome sert de matrice pour la transcription d'ARNm du HDV par l'ARN polymérase II. Une première forme d'ARNm est exportée dans le cytoplasme permettant la production de la forme courte S-AgHD. La protéine S-AgHD est importée dans le noyau et stimule la réplcation virale [32], au cours de laquelle le génome





**Figure 2. Cycle viral du HDV.** (1) Le cycle viral débute par l'attachement aux protéoglycanes à sulfate d'héparane (HSPG), dont Glypican 5 (GPC5) à la surface des hépatocytes. La région pré-S1 de L-AgHBs se lie ensuite au récepteur spécifique du HBV et du HDV, le transporteur d'acides biliaires, NTCP. La particule virale est endocytée et la RNP virale est libérée dans le cytoplasme. (2) Elle est ensuite acheminée au noyau grâce au signal de localisation nucléaire présent sur les AgHD. (3) L'ARN polymérase II transcrit l'ARNm AgHD qui est ensuite exporté dans le cytoplasme où il est traduit pour produire la forme courte « *Small* » de AgHD (S-AgHD). (4) L'ARN polymérase II synthétise l'ARN antigénomique du HDV qui est ensuite transféré dans le nucléole. (5) Dans le nucléole, l'ARN antigénomique sert de matrice pour la néosynthèse d'ARN génomique par un mécanisme de cercle roulant. (6) L'ARN antigénomique est édité par l'action de l'enzyme ADAR1, supprimant le codon stop de l'ORF S-AgHD. (7) L'ARN antigénomique édité est transcrit en ARN génomique, induisant la synthèse de l'ARNm AgHD édité plus long. Ce dernier est exporté dans le cytoplasme où il est traduit en forme longue « *Large* » de AgHD (L-AgHD). (8) L-AgHD contient un site de prénylation qui est farnésylé par une farnésyltransférase cellulaire avant d'être transféré dans le noyau. (9) Les deux formes d'AgHD interagissent avec les ARN génomiques néo-synthétisés afin de former de nouvelles ribonucléoprotéines virales qui sont exportées dans le cytoplasme. (10) Les ribonucléoprotéines virales interagissent, *via* une cystéine farnésylée de L-AgHD, avec la partie cytosolique de l'enveloppe du HBV au niveau du réticulum endoplasmique permettant leur enveloppement. (11) Les virions néoformés sont excrétés de la cellule infectée. Les différents antiviraux sont indiqués en rouge. L'hépatocyte représenté est également infecté par le HBV, indiqué par la présence l'ADNccc ou de son génome intégré (cycle complet non représenté).

du HDV recrute l'ARN polymérase II pour la synthèse de multimères d'ARN antigénomique par un mécanisme en cercle roulant [4]. L'activité ribozyme de l'antigénome catalyse l'auto-clivage des multimères en monomères, qui sont ensuite circularisés. L'ARN antigénomique du HDV sert ensuite de matrice à la synthèse de nouvelles copies de génome du HDV, également selon un mécanisme en cercle roulant *via* l'ARN polymérase II [4]. Bien que la réplication du HDV soit majoritairement portée par l'ARN polymérase II, plusieurs études suggèrent une implication des ARN polymérases I et III dans la synthèse de l'ARN

antigénomique [33, 34]. Il est généralement admis que la transcription des ARN génomiques et antigénomiques a lieu dans deux compartiments nucléaires différents. La transcription de l'ARN génomique a lieu dans le nucléoplasme et celle de l'ARN antigénomique se déroule dans le nucléole [35]. Durant la phase tardive de réplication, l'ARN antigénomique du HDV est édité *via* l'enzyme ADAR1 (*adenosine deaminase acting on RNA 1*), une adénosine du codon stop de S-AgHD étant transformée en inosine [4] ce qui conduira après réplication au remplacement du codon stop en codon tryptophane (figure 2). La nouvelle version

des ARN génomiques produits sert alors de matrice pour la transcription d'ARNm codant pour une forme plus longue, L-AgHD. Cette allongement C-terminale de 19 acides aminés contient notamment un signal d'export nucléaire (NES) [36] et un site de prénylation sur la cystéine 211 qui est farnésylée *via* une farnésyltransférase cellulaire après traduction [37, 38]. L-AgHD farnésylé est transféré dans le noyau et inhibe la réplication virale, ce qui a pour conséquence d'orienter le cycle viral vers l'assemblage de nouveaux virions [39]. L'ARN génomique néoformé s'associe avec S-AgHD et L-AgHD afin de former de nouvelles RNP qui sont ensuite exportées du noyau, grâce au NES de L-AgHD, par la voie TAP/Aly [40]. La cystéine farnésylée de L-AgHD permet l'interaction des RNP virales avec la partie cytosolique de AgHBs à la membrane du réticulum endoplasmique. L'interaction entre la RNP virale et AgHBs induit l'enveloppement des nouveaux virions ainsi que leur sécrétion, selon des voies encore inconnues [41]. HDV est donc tributaire de l'expression d'AgHBs pour achever son cycle viral. Au sein d'un hépatocyte préalablement infecté par le HBV, AgHBs peut être exprimé à partir de l'ADNccc du HBV mais aussi d'une version intégrée du génome viral dans le génome de la cellule hôte (*figure 2*). La réplication active du HBV ne semble en effet pas indispensable à la production de virions HDV [42].

Si les grandes étapes du cycle sont à présent bien décrites, de nombreuses zones d'ombre persistent quant aux interactions moléculaires entre HDV et facteurs hépatiques dont le virus dépend fortement pour l'achèvement de son cycle réplcatif.

### Interactions virus-hôte

Comme le démontre la réplication de son génome, entièrement réalisée *via* les ARN polymérases, le HDV est extrêmement dépendant des facteurs des cellules hôtes pour l'accomplissement de son cycle réplcatif. Même si ces interactions sont encore largement méconnues, un certain nombre d'acteurs cellulaires ont été décrits. Récemment, un criblage génomique a montré la forte dépendance du HDV à la biosynthèse des pyrimidines. En effet, l'inhibition de l'enzyme CAD (pour *Carbamyl-phosphate synthase II / Aspartate carbamoyltransférase / Dihydroorotase*), qui catalyse les trois premières étapes de la synthèse de l'uridine, inhibe fortement la réplication du HDV *in vitro*, sans impacter la réplication du HBV dans le laps de temps de l'étude [43]. De plus, l'inhibition du récepteur alpha des œstrogènes (ESR1), qui régule l'expression de CAD, inhibe également la réplication virale. En plus de la synthèse des pyrimidines, les résultats du criblage suggèrent l'importance d'autres facteurs cellulaires dans l'infection HDV, notamment des gènes impliqués dans la résistance à l'insuline ou la voie de signalisation HIF-1 (*hypoxia-*

*inducible factor*). À ce jour, il est décrit que l'AgHD interagit avec plus d'une centaine de protéines, dont beaucoup sont impliquées dans le métabolisme de l'ARN, mais également des hélicases à ARN, l'histone H1, ou encore des sous-unités de l'ARN polymérase II [44]. Les ARNs du HDV se lient également avec des protéines cellulaires, notamment la protéine kinase R ou ADAR1, qui catalyse l'édition de l'ARN antigénomique [45-47]. Au niveau des senseurs de l'immunité innée, le génome du HDV est détecté par l'hélicase MDA5 (*melanoma differentiation associated gene 5*), induisant la production d'IFN de type I et III [48]. La production d'ISG (*interferon stimulated gene*) n'a cependant que peu d'effet sur la réplication du HDV, qui inhibe activement les voies de signalisation de la réponse immunitaire innée en inhibant la phosphorylation de STAT1 (*signal transducer and activator of transcription 1*) et STAT2, et par conséquent l'expression des ISG [49]. En revanche, L-AgHD induit l'activation de NF- $\kappa$ B (*nuclear factor kappa B*) et de STAT3 *via* le stress oxydatif [50]. L'activation de ces facteurs de transcription participe à l'état inflammatoire hépatique et pourrait expliquer l'aggravation et l'accélération de la maladie hépatique chez le patient.

Par ailleurs, l'infection par le HDV induit chez les patients une réponse immunitaire adaptative soutenue. En effet, une étude menée sur des patients atteints d'hépatite virale HBV, HBV/HDV ou virus de l'hépatite C (HCV pour *hepatitis C virus*) a montré que les patients HDV présentent les taux les plus élevés de lymphocytes T CD4<sup>+</sup> perforine-positifs [51]. Ces lymphocytes éliminent les cellules infectées et jouent probablement un rôle dans la progression plus rapide de la maladie hépatique chez les patients HDV. L'infection induit également une réponse immunitaire *via* les lymphocytes T CD8<sup>+</sup> [52], mais un échappement viral par mutations est observé chez les patients, limitant le contrôle de l'infection. Ces lymphocytes T CD8<sup>+</sup> reconnaissant des épitopes du HDV sont retrouvés à la fois chez les patients chroniques et chez les patients ayant résolu l'infection, suggérant que ces mutations pourraient expliquer au moins en partie la mise en place d'une infection chronique, les mutants HDV échappant au contrôle des cellules T CD8<sup>+</sup> [52].

## Importance clinique

### Histoire naturelle

Chez le patient, il existe deux types d'infection par le HDV : 1) la co-infection initiale simultanée d'un patient sain par le HBV et le HDV, qui évolue chez la majorité des adultes vers l'élimination spontanée des virus (> 95 % des cas), comme lors d'une mono-infection HBV [53], avec cependant un risque accru d'hépatite fulminante [54, 55] ; 2) la surinfection HDV chez un patient HBC, qui évolue majoritairement vers une infection persistante et une

hépatite D chronique (environ 80 % des cas), caractérisée par une inflammation et une fibrose hépatiques progressant avec un risque trois fois plus élevé vers la cirrhose que chez les patients HBC [53, 56]. De plus, le risque de survenue d'un carcinome hépatocellulaire est trois fois plus élevé que pour des patients HBC [57]. En comparaison des autres hépatites chroniques, l'hépatite D est marquée par une progression rapide vers la cirrhose, et un risque de mortalité plus élevé [4].

### Diagnostic

Actuellement, il existe trois méthodes de détection du HDV mais ces méthodes restent inégalement disponibles et de fiabilité variable. Lors de l'infection aiguë, le HDV exprime et sécrète fortement AgHD qui peut être détecté par ELISA. Cependant, ce test n'est réalisable que lors des deux premières semaines d'infection, l'AgHD n'étant ensuite exprimé que de manière transitoire [58, 59].

La seconde méthode consiste à détecter les anticorps anti-HDV [60]. Les IgM sont les premières produites, concomitamment aux IgM anti-HBc (anti-*HBV core protein*) dans le cas d'une co-infection. Les IgG anti-HDV produites ensuite persistent dans le sérum des patients, que l'infection aiguë ait été résolue ou qu'elle soit devenue chronique. Chez les patients positifs pour AgHBs, la détection des IgG et IgM (Ig totaux) anti-HDV est généralement la première phase de diagnostic d'une infection HDV, même si le risque de faux-négatifs existe [10]. De plus, la détection d'IgG anti-HDV ne permet actuellement pas de conclure à une répllication active du virus, les IgG persistant en cas d'infection résolue spontanément. Une nouvelle méthode récemment développée, Q-MAC (*quantitative microarray antibody capture*), en plus d'être largement plus sensible et spécifique que les précédentes, permettrait, en fonction de l'intensité du signal, de discriminer les patients présentant ou non une activité répllicative du HDV, même si ce test nécessite encore des études sur des cohortes plus larges [61, 62].

La troisième méthode consiste à détecter l'ARN viral du HDV dans le sérum des patients par qRT-PCR, seul marqueur d'une répllication active du HDV dans les hépatocytes. La détection de l'ARN viral, après détection des Ig totaux anti-HDV, dans le sérum permet de distinguer une infection aiguë résolue, d'une infection chronique. Cependant, les différents tests disponibles en fonction des pays n'ont pas la même sensibilité ni la même précision dans la détermination de la charge virale. De plus, le HDV ayant une forte variabilité génétique, les tests ne détectent pas efficacement les 8 génotypes viraux [63, 64]. L'Organisation mondiale de la santé (OMS) a récemment mis en place un ARN standard international afin de permettre aux laboratoires d'exprimer leurs résultats en unité internationale

(UI). Depuis, plusieurs kits de détection de la charge virale ont été développés permettant une meilleure détection de tous les génotypes ainsi qu'une meilleure reproductibilité. L'un de ces kits, l'Eurobioplex HDV kit, a montré sur des échantillons de patients une forte sensibilité, précision et reproductibilité, dans la détection de tous les génotypes du HDV [65]. Une mesure précise de la charge virale du HDV chez les patients est nécessaire car elle est le seul moyen de mesurer et d'analyser la répllication virale afin de suivre l'efficacité des traitements antiviraux.

### Traitements actuels

Les méthodes de détection du HDV et le suivi des patients évoluent mais les traitements actuels reposent toujours sur l'utilisation de l'interféron-alpha pégylé (PEG-IFN $\alpha$ ), peu spécifique et responsable d'effets secondaires marqués tels qu'état grippal, anémie, dépression, conduisant parfois à un arrêt anticipé du traitement. Les réponses au traitement, de l'ordre de 30 %, sont partielles et conduisent rarement à l'élimination persistante de la charge virale [66-68]. Il est à noter que les analogues de nucléo(s)tides (NUC) utilisés dans le traitement contre le HBV sont inefficaces contre le HDV [12]. Plusieurs éléments peuvent expliquer la difficulté de traitement du HDV. Premièrement, le HDV ne code pour aucune enzyme dans son génome dont la fonction pourrait être ciblée par un traitement antiviral direct. Si le ribozyme du génome présente bien une activité enzymatique, les tentatives d'utilisation d'inhibiteurs de ribozyme se sont heurtées à une très forte toxicité *in vitro*, limitant leur caractérisation antivirale [69]. Par ailleurs, en plus d'une forte diminution de la répllication du HDV, il est nécessaire d'inhiber l'expression d'AgHBs, responsable du rebond de charge virale HDV après arrêt du traitement PEG-IFN $\alpha$ , même si le sujet a développé une réponse virologique soutenue [70]. Cette inhibition n'est cependant observée que chez une minorité de patients à l'aide des traitements PEG-IFN $\alpha$  actuels (10 % environ) [70].

Par conséquent, de nouvelles stratégies de traitement du HDV sont attendues. Dans ce contexte, les molécules ciblant les facteurs hépatiques peuvent être de nouvelles armes antivirales, dont l'efficacité a été démontrée pour d'autres virus hépatiques [71-73]. Ce type de stratégie nécessite toutefois une connaissance approfondie des cycles viraux et des interactions moléculaires entre virus et facteurs cellulaires.

### Perspectives thérapeutiques : molécules ciblant l'hôte

Actuellement, plusieurs traitements sont en phase avancée d'essai clinique (*tableau 1*). Ils ciblent différentes étapes

Tableau 1 Molécules antivirales en essai clinique.

|                                       | Cible cellulaire /<br>étape du cycle viral | Stade essai<br>clinique | Administration   | Avantages  | Inconvénients  | Références  |
|---------------------------------------|--|-------------------------|--|--|--|---|
| Myrcludex B                           | NTCP / inhibiteur<br>d'entrée              | Phase II                | Injection sous-cutanée<br>2, 5 ou 10 mg + 245 mg<br>Ténofovir par jour<br>(24 semaines)  | 1,6-2,7 log réduction<br>ARN HDV   | Rechute chez<br>60-83 % des patients   | Wedemeyer<br><i>et al.</i> , 2018                   |
| Lonafarnib (LNF)<br>+ Ritonavir (RTV) | Farnésylation /<br>inhibiteur d'assemblage | Phase II                | Orale<br>50-100 mg LNF + 100 mg<br>RTV par jour<br>(24 semaines)   | 1,6 log réduction ARN<br>HDV ; ARN HDV<br>indétectable chez<br>1 patient | Forte augmentation<br>des ALT après arrêt<br>du traitement chez<br>30 % des patients ;<br>pas de données sur<br>rechutes | Wedemeyer<br><i>et al.</i> , 2017                   |
| REP2139                               | ? / Sécrétion d'AgHBs                      | Phase II                | Injection intraveineuse<br>500 mg par semaine<br>(15 semaines), 250 mg<br>+ 180 µg Peg-IFN<br>(15 semaines), 180 µg<br>Peg-IFN (33 semaines) | ARN HDV indétectable<br>chez 7 patients sur 11                           | Forte augmentation<br>des ALT, pas de<br>données sur rechutes  | Bazinet <i>et al.</i> ,<br>2017 ; Vaillant,<br>2018 |

du cycle viral à savoir : l'entrée virale, l'assemblage des particules virales et la sécrétion d'AgHBs.

#### *Inhibiteur de l'entrée virale : l'exemple du Myrcludex B*

Le HBV et le HDV utilisent la même enveloppe virale et, de ce fait, le même récepteur, NTCP [74]. Ce récepteur spécifique du foie, a rapidement été considéré comme une cible thérapeutique d'intérêt après la découverte de son rôle dans l'entrée des deux virus (pour des revues, voir [74, 75]). Si de nombreux inhibiteurs de NTCP décrits ont montré une activité antivirale prometteuse, c'est un lipopeptide dérivé de la partie pré-S1 de l'enveloppe du HBV qui concentre l'essentiel des attentions. Même avant la découverte du récepteur, ce peptide était connu pour son activité préventive *in vivo*, inhibant l'infection par le HBV et le HDV dans un modèle murin [76]. Le peptide pré-S1, site de liaison de l'enveloppe virale au récepteur, se fixe spécifiquement sur NTCP [77-79]. La forme commerciale du peptide, le Myrcludex B, peptide myristoylé dérivant des 47 acides aminés en N-terminal du domaine pré-S1 d'AgHBs, a été testé pour son activité antivirale dans de nombreux modèles *in vitro* et *in vivo* et en essai clinique [80]. Dans cet essai, 24 patients HDV co-infectés HBV ont été divisés en trois groupes afin de recevoir, pendant 24 semaines, des injections sous-cutanées quotidiennes de 2 mg de Myrcludex B, couplé ou non au Peg-IFN $\alpha$ , comparé à un traitement Peg-IFN $\alpha$  seul. Le critère principal de l'étude reposait sur la mesure du taux d'AgHBs. Aucune baisse significative d'AgHBs n'a été observée chez les patients après traitement. Toutefois, dans le groupe de patients traités uniquement avec Myrcludex B, 6 patients (75 %) ont montré une stabilisation des ALT (alanine aminotransférase) et 4 patients (50 %) ont montré une baisse de l'ARN HDV sérique de plus d'un log. De plus, l'élimination du virus a été atteinte chez 2 patients (25 %). Le groupe traité avec Myrcludex B en combinaison au Peg-IFN $\alpha$  a montré de meilleurs résultats avec l'ARN HDV devenu indétectable chez 5 patients (62,5 %), en revanche l'ARN HDV est réapparu chez tous les patients après arrêt du traitement, quel que soit le traitement administré [81]. Plus récemment, les résultats d'un essai multicentrique ouvert de phase II ont été rendus publics [82]. Cet essai, réalisé sur une cohorte de 120 patients co-infectés HBV/HDV, avait pour but de déterminer la tolérance et l'efficacité d'un traitement composé de différentes doses de Myrcludex B en combinaison avec du ténofovir, un inhibiteur de la transcriptase inverse du HBV [83]. Les patients, divisés en quatre groupes, ont reçu quotidiennement 2, 5 ou 10 mg de Myrcludex B par injection sous-cutanée en combinaison avec du ténofovir (245 mg/jour), comparé à un traitement au ténofovir seul pendant 24 semaines. Après cette phase, tous les patients ont ensuite suivi un traitement au ténofovir pen-



dant 24 semaines. Le critère principal était une diminution de l'ARN HDV de 2 log ou une absence d'ARN viral détectable. À la fin du traitement, l'ARN HDV avait diminué en moyenne entre 1,6 et 2,7 log, la plus forte dose de Mycludex B étant reliée à la baisse d'ARN HDV la plus forte. Un suivi à 12 semaines sur une partie des patients a révélé la rechute de l'infection chez 60 % à 83 % des patients en fonction des groupes. Plus récemment, les résultats provisoires de l'étude de phase II MYR203 indiquent des niveaux indétectables de HDV chez 9 des 15 patients traités 48 semaines avec une dose quotidienne de Mycludex B (2 mg) en combinaison avec le PEG-IFN $\alpha$  [84]. Probablement insuffisant en monothérapie, l'utilisation du Mycludex B semble donc bénéfique chez les patients HDV en combinaison avec du ténofovir ou du PEG-IFN $\alpha$ . Cependant, quelques effets secondaires ont été enregistrés (démangeaisons, augmentation des acides sériques, etc.) et les questions en suspens restent la sécurité et la tolérance à long terme, notamment chez les patients cirrhotiques.

#### *Inhibiteur de l'assemblage : lonafarnib*

Durant le cycle viral du HDV, la protéine L-AgHD est farnésylée et cette étape précède l'interaction avec AgHBs. L'inhibition de la farnésylation de L-AgHD inhibe l'assemblage des nouvelles particules virales dans des modèles cellulaires et murins [85, 86]. Le lonafarnib est un inhibiteur de farnésylation qui a d'abord été testé comme anti-cancéreux et ayant des effets bénéfiques chez les patients atteints du syndrome de Hutchinson-Gilford (progeria) [87]. Bien que son efficacité n'ait pas été démontrée dans ce contexte, les premières études ont fourni des données de tolérance chez le patient. Un premier essai chez des patients HDV a été réalisé durant lequel deux doses de lonafarnib (100 mg et 200 mg) ont été administrées deux fois par jour par voie orale pendant 28 jours. Comparés à un placebo, les deux groupes de patients ont montré une baisse significative de l'ARN HDV (0,73 log pour le groupe 100 mg ; 1,54 log pour le groupe 200 mg). En revanche, le traitement n'a induit aucune baisse ni des ALT, ni des AgHBs, et le taux d'ARN HDV est revenu à la normale chez la totalité des patients à la fin de la période de suivi. De plus, tous les patients ayant reçu la plus forte dose de lonafarnib ont subi de forts effets secondaires tels que des diarrhées, des nausées et une perte de poids [88]. Afin d'améliorer l'absorption dans le sang de lonafarnib et diminuer les effets secondaires, quatre études de phases II appelées LOWR-HDV ont été réalisées avec un traitement lonafarnib en combinaison avec le ritonavir, un inhibiteur du cytochrome P450-3A4 qui est le principal acteur du métabolisme du lonafarnib et améliore sa stabilité sans effet antiviral direct [89]. De manière générale, ces études ont montré que la combinaison des deux traitements

permet de diminuer la dose de lonafarnib administrée quotidiennement et d'améliorer la tolérance chez le patient [90, 91]. Cette combinaison de traitement a une meilleure efficacité sur la diminution de l'ARN HDV mais cette forte baisse est en général observée pour la moitié des patients seulement, et aucune information sur le suivi de ces patients après arrêt du traitement n'est encore disponible.

#### *Les polymères d'acides nucléiques*

Les polymères d'acides nucléiques (NAP) sont des oligonucléotides phosphorothioés leur conférant une résistance à la dégradation et à la dénaturation *in vivo*. Ils possèdent une activité inhibitrice à large spectre contre plusieurs virus comme le HCV [92] ou le virus herpes simplex [93]. Bien que leur mécanisme d'action précis ne soit pas connu avec précision, différents NAPs inhibent l'entrée des particules HDV *in vitro* [94]. De plus, les NAPs semblent inhiber la sécrétion d'AgHBs, affectant potentiellement le cycle du HDV *via* divers mécanismes [95]. Leur activité est indépendante de leur séquence mais dépendante de leur taille et de leur hydrophobicité [95]. La première étude *in vivo* a été conduite chez des canards infectés par le HBV du canard (DHBV) et traités pendant 28 jours avec le NAP REP2055. L'étude a montré une baisse d'AgHBs dans le sérum et de l'ADN DHBV, jusqu'à 16 semaines post-traitement, ainsi qu'une augmentation des anticorps anti-DHBV [96]. La tolérance et l'efficacité de REP2055 ainsi que de REP2139, un dérivé de REP2055, ont été étudiées dans une étude portant sur des patients HBV positifs à AgHBe. Pour chacun des composés, le traitement en monothérapie a montré une baisse de AgHBs dans le sérum de 2 à 7 log et de l'ADN HBV de 3 à 9 log. De plus, les traitements ont été accompagnés d'une production d'anticorps anti-HBs [97]. Le NAP REP2139 ayant montré une meilleure tolérance chez les patients ainsi qu'une forte efficacité, il a ensuite été utilisé dans une nouvelle étude afin d'évaluer sa tolérance et son efficacité en co-traitement avec Peg-IFN $\alpha$  sur des patients co-infectés avec HBV et HDV. Les patients ont été injectés une fois par semaine par voie intraveineuse avec 500 mg de REP2139-Ca seul pendant 15 semaines. Le traitement a été suivi par 15 semaines de traitement avec 250 mg de REP2139-Ca combiné à 180  $\mu$ g de Peg-IFN $\alpha$  puis 33 semaines de traitement avec 180  $\mu$ g de Peg-IFN $\alpha$  seul [98]. Après traitement, les patients ont été suivis pendant deux ans et l'ARN HDV est resté indétectable chez 7 patients (64 %), avec les AgHBs et l'ADN HBV en dessous du seuil de détection chez 4 (36 %) et 6 (54 %) patients, respectivement [99]. Les résultats obtenus lors de ces études sont pour l'instant les plus convaincants, cependant les cohortes étudiées restent de petite taille et les effets restent à confirmer sur de plus grandes populations. De plus, le mode d'administration n'est pas adapté à un traitement

de longue durée. De nouvelles études seront menées afin de tester la tolérance du traitement en administration sous-cutanée. Enfin, le traitement au Peg-IFN $\alpha$  a induit chez 5 patients une forte augmentation des ALT [99]. Les patients de l'étude étant non-cirrhotiques, cette augmentation est restée asymptomatique et s'est résolue après arrêt du traitement. Le traitement pourrait cependant être plus délétère chez des patients cirrhotiques.

## Conclusion

L'hépatite D reste aujourd'hui incurable et représente une menace de santé publique majeure pour des millions de patients à travers le monde. Les données épidémiologiques actuelles sont des estimations approximatives car la présence du HDV chez les patients HBV est encore trop peu souvent recherchée dans certains pays. En plus d'une meilleure détection, les traitements nécessitent également d'être améliorés. Les avancées récentes sur le cycle viral ont permis l'émergence de nouvelles solutions thérapeutiques qui démontrent l'intérêt des molécules ciblant l'hôte dans le traitement contre ce virus. La caractérisation exhaustive des facteurs hépatocytaires impliqués dans le cycle viral permettra à terme le développement de nouvelles solutions thérapeutiques pour l'éradication de ce virus hépatique majeur.

**Remerciements.** Les travaux de l'unité résumés dans cette revue ont notamment été financés par l'Agence nationale de la recherche sur le sida et les hépatites virales (ANRS). Cette revue est écrite dans le cadre du LabEx HepSYS (ANR-10-LAB-28). Vincent Turon-Lagot bénéficie d'un contrat doctoral (MRES) du ministère de l'Enseignement supérieur et de la Recherche.

Figures : certains éléments de figure ont été reproduits ou modifiés avec l'autorisation de Servier Medical Art (licence : <https://creativecommons.org/licenses/by/3.0/fr/>).

**Liens d'intérêt :** les auteurs déclarent n'avoir aucun lien d'intérêt en rapport avec l'article.

## Références

- Rizzetto M, Canese MG, Aricò S, Crivelli O, Trepo C, Bonino F, Verme G. Immunofluorescence detection of new antigen-antibody system (delta/anti-delta) associated to hepatitis B virus in liver and in serum of HBsAg carriers. *Gut* 1977 ; 18 : 997-1003.
- Rizzetto M, Hoyer B, Canese MG, Shih JW, Purcell RH, Gerin JL. Delta Agent : association of delta antigen with hepatitis B surface antigen and RNA in serum of delta-infected chimpanzees. *Proc Natl Acad Sci U S A* 1980 ; 77 : 6124-8.
- Sureau C, Guerra B, Lanford RE. Role of the large hepatitis B virus envelope protein in infectivity of the hepatitis delta virion. *J Virol* 1993 ; 67 : 366-72.
- Sureau C, Negro F. The hepatitis delta virus: Replication and pathogenesis. *J Hepatol* 2016 ; 64 : S102-16.
- Trépo C, Chan HLY, Lok A. Hepatitis B virus infection. *Lancet* 2014 ; 384 : 2053-63.
- Chen HY, Shen DT, Ji DZ, Han PC, Zhang WM, *et al.* Prevalence and burden of hepatitis D virus infection in the global population: a systematic review and meta-analysis. *Gut* 2018 ; sous presse. doi:10.1136/gutjnl-2018-316601.
- Amini N, Alavian SM, Kabir A, Aalaei-Andabili SH, Saiedi Hosseini SY, Rizzetto M. Prevalence of hepatitis d in the eastern mediterranean region: systematic review and meta analysis. *Hepat Mon* 2013 ; 13 : e8210.
- Daw MA, Daw AM, Sifennasr NEM, Draha AM, Daw AM, Daw AM, *et al.* The Epidemiology of Hepatitis D Virus in North Africa: A Systematic Review and Meta-Analysis. *Scientific World Journal* 2018 : 9312650.
- Stockdale AJ, Chaponda M, Beloukas A, Phillips RO, Matthews PC, Papadimitropoulos A, *et al.* Prevalence of hepatitis D virus infection in sub-Saharan Africa : a systematic review and meta-analysis. *Lancet Glob Health* 2017 ; 5 : e992-1003.
- Noureddin M, Gish R. Hepatitis delta : epidemiology, diagnosis and management 36 years after discovery. *Curr Gastroenterol Rep* 2014 ; 16 : 365.
- Bonino F, Heermann KH, Rizzetto M, Gerlich WH. Hepatitis delta virus : protein composition of delta antigen and its hepatitis B virus-derived envelope. *J Virol* 1986 ; 58 : 945-50.
- Koh C, Heller T, Glenn JS. Pathogenesis of and New Therapies for Hepatitis D. *Gastroenterology* 2019 ; 156 : 461-76.
- Farci P, Niro GA. Clinical features of hepatitis D. *Semin Liver Dis* 2012 ; 32 : 228-36.
- Tsagris EM, Martínez de Alba AE, Gozmanova M, Kalantidis K. Viroids. *Cell Microbiol* 2008 ; 10 : 2168-79.
- Tabler M, Tsagris M. Viroids : petite RNA pathogens with distinguished talents. *Trends Plant Sci* 2004 ; 9 : 339-48.
- Flores R, Gas ME, Molina-Serrano D, Nohales MÁ D, Carbonell A, Gago S, *et al.* Viroid replication : rolling-circles, enzymes and ribozymes. *Viruses* 2009 ; 1 : 317-34.
- AbouHaidar MG, Venkataraman S, Golshani A, Liu B, Ahmad T. Novel coding, translation, and gene expression of a replicating covalently closed circular RNA of 220 nt. *PNAS* 2014 ; 111 : 14542-7.
- Takeda R, Ding B. Viroid intercellular trafficking: RNA motifs, cellular factors and broad impacts. *Viruses* 2009 ; 1 : 210-21.
- Brazas R, Ganem D. A cellular homolog of hepatitis delta antigen: implications for viral replication and evolution. *Science* 1996 ; 274 : 90-4.
- Salehi-Ashtiani K, Lupták A, Litovchick A, Szostak JW. A genome-wide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. *Science* 2006 ; 313 : 1788-92.
- Nigro JM, Cho KR, Fearon ER, Kern SE, Ruppert JM, Oliner JD, *et al.* Scrambled exons. *Cell* 1991 ; 64 : 607-13.
- Zhang J, Liu H, Hou L, Wang G, Zhang R, Huang Y, *et al.* Circular RNA\_LARP4 inhibits cell proliferation and invasion of gastric cancer by sponging miR-424-5p and regulating LATS1 expression. *Mol Cancer* 2017 ; 16 : 151.
- Fu L, Chen Q, Yao T, Li T, Ying S, Hu Y, Guo J. Hsa\_circ\_0005986 inhibits carcinogenesis by acting as a miR-129-5p sponge and is used as a novel biomarker for hepatocellular carcinoma. *Oncotarget* 2017 ; 8 : 43878-88.
- Han D, Li J, Wang H, Su X, Hou J, Gu Y, *et al.* Circular RNA circMTO1 acts as the sponge of microRNA-9 to suppress hepatocellular carcinoma progression. *Hepatology* 2017 ; 66 : 1151-64.
- Wille M, Netter HJ, Littlejohn M, Yuen L, Shi M, Eden JS, *et al.* A Divergent Hepatitis D-Like Agent in Birds. *Viruses* 2018 ; 10 : E720.

26. Schulze A, Gripon P, Urban S. Hepatitis B virus infection initiates with a large surface protein-dependent binding to heparan sulfate proteoglycans. *Hepatology* 2007 ; 46 : 1759-68.
27. Verrier ER, Colpitts CC, Bach C, Heydmann L, Weiss A, Renaud M, *et al.* A targeted functional RNA interference screen uncovers glypican 5 as an entry factor for hepatitis B and D viruses. *Hepatology* 2016 ; 63 : 35-48.
28. Li W, Urban S. Entry of hepatitis B and hepatitis D virus into hepatocytes : Basic insights and clinical implications. *J Hepatol* 2016 ; 64 : S32-40.
29. Yan H, Zhong G, Xu G, He W, Jing Z, Gao Z, *et al.* Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. *Elife* 2012 ; 1 : e00049.
30. Ni Y, Lempp FA, Mehrle S, Nkongolo S, Kaufman C, Fälth M, *et al.* Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes. *Gastroenterology* 2014 ; 146 : 1070-83.
31. Alves C, Freitas N, Cunha C. Characterization of the nuclear localization signal of the hepatitis delta virus antigen. *Virology* 2008 ; 370 : 12-21.
32. Chou HC, Hsieh TY, Sheu GT, Lai MM. Hepatitis delta antigen mediates the nuclear import of hepatitis delta virus RNA. *J Virol* 1998 ; 72 : 3684-90.
33. Modahl LE, Lai MM. Transcription of hepatitis delta antigen mRNA continues throughout hepatitis delta virus (HDV) replication: a new model of HDV RNA transcription and replication. *J Virol* 1998 ; 72 : 5449-56.
34. Macnaughton TB, Shi ST, Modahl LE, Lai MMC. Rolling circle replication of hepatitis delta virus RNA is carried out by two different cellular RNA polymerases. *J Virol* 2002 ; 76 : 3920-7.
35. Huang WH, Chen YS, Chen PJ. Nucleolar targeting of hepatitis delta antigen abolishes its ability to initiate viral antigenomic RNA replication. *J Virol* 2008 ; 82 : 692-9.
36. Lee CH, Chang SC, Wu CH, Chang MF. A novel chromosome region maintenance 1-independent nuclear export signal of the large form of hepatitis delta antigen that is required for the viral assembly. *J Biol Chem* 2001 ; 276 : 8142-8.
37. Glenn JS, Watson JA, Havel CM, White JM. Identification of a prenylation site in delta virus large antigen. *Science* 1992 ; 256 : 1331-3.
38. Otto JC, Casey PJ. The hepatitis delta virus large antigen is farnesylated both in vitro and in animal cells. *J Biol Chem* 1996 ; 271 : 4569-72.
39. Hwang SB, Lai MM. Isoprenylation masks a conformational epitope and enhances trans-dominant inhibitory function of the large hepatitis delta antigen. *J Virol* 1994 ; 68 : 2958-64.
40. Huang HC, Lee CP, Liu HK, Chang MF, Lai YH, Lee YC, Huang C. Cellular Nuclear Export Factors TAP and Aly Are Required for HDV-L-mediated Assembly of Hepatitis Delta Virus. *J Biol Chem* 2016 ; 291 : 26226-38.
41. Hwang SB, Lai MM. Isoprenylation mediates direct protein-protein interactions between hepatitis large delta antigen and hepatitis B virus surface antigen. *J Virol* 1993 ; 67 : 7659-62.
42. Freitas N, Cunha C, Menne S, Gudima SO. Envelope proteins derived from naturally integrated hepatitis B virus DNA support assembly and release of infectious hepatitis delta virus particles. *J Virol* 2014 ; 88 : 5742-54.
43. Verrier ER, Weiss A, Bach C, Heydmann L, Turon-Lagot V, Kopp A, *et al.* Combined small molecule and loss-of-function screen uncovers estrogen receptor alpha and CAD as host factors for HDV infection and antiviral targets. *Gut* 2019 ; sous presse. doi:10.1136/gutjnl-2018-317065.
44. Cao D, Haussecker D, Huang Y, Kay MA. Combined proteomic-RNAi screen for host factors involved in human hepatitis delta virus replication. *RNA* 2009 ; 15 : 1971-9.
45. Mota S, Mendes M, Penque D, Coelho AV, Cunha C. Changes in the proteome of Huh7 cells induced by transient expression of hepatitis D virus RNA and antigens. *J Proteomics* 2008 ; 71 : 71-9.
46. Mota S, Mendes M, Freitas N, Penque D, Coelho AV, Cunha C. Proteome analysis of a human liver carcinoma cell line stably expressing hepatitis delta virus ribonucleoproteins. *J Proteomics* 2009 ; 72 : 616-27.
47. Sikora D, Greco-Stewart VS, Miron P, Pelchat M. The hepatitis delta virus RNA genome interacts with eEF1A1, p54(nrb), hnRNP-L, GAPDH and ASF/SF2. *Virology* 2009 ; 390 : 71-8.
48. Zhang Z, Filzmayer C, Ni Y, Sülthmann H, Mutz P, Hiet MS, *et al.* Hepatitis D virus replication is sensed by MDA5 and induces IFN- $\beta$ / $\lambda$  responses in hepatocytes. *J Hepatol* 2018 ; 69 : 25-35.
49. Pugnale P, Pazienza V, Guilloux K, Negro F. Hepatitis delta virus inhibits alpha interferon signaling. *Hepatology* 2009 ; 49 : 398-406.
50. Williams V, Brichler S, Khan E, Chami M, Dény P, Kremsdorf D, Gordien E. Large hepatitis delta antigen activates STAT-3 and NF- $\kappa$ B via oxidative stress. *J Viral Hepat* 2012 ; 19 : 744-53.
51. Aslan N, Yurdaydin C, Wiegand J, Greten T, Ciner A, Meyer MF, *et al.* Cytotoxic CD4+ T cells in viral hepatitis. *J Viral Hepat* 2006 ; 13 : 505-14.
52. Karimzadeh H, Kiraithe MM, Oberhardt V, Alizei ES, Bockmann J, Zur Wiesch JS, *et al.* Mutations in Hepatitis D Virus Allow it to Escape Detection by CD8+ T Cells and Evolve at the Population Level. *Gastroenterology* 2019. Sous presse. doi:10.1053/j.gastro.2019.02.003.
53. Negro F. Hepatitis D virus coinfection and superinfection. *Cold Spring Harb Perspect Med* 2014 ; 4 : a021550.
54. Smedile A, Verme G, Cargnel A, Dentico P, Opolon P, Vergani D, *et al.* Influence of Delta Infection on Severity of Hepatitis B. *The Lancet* 1982 ; 320 : 945-7.
55. Govindarajan S, Chin KP, Redeker AG, Peters RL. Fulminant B viral hepatitis: role of delta agent. *Gastroenterology* 1984 ; 86 : 1417-20.
56. Fattovich G, Boscaro S, Noventa F, Pomaro E, Stenico D, Alberti A, *et al.* Influence of hepatitis delta virus infection on progression to cirrhosis in chronic hepatitis type B. *J Infect Dis* 1987 ; 155 : 931-5.
57. Fattovich G, Giustina G, Christensen E, Pantalena M, Zagni I, Realdi G, Schalm SW. Influence of hepatitis delta virus infection on morbidity and mortality in compensated cirrhosis type B. The European Concerted Action on Viral Hepatitis (Eurohep). *Gut* 2000 ; 46 : 420-6.
58. Shattock AG, Morgan BM. Sensitive enzyme immunoassay for the detection of delta antigen and anti-delta, using serum as the delta antigen source. *J Med Virol* 1984 ; 13 : 73-82.
59. Shattock AG, Morris MC. Evaluation of commercial enzyme immunoassays for detection of hepatitis delta antigen and anti-hepatitis delta virus (HDV) and immunoglobulin M anti-HDV antibodies. *J Clin Microbiol* 1991 ; 29 : 1873-6.
60. Aragona M, Macagno S, Caredda F, Crivelli O, Lavarini C, Maran E, *et al.* Serological response to the hepatitis delta virus in hepatitis D. *Lancet* 1987 ; 1 : 478-80.
61. Chen X, Oidovsambuu O, Liu P, Grosely R, Elazar M, Winn VD, *et al.* A novel quantitative microarray antibody capture assay identifies an extremely high hepatitis delta virus prevalence among hepatitis B virus-infected mongolians. *Hepatology* 2017 ; 66 : 1739-49.
62. Mahale P, Aka PV, Chen X, Liu P, Fram BJ, Wang AS, *et al.* Hepatitis D Viremia Among Injection Drug Users in San Francisco. *J Infect Dis* 2018 ; 217 : 1902-6.
63. Brichler S, Le Gal F, Butt A, Chevret S, Gordien E. Commercial real-time reverse transcriptase PCR assays can underestimate or fail to quantify hepatitis delta virus viremia. *Clin Gastroenterol Hepatol* 2013 ; 11 : 734-40.
64. Brichler S, Le Gal F, Neri-Pinto F, Mansour W, Roulot D, Laperche S, Gordien E. Serological and molecular diagnosis of hepatitis delta virus infection: results of a French national quality control study. *J Clin Microbiol* 2014 ; 52 : 1694-7.
65. Le Gal F, Dziri S, Gerber A, Alloui C, Ben Abdesselam Z, Roulot D, *et al.* Performance Characteristics of a New Consensus Commercial Kit for Hepatitis D Virus RNA Viral Load Quantification. *J Clin Microbiol* 2017 ; 55 : 431-41.



66. Lempp FA, Ni Y, Urban S. Hepatitis delta virus : insights into a peculiar pathogen and novel treatment options. *Nat Rev Gastroenterol Hepatol* 2016 ; 13 : 580-9.
67. Wedemeyer H, Yurdaydin C, Dalekos GN, Erhardt A, Çakaloğlu Y, Değertekin H, *et al.* Peginterferon plus Adefovir versus Either Drug Alone for Hepatitis Delta. *N Engl J Med* 2011 ; 364 : 322-31.
68. Yurdaydin C. New treatment option for delta virus: is a cure in sight ? *J Viral Hepat* 2019 ; sous presse. doi:10.1111/jvh.13081.
69. Buchmann B, Döhner K, Schirdewahn T, Sodeik B, Manns MP, Wedemeyer H, *et al.* A screening assay for the identification of host cell requirements and antiviral targets for hepatitis D virus infection. *Antiviral Res* 2017 ; 141 : 116-23.
70. Heidrich B, Yurdaydin C, Kabaçam G, Ratsch BA, Zachou K, Bremer B, *et al.* Late HDV RNA relapse after peginterferon alpha-based therapy of chronic hepatitis delta. *Hepatology* 2014 ; 60 : 87-97.
71. Baumert TF, Verrier ER, Nassal M, Chung RT, Zeisel MB. Host-targeting agents for treatment of hepatitis B virus infection. *Curr Opin Virol* 2015 ; 14 : 41-6.
72. Colpitts CC, Verrier ER, Baumert TF. Targeting Viral Entry for Treatment of Hepatitis B and C Virus Infections. *ACS Infect Dis* 2015 ; 1 : 420-7.
73. Crouchet E, Wrensch F, Schuster C, Zeisel MB, Baumert TF. Host-targeting therapies for hepatitis C virus infection: current developments and future applications. *Therap Adv Gastroenterol* 2018 ; 11 : 1-15.
74. Verrier ER, Heydmann L, Baumert TF, Schuster C. Le transporteur d'acides biliaires NTCP, un acteur majeur dans l'infection par les virus humains des hépatites offrant de nouvelles perspectives thérapeutiques. *Virologie* 2018 ; 22 : 55-66.
75. Verrier ER, Colpitts CC, Sureau C, Baumert TF. Hepatitis B virus receptors and molecular drug targets. *Hepatol Int* 2016 ; 10 : 567-73.
76. Petersen J, Dandri M, Mier W, Lütgehetmann M, Volz T, Weizsäcker F, von F, *et al.* Prevention of hepatitis B virus infection *in vivo* by entry inhibitors derived from the large envelope protein. *Nature Biotechnol* 2008 ; 26 : 335-41.
77. Barrera A, Guerra B, Notvall L, Lanford RE. Mapping of the hepatitis B virus pre-S1 domain involved in receptor recognition. *J Virol* 2005 ; 79 : 9786-98.
78. Gripon P, Cannie I, Urban S. Efficient Inhibition of Hepatitis B Virus Infection by Acylated Peptides Derived from the Large Viral Surface Protein. *J Virol* 2005 ; 79 : 1613-22.
79. Lütgehetmann M, Mancke LV, Volz T, Helbig M, Allweiss L, Bornscheuer T, *et al.* Humanized chimeric uPA mouse model for the study of hepatitis B and D virus interactions and preclinical drug evaluation. *Hepatology* 2012 ; 55 : 685-94.
80. Bogomolov P, Alexandrov A, Voronkova N, Macievich M, Kokina K, Petrachenkova M, *et al.* Treatment of chronic hepatitis D with the entry inhibitor myrcludex B: First results of a phase Ib/IIa study. *J Hepatol* 2016 ; 65 : 490-8.
81. Farci P, Niro GA. Current and Future Management of Chronic Hepatitis D. *Gastroenterol Hepatol (N Y)* 2018 ; 14 : 342-51.
82. Wedemeyer H, Bogomolov P, Blank A, Allweiss L, Dandri-Petersen M, Bremer B, *et al.* Final results of a multicenter, open-label phase 2b clinical trial to assess safety and efficacy of Myrcludex B in combination with Tenofovir in patients with chronic HBV/HDV co-infection. *J Hepatol* 2018 ; 68 : S3.
83. Lou L. Advances in Nucleotide Antiviral Development from Scientific Discovery to Clinical Applications: Tenofovir Disoproxil Fumarate for Hepatitis B. *J Clin Transl Hepatol* 2013 ; 1 : 33-8.
84. Wedemeyer H, Schöneweis K, Bogomolov P, Voronkova N, Chulanov V, Stepanova T, *et al.* Interim Results of a Multicentre. Open-Label Phase 2 Clinical Trial (MYR203) to Assess Safety and Efficacy of Myrcludex B in Combination with Peg-Interferon Alpha 2a in Patients with Chronic HBV/HDV Co-Infection. *Hepatology* 2018 ; 68 : S11A.
85. Bordier BB, Marion PL, Ohashi K, Kay MA, Greenberg HB, Casey JL, Glenn JS. A prenylation inhibitor prevents production of infectious hepatitis delta virus particles. *J Virol* 2002 ; 76 : 10465-72.
86. Bordier BB, Ohkanda J, Liu P, Lee SY, Salazar FH, Marion PL, *et al.* In vivo antiviral efficacy of prenylation inhibitors against hepatitis delta virus. *J Clin Invest* 2003 ; 112 : 407-14.
87. Berndt N, Hamilton AD, Sebt SM. Targeting protein prenylation for cancer therapy. *Nat Rev Cancer* 2011 ; 11 : 775-91.
88. Koh C, Canini L, Dahari H, Zhao X, Upprichard SL, Haynes-Williams V, *et al.* Oral prenylation inhibition with lonafarnib in chronic hepatitis D infection: a proof-of-concept randomised, double-blind, placebo-controlled phase 2A trial. *Lancet Infect Dis* 2015 ; 15 : 1167-74.
89. Yurdaydin C, Keskin O, Kalkan Ç, Karakaya F, Çalışkan A, Karataylı E, *et al.* Optimizing lonafarnib treatment for the management of chronic delta hepatitis: The LOWR HDV-1 study. *Hepatology* 2018 ; 67 : 1224-36.
90. Wedemeyer H, Port K, Deterding K, Wranke A, Kirschner J, Bruno B, *et al.* A phase 2 dose-escalation study of lonafarnib plus ritonavir in patients with chronic hepatitis D : final results from the lonafarnib with ritonavir in HDV-4 (LOWR HDV-4) study. *J Hepatol* 2017 ; 66 : S24.
91. Koh C, Surana P, Han T, Fryzek N, Kapuria D, Etzion O, *et al.* A phase 2 study exploring once daily dosing of ritonavir boosted lonafarnib for the treatment of chronic delta hepatitis – end of study results from the LOWR HDV-3 study. *J Hepatol* 2017 ; 66 : S101-2.
92. Matsumura T, Hu Z, Kato T, Dreux M, Zhang YY, Imamura M, *et al.* Amphipathic DNA polymers inhibit hepatitis C virus infection by blocking viral entry. *Gastroenterology* 2009 ; 137 : 673-81.
93. Bernstein DI, Goyette N, Cardin R, Kern ER, Boivin G, Ireland J, *et al.* Amphipathic DNA polymers exhibit antiherpetic activity in vitro and in vivo. *Antimicrob Agents Chemother* 2008 ; 52 : 2727-33.
94. Beilstein F, Blanchet M, Vaillant A, Sureau C. Nucleic Acid Polymers Are Active against Hepatitis Delta Virus Infection In Vitro. *J Virol* 2018 ; 92 : e01416-1417.
95. Vaillant A. Nucleic acid polymers : Broad spectrum antiviral activity, antiviral mechanisms and optimization for the treatment of hepatitis B and hepatitis D infection. *Antiviral Res* 2016 ; 133 : 32-40.
96. Noordeen F, Scougall CA, Grosse A, Qiao Q, Ajilian BB, Reaiche-Miller G, *et al.* Therapeutic Antiviral Effect of the Nucleic Acid Polymer REP 2055 against Persistent Duck Hepatitis B Virus Infection. *PLoS ONE* 2015 ; 10 : e0140909.
97. Al-Mahtab M, Bazinet M, Vaillant A. Safety and Efficacy of Nucleic Acid Polymers in Monotherapy and Combined with Immunotherapy in Treatment-Naïve Bangladeshi Patients with HBsAg+ Chronic Hepatitis B Infection. *PLoS ONE* 2016 ; 11 : e0156667.
98. Bazinet M, Pântea V, Cebotarescu V, Cojuhari L, Jimbei P, Albrecht J, *et al.* Safety and efficacy of REP 2139 and pegylated interferon alfa-2a for treatment-naïve patients with chronic hepatitis B virus and hepatitis D virus co-infection (REP 301 and REP 301-LTF): a non-randomised, open-label, phase 2 trial. *Lancet Gastroenterol Hepatol* 2017 ; 2 : 877-89.
99. Vaillant A. REP 2139: Antiviral Mechanisms and Applications in Achieving Functional Control of HBV and HDV Infection. *ACS Infect Dis* 2018. Sous presse. doi:10.1021/acsinfecdis.8b00156.



---

7.5 **Saviano A**, Baumert TF. Mortality from liver cirrhosis and HCC in the DAA era: success in viral control is darkened by raise of metabolic disease. *Hepatobiliary Surgery and Nutrition*, 2019 Jun;8(3):307-310.

# Mortality from liver cirrhosis and HCC in the DAA era: success in viral control is darkened by raise of metabolic disease

Antonio Saviano<sup>1,2</sup>, Thomas F. Baumert<sup>1,2</sup>

<sup>1</sup>Inserm U1110, Institut de Recherche sur les Maladies Virales et Hépatiques, Université de Strasbourg, F-67000 Strasbourg, France; <sup>2</sup>Pôle Hépatodigestif, Institut Hospitalo-Universitaire, Hôpitaux Universitaires, F-67000 Strasbourg, France

Correspondence to: Prof. Thomas F. Baumert, MD. Inserm U1110, Université de Strasbourg, 3 Rue Koeberlé, F-67000 Strasbourg, France.

Email: thomas.baumert@unistra.fr.

Comment on: Kim D, Li AA, Perumpail BJ, *et al.* Changing Trends in Etiology- and Ethnicity-Based Annual Mortality Rates of Cirrhosis and Hepatocellular Carcinoma in the United States. *Hepatology* 2019;69:1064-74.

Submitted Jan 14, 2019. Accepted for publication Jan 26, 2019.

doi: 10.21037/hbsn.2019.01.21

View this article at: <http://dx.doi.org/10.21037/hbsn.2019.01.21>

Liver cirrhosis and primary liver cancers are leading causes of mortality worldwide with over than 1.7 million of deaths in 2010 representing about 3.4% of the overall deaths (1). Global liver cirrhosis deaths ceaselessly increased by 52.2% from 1980 to 2010 (1). Similarly, primary liver cancers global deaths increased by 62.4% from 1990 to 2010 (2). Chronic viral (i.e., HBV and HCV), alcoholic (ALD) and non-alcoholic fatty liver (NAFLD) diseases are the most important etiologies of liver cirrhosis and hepatocellular carcinoma (HCC). In the last years, the introduction of direct antiviral agents (DAAs) has revolutionized HCV care. Indeed, DAAs are highly effective (cure rates higher than 90%) and well tolerated even by difficult-to-treat candidates as patients with advanced liver cirrhosis. However, the global impact of these progresses on advanced liver disease and HCC remains to be determined. While viral cure has been shown to decrease the overall HCC risk in HCV-infected patients, accumulating clinical evidences in large cohort studies demonstrate that HCC risk persists after HCV cure especially in advanced fibrosis (3,4) with an annual HCC incidence between 1–12% per year. Furthermore, an unexpected high rate of early HCC recurrence following DAA treatment in some studies but not in others has raised concerns on the effect of these drugs in HCC prevention (5). Finally, a large majority of HCV-infected patients has no access to DAAs due to high costs.

Highly effective and tolerate antiviral agents for treatment of chronic hepatitis B were also introduced in

clinical practice in the last 15 years. Indeed, entecavir was FDA-approved in 2005 and tenofovir disoproxil fumarate in 2008. While antiviral drugs have been shown to prevent or reverse hepatic decompensation and to reduce the prevalence of virus-induced end-stage liver disease (6,7), data on their effect on mortality on a population level are largely lacking.

At the same time, the incidence of NAFLD-related chronic liver failure and HCC has increased dramatically and ALD has been reported as the most common cause of chronic liver disease (7,8). Data from United Network for Organ Sharing (UNOS) cohort revealed a significant decrease in the prevalence of HCV and an increase of NAFLD or ALD among patients new to the liver transplant waitlist or undergoing liver transplantation for liver cirrhosis (7). However, among patients transplanted for HCC, the proportions of HCV infection, NAFLD and ALD did not change between 2003 and 2015 (7).

To better understand the epidemiological changes in the etiologies of advanced liver disease and evaluate the impact of the novel antiviral treatments on a population level, updated incidence and mortality data of the last decade are essentially needed.

In the article by Kim *et al.* recently published in *Hepatology* (9), the authors performed an elegant analysis of the mortality trends for liver cirrhosis and HCC in the United States over the last 10 years in adults aged  $\geq 20$  years. Using mortality records from the Centers for Disease Control and Prevention's National Vital Statics System

that uses the International Classification of Disease, Tenth Revision (ICD-10) to codify diseases, they calculated age-specific mortality rates and used joinpoint regression to determine the annual percentage change (APC) of mortality for both cirrhosis and HCC. APCs of mortality rates according to the major etiologies, ethnicities and sex were also assessed.

The authors found that in U.S. the age-standardized cirrhosis-related mortality rates increased from 19.77/100,000 persons in 2007 to 23.67 in 2016 with an average annual increase of 2.3%, and similarly, HCC-related mortality increased from 3.48 persons in 2007 to 4.41 in 2016 at an annual rate of 2.0% (9).

Interestingly, the APC of HCV-related cirrhosis mortality shifted from an increase of 2.9% per year during 2007–2014 to a reduction of 6.5% per year during 2014–2016. In parallel, mortality for ALD-related and NAFLD-related cirrhosis increased over the study period with an APC of, respectively, 4.5% and 15.4% per year. Inversely, mortality for HBV-related cirrhosis decreased with an average APC of –1.1% (9).

Regarding HCC, age-standardized mortality rates in U.S. for HCV-related HCC increased (APC 7.0%) and reached a plateau in 2012 remaining stable from 2012 to 2016 (APC 0.4%). A linear increase in the age-standardized HCC-related mortality rates for ALD and NAFLD was observed from 2007 to 2016 (APC respectively 7.4% and 19.1%). Concerning HBV, a trend in increased HCC mortality (APC +6.3%) was observed between 2007 and 2010 while an inverse trend (–2.2%) was noted in the period 2010–2016 (9).

Here, Kim *et al.* showed that since the introduction of DAAs in U.S. in late 2013, there has been a significant decrease in HCV-related cirrhosis mortality rates compared with the pre-DAA era. Importantly—since an alert on the association between DAAs and early HCC recurrence was raised in 2016 (10,11)—no increase in HCV-related HCC mortality after the DAA introduction was observed. It is of interest to note that HCC is a late event in the natural history of chronic liver diseases. Long-term follow-up studies have observed that 1–8% of patients with cirrhosis develop HCC per year (3,12). This implies that long-term studies (>10 years) are needed to evaluate any benefit of antiviral treatments on HCC mortality. Kim *et al.* showed that HCV-related HCC mortality did not raise after 2012 and that HBV-related HCC mortality started to steadily decrease after 2010.

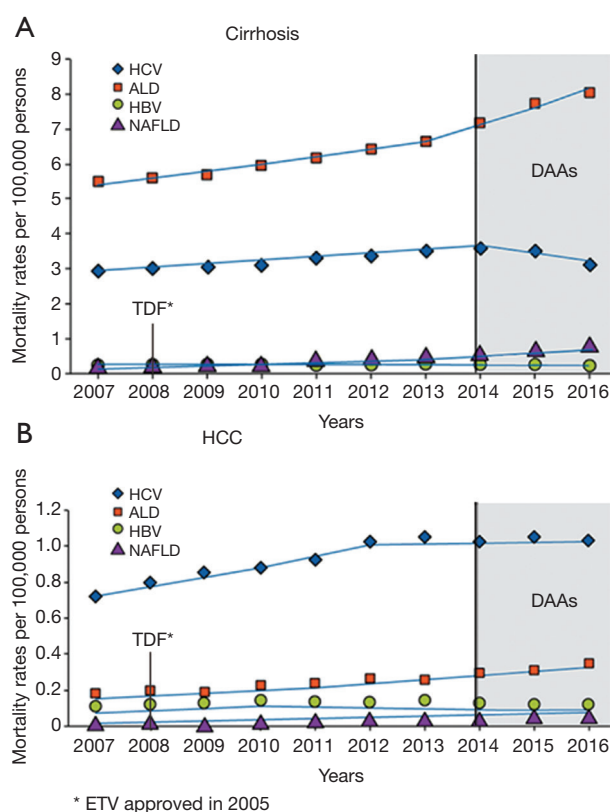
Other relevant findings of this study are the data on ALD and NAFLD. The mortality rates for ALD and

NAFLD dramatically increased during the 10-year study period. In 2016, among patients with cirrhosis, ALD had the highest age-standardized mortality rate, more than the double of HCV (8.23/100,000 persons *vs.* 3.20) with an APC of +4.5%. At the same time, NAFLD is the third cause of cirrhosis mortality (0.82/100,000 persons in 2016) and the fourth cause of HCC mortality (0.06/100,000 persons in 2016) with impressive APCs of respectively +15.4% and +19.1%. These trends could be partly explained by the increased awareness and improvement in diagnosis in the last 10 years even though the ICD coding system has shown to underestimate NAFLD prevalence and mortality. However, it is important to note that in spite of the ALD and NAFLD trends, HCV still accounted for most of HCC deaths during the study period confirming the different cancer risk between viral and metabolic diseases (12).

Several U.S. studies also demonstrated that, in the last 15 years, inpatient mortality from liver cirrhosis significantly declined over time (13,14). Then, the trends in cirrhosis mortality presented by Kim *et al.* confirms that some burden of mortality cirrhosis shifted from the in-hospital to the outpatient setting as already suggested by previous Veterans Administration hospitals' data (14).

Finally, a subgroup analysis of mortality by ethnical groups was conducted. In 2016, non-Hispanic whites and Hispanic had the highest age-standardize cirrhosis mortality while non-Hispanic blacks and non-Hispanic Asian had the highest HCC mortality. In terms of APC, non-Hispanic whites had the highest APC in cirrhosis and HCC-related mortality (respectively +3.5% and +2.4% per year) while non-Hispanic Asians were the only ethnic group showing a significant APC reduction in HCC mortality with a decline of 3.5% per year. The interpretation of these data is difficult because no adjustment for etiology's prevalence, income and/or access to care was performed. A recent study showed that a significant variability exists in U.S. in liver disease-related mortality among states and is independent from the prevalence of alcohol consumption and obesity while strongly correlates with high prevalence of Hispanic individuals, viral hepatitis and low income (15).

In conclusion, the work by Kim *et al.* provides the liver community with relevant data about the changes in cirrhosis and HCC mortality in U.S. over the last 10 years (Figure 1). In the DAA era, HCV-cirrhosis mortality significantly decreased while HCV-HCC mortality did not increase after 2012. After the introduction of high-barrier nucleoside and nucleotide analogues for HBV, HBV-cirrhosis mortality has constantly decreased. Studies with longer follow-up are



**Figure 1** Age-standardized mortality rates for cirrhosis (A) and HCC (B) in United States between 2007 and 2016. HCV-related cirrhosis mortality significantly decreased after the introduction of DAAs while HCV-related HCC mortality did not increase. After the introduction of entecavir and tenofovir, mortality of HBV-cirrhosis significantly decreased and HBV-HCC mortality showed a decline trend. ALD and NAFLD mortality rates dramatically raised over the study period. Adapted from Kim *et al.* (9). ALD, alcoholic liver disease; DAA, direct antiviral agent; ETV, entecavir; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; NAFLD, non-alcoholic liver disease, TDF, tenofovir.

needed to capture changes in the mortality of HBV and HCV-related HCC. Concurrently, mortality rates for ALD and NAFLD associated cirrhosis and HCC dramatically increased in the last years. Public policies and treatment strategies are urgently required to further reduce liver disease mortality with a special focus on metabolic liver disease.

## Acknowledgments

**Funding:** This work was supported by Inserm, the University

of Strasbourg, the European Union (Infect-ERA hepBccc, ERC-2014-AdG-671231-HEPCIR and Horizon 2020 research and innovation programme under grant agreement 667273 - HEPCAR), the Agence Nationale de Recherches sur le Sida et les Hépatites Virales (ANRS 15/1099), the French Cancer Agency (ARC IHU201301187) and the US National Institutes of Health (NIH/NIAID U19 AI123862-01, NIH/NIAID R03 AI131066, NIH/NCI R21 CA209940). This work has been published under the framework of the LabEx ANR-10-LAB-28 and benefits from a funding from the state managed by the French National Research Agency as part of the Investments for the Future (Investissements d'Avenir) program. A.S. is the recipient of a fellowship co-funded by the Région Alsace, France, the LabEx HepSys and IHU Strasbourg.

## Footnote

**Conflicts of Interest:** The authors have no conflicts of interest to declare.

## References

1. Mokdad AA, Lopez AD, Shahraz S, et al. Liver cirrhosis mortality in 187 countries between 1980 and 2010: a systematic analysis. *BMC Med* 2014;12:145.
2. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380:2095-128.
3. Kanwal F, Kramer J, Asch SM, et al. Risk of Hepatocellular Cancer in HCV Patients Treated With Direct-Acting Antiviral Agents. *Gastroenterology* 2017;153:996-1005.e1.
4. van der Meer AJ, Feld JJ, Hofer H, et al. Risk of cirrhosis-related complications in patients with advanced fibrosis following hepatitis C virus eradication. *J Hepatol* 2017;66:485-93.
5. Llovet JM, Villanueva A. Liver cancer: Effect of HCV clearance with direct-acting antiviral agents on HCC. *Nat Rev Gastroenterol Hepatol* 2016;13:561-2.
6. Lim YS, Han S, Heo NY, et al. Mortality, liver transplantation, and hepatocellular carcinoma among patients with chronic hepatitis B treated with entecavir vs lamivudine. *Gastroenterology* 2014;147:152-61.
7. Goldberg D, Ditah IC, Saeian K, et al. Changes in the Prevalence of Hepatitis C Virus Infection, Nonalcoholic

- Steatohepatitis, and Alcoholic Liver Disease Among Patients With Cirrhosis or Liver Failure on the Waitlist for Liver Transplantation. *Gastroenterology* 2017;152:1090-9.e1.
8. Beste LA, Leipertz SL, Green PK, et al. Trends in burden of cirrhosis and hepatocellular carcinoma by underlying liver disease in US veterans, 2001-2013. *Gastroenterology* 2015;149:1471-82.e5; quiz e17-8.
  9. Kim D, Li AA, Perumpail BJ, et al. Changing Trends in Etiology- and Ethnicity-Based Annual Mortality Rates of Cirrhosis and Hepatocellular Carcinoma in the United States. *Hepatology* 2019;69:1064-74.
  10. Conti F, Buonfiglioli F, Scuteri A, et al. Early occurrence and recurrence of hepatocellular carcinoma in HCV-related cirrhosis treated with direct-acting antivirals. *J Hepatol* 2016;65:727-33.
  11. Reig M, Mariño Z, Perelló C, et al. Unexpected high rate of early tumor recurrence in patients with HCV-related HCC undergoing interferon-free therapy. *J Hepatol* 2016;65:719-26.
  12. Ioannou GN, Green P, Lowy E, et al. Differences in hepatocellular carcinoma risk, predictors and trends over time according to etiology of cirrhosis. *PLoS One* 2018;13:e0204412.
  13. Schmidt ML, Barritt AS, Orman ES, et al. Decreasing mortality among patients hospitalized with cirrhosis in the United States from 2002 through 2010. *Gastroenterology* 2015;148:967-77.e2.
  14. Kanwal F, Tansel A, Kramer JR, et al. Trends in 30-Day and 1-Year Mortality Among Patients Hospitalized With Cirrhosis From 2004 to 2013. *Am J Gastroenterol* 2017;112:1287-97.
  15. Desai AP, Mohan P, Roubal AM, et al. Geographic Variability in Liver Disease-Related Mortality Rates in the United States. *Am J Med* 2018;131:728-34.

**Cite this article as:** Saviano A, Baumert TF. Mortality from liver cirrhosis and HCC in the DAA era: success in viral control is darkened by raise of metabolic disease. *HepatoBiliary Surg Nutr* 2019;8(3):307-310. doi: 10.21037/hbsn.2019.01.21

---

7.6 **Saviano A, Roehlen N, Virzì A, Roca Suarez AA, Hoshida Y, Lupberger J, Baumert TF.** Stromal and Immune Drivers of Hepatocarcinogenesis. Hepatocellular carcinoma. Springer 2019.

# Chapter 15

## Stromal and Immune Drivers of Hepatocarcinogenesis



Antonio Saviano, Natascha Roehlen, Alessia Virzi,  
Armando Andres Roca Suarez, Yujin Hoshida, Joachim Lupberger,  
and Thomas F. Baumert

### Introduction

The liver is a multifunctional organ that plays a key role in metabolism and detoxification as well as in regulation of immune response and tolerance. The liver is physiologically exposed to many pathogens and toxic substances derived from the gut and has the largest population of resident macrophages (i.e., Kupffer cells, KCs) in the body and a high prevalence of natural killer cells (NK), natural killer T cells (NKT), and T cells. In normal conditions, the liver removes a large amount of microbes and pathogen-associated and damage-associated molecular patterns (PAMPs and DAMPs) and maintains an immunosuppressive environment [1].

Following chronic hepatocyte damage, immune and stromal cells modify a liver environment, which triggers chronic inflammation and ultimately promotes hepatocellular carcinoma (HCC) [2]. Indeed, independently from the etiology, chronic liver disease is characterized by a deregulation in the liver immune network

---

\* Antonio Saviano and Natascha Roehlen are co-first authors of this chapter

A. Saviano · T. F. Baumert (✉)  
Inserm U1110, Institut de Recherche sur les Maladies Virales et Hépatiques, Université de  
Strasbourg, Strasbourg, France

Pôle Hépato-digestif, Institut Hopitalo-Universitaire, Hôpitaux Universitaires,  
Strasbourg, France  
e-mail: [thomas.baumert@unistra.fr](mailto:thomas.baumert@unistra.fr)

N. Roehlen · A. Virzi · A. A. R. Suarez · J. Lupberger  
Inserm U1110, Institut de Recherche sur les Maladies Virales et Hépatiques, Université de  
Strasbourg, Strasbourg, France

Y. Hoshida  
Liver Tumor Translational Research Program, Simmons Comprehensive Cancer Center,  
Division of Digestive and Liver Diseases, Department of Internal Medicine,  
University of Texas Southwestern Medical Center, Dallas, TX, USA

© Springer Nature Switzerland AG 2019  
Y. Hoshida (ed.), *Hepatocellular Carcinoma*, Molecular and Translational  
Medicine, [https://doi.org/10.1007/978-3-030-21540-8\\_15](https://doi.org/10.1007/978-3-030-21540-8_15)



that stimulates cellular stress and death favoring liver fibrosis, hepatocyte proliferation, and epithelial-to-mesenchymal transition (EMT) [2]. A combination of EMT, genetic mutations, and epigenetic alterations that accumulate during cell proliferation is the most important driver of hepatocarcinogenesis [3].

Once HCC has developed, liver microenvironment greatly affects tumor progression and response to therapy [4]. This is the reason why gene expression signatures in liver tissues adjacent to the HCC—and the not in tumor itself—highly correlate with long-term survival of patients with liver fibrosis [5]. Similarly, HCC infiltration by non-parenchymal cells (e.g., regulatory T cells,  $T_{reg}$ ) has been associated with tumor progression [5–8]. New therapies targeting liver microenvironment are recently developed or under clinical investigation for both chronic liver disease (e.g., nonalcoholic steatohepatitis, NASH) and HCC.

Hence, liver microenvironment plays an essential role in both hepatocarcinogenesis and tumor progression and it is an important therapeutic target for HCC prevention and treatment.

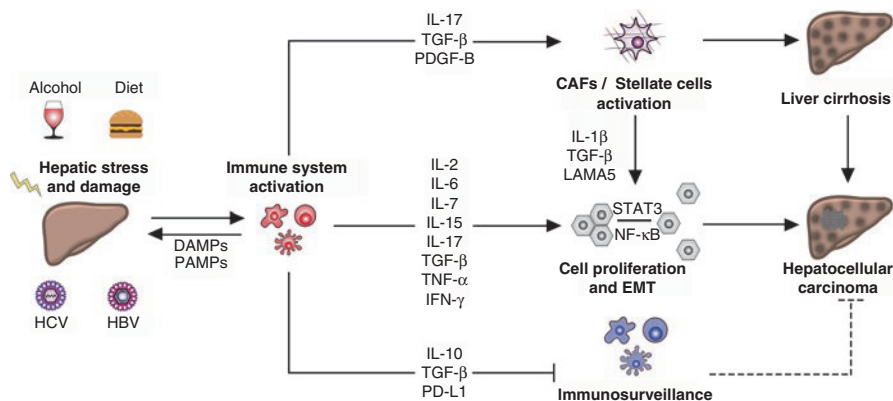
## From Chronic Inflammation to Hepatocellular Carcinoma

HCC almost universally evolves on the background of chronic liver inflammation and liver fibrosis [9]. Chronic hepatocyte cell injury induces activation of the immune system that initiates and supports chronic inflammation by generation of proinflammatory cytokines and chemokines and activation of hepatic stellate cells (HSCs), finally resulting in liver fibrosis, cirrhosis, and cancer [10] (Fig. 15.1).

During chronic infections (e.g., hepatitis B virus, HBV, or hepatitis C virus, HCV) as well as metabolic (e.g., NASH) or toxic diseases (e.g., alcoholic steatohepatitis, ASH), immune cells—first of all KCs—are activated by the release of PAMPs and DAMPs produced by hepatocyte apoptosis and death. Activated KCs present viral antigens to T cells and/or secrete cytokines and chemokines that recruit circulating monocytes, lymphocytes, and neutrophils [11]. Proinflammatory signals are mainly mediated by the accumulation of tumor necrosis factor alpha (TNF- $\alpha$ ); interleukins (IL) such as IL-6, IL-1 $\beta$ , IL-2, IL-7, IL-15, IL-17; C-C motif chemokine ligand 2 (CCL2); and interferon gamma (IFN- $\gamma$ ).

Following activation by antigen-presenting cells, T cells and especially T-helper 17 (Th17) cells and the mucosal-associated invariant T (MAIT) cells are major promoters of liver inflammation primarily by secretion of IL-17 [12, 13]. IL-17 secreted by T cells as well as transforming growth factor beta 1 (TGF- $\beta$ 1) and platelet-derived growth factor subunit B (PDGF-B) secreted by KCs and monocyte-derived macrophages are able to activate and differentiate HSC into collagen-producing myofibroblasts [12, 13]. Finally, also DAMPs can directly activate HSC and participate in fibrosis [7, 14]. HSC-derived myofibroblasts account for abnormal production of collagen in the liver and are main components of the hepatic precancerous microenvironment [15].

The inflammatory microenvironment causes hepatocellular stress, accompanied by epigenetic modifications, mitochondrial alterations, DNA damage, and



**Fig. 15.1** Chronic inflammation is a pan-etiological driver of hepatocarcinogenesis. Hepatocarcinogenesis can be induced by multiple etiological and environmental conditions. Chronic HBV and HCV infections, as well as chronic alcohol abuse and metabolic syndrome trigger the activation of the innate immune system via release of Damage-Associated Molecular Patterns (DAMPs) and Pathogen Associated Molecular Patterns (PAMPs). The persistent dysregulation of the immunological network of the liver, promoted by the secretion of pro-inflammatory cytokines/chemokines (e.g. IL-2, IL-6, IL-7, IL-15, IL-17, TGF- $\beta$ , TNF- $\alpha$ , IFN- $\gamma$ ), leads to cells death, compensatory hepatocellular proliferation, activation of cancer-associated fibroblasts (CAFs) and hepatic stellate cells (HSCs) as well as epithelial-to-mesenchymal transition (EMT). Moreover, sustained necro-inflammatory status attenuates immune-surveillance and anti-tumor immune response, by secretion of anti-inflammatory molecules (e.g. IL-10, TGF- $\beta$ , PD-L1). In addition, the activation of HSCs contributes significantly to cell proliferation (by the release of IL-1 $\beta$ , TGF- $\beta$  and LAMA5) and cirrhosis. In conclusion, cellular proliferation and EMT, further sustained by STAT3/NF- $\kappa$ B pathway activation, cirrhosis and impaired immunosurveillance activity collectively contribute to HCC development

chromosomal alterations that determine cell transformations [7]. Inflammation has been shown to upregulate nuclear factor kappa B (NF- $\kappa$ B) and signal transducer and activator of transcription 3 (STAT3) thereby affecting cell proliferation, survival, angiogenesis, and chemotaxis [16–18]. STAT3 is further induced by several other cytokines and growth factors that are known to be upregulated under conditions of chronic liver inflammation [19]. Regarding chronic HBV and HCV infection, upregulation of the cytokines lymphotoxin beta and TNF- $\alpha$  in CD4<sup>+</sup> and CD8<sup>+</sup> T cells has been shown to promote hepatocarcinogenesis [20, 21].

Collectively, persistence of infection by hepatotropic viruses or toxic condition may cause a chronic inflammatory state, accompanied by continual cell death and promotion of compensatory tissue repair mechanisms, finally resulting in liver cirrhosis and cell transformation. Since chronic inflammation induces impaired immune surveillance due to exhausted T cells, chronic inflammatory liver status not only provokes cell transformation but also attenuates physiological antitumor defense mechanisms by the immune system. Thus, tumor cell attack by cytolytic T cells is weakened in chronic inflammatory liver tissue and HCC microenvironment [22–24].

Moreover, upregulation of immunosuppressive T<sub>reg</sub> cells has been related to chronic inflammation associated with attenuated immune surveillance contributing to risk of HCC development [25, 26]. The inducible type 1 T regulatory (Tr1) cells

possess many immunosuppressive functions by secretion of the cytokines IL-10 and TGF- $\beta$ , as well as by expression of the checkpoint inhibitors cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) and programmed death 1 (PD1) on the cell surface [27–29]. T<sub>reg</sub> or KC-secreted IL-10 was reported to reduce immune surveillance by suppressing macrophage activation, T-cell proliferation, and IFN- $\gamma$  production, hereby inhibiting antitumor response mediated by the immune system [30–32]. Moreover, TGF- $\beta$  is known to inhibit IL-2-dependent T-cell proliferation as well as production of proinflammatory cytokines and performance of cytolytic functions by effector cells [33–35]. Suggesting its involvement in chronic inflammatory liver disease and contribution to hepatocarcinogenesis, levels of the immunoregulatory cytokine IL-10 and TGF- $\beta$  have been reported to be elevated in patients with chronic liver disease and related to disease progression and patients' survival [30, 36, 37].

## Immune Cells in HCC Microenvironment

Leukocytes are one of the main drivers in chronic inflammation. They are highly enriched in both the precancerous state of liver cirrhosis and in malignant tissue of HCC. Indeed, liver carcinoma is characterized by an immunogenic microenvironment, consisting of high amounts of lymphocytes, including NK cells, NKT cells, B cells, and T cells [38]. T-cell exhaustion due to chronic inflammation hereby shapes an immunogenic microenvironment that is characterized by an enhanced immunotolerance. Thus, the endogenous antitumor function of cytotoxic lymphocytes can be restored by antigen-presenting cells, which are typically reduced in the HCC microenvironment [39]. Indeed, decreased activity of NK cells, one of the most important antigen-presenting cells, correlates with an increased incidence of HCC in patients with liver cirrhosis [40]. Moreover, infiltration and density of T cells in human HCCs correlate with better patient prognosis, whereas tumor-infiltrating B cells reduce tumor viability [41].

Macrophages perpetuate chronic inflammation following liver injury and promote fibrogenesis via HSC activation. This therefore represents a significant component of HCC microenvironment. Of note, tumor-associated macrophages (TAMs) are considered to promote tumor development and favor angiogenesis and tumor cell migration [42, 43]. Moreover, TAMs may stimulate tumor growth by suppression of the adaptive immune system. They express high levels of cell death-ligand 1 (PD-L1), thereby suppressing the antitumor cytotoxic T-cell responses [44]. TAMs provide cytokines and growth factors that enhance tumor cell proliferation and NF- $\kappa$ B-mediated protection from cancer cell apoptosis and angiogenesis [45]. Accordingly, TAM infiltration correlates with HCC progression and poor survival [46, 47].

Dendritic cells (DCs) are a heterogeneous cell population and one of the most powerful antigen-presenting cells which regulate the primary immune response and the immune homeostasis in the liver [48]. By forming a bridge between the innate and the adaptive immune system [49], DCs are regarded as key players in immune

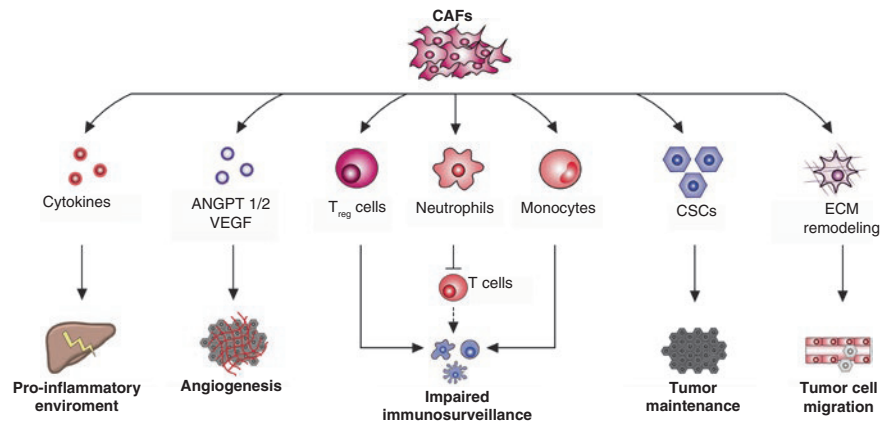
regulation [50, 51]. An impaired DC function has frequently been suggested as an important factor contributing to an immunosuppressive microenvironment in chronic liver disease, which is favoring tumor development. Accordingly, several studies report lower DC numbers in both the peripheral blood and liver tissue of patients with HCC [52, 53]. A reduced IL-12 secretion by DCs is hereby attributed to an attenuated stimulation of T cells [54]. Moreover, DC inhibition and its effects on downstream effector cells have further been identified as immune escape mechanisms of HCC [55, 56].

### **Stromal Cells Participate in HCC Development and Progression**

Liver cirrhosis is one of the main risk factors for hepatocarcinogenesis and therefore regarded as a precancerous liver state [57]. Thus, the lifetime risk of HCC development in patients with advanced liver cirrhosis is approximately 30%, and 80–90% of HCCs evolve in cirrhotic liver tissue [58, 59]. Considering HSCs as the most important progenitor cells of myofibroblasts that account for enhanced production of the extracellular matrix in liver fibrosis and liver cirrhosis, HSC-derived myofibroblasts are the main components of the hepatic precancerous microenvironment as well as the HCC tumor environment. Indeed, differentiation of HSCs from pericyte-like cells to collagen-producing myofibroblasts provides 85–95% of the myofibroblasts in liver fibrosis and liver cirrhosis, independent of the underlying trigger [15]. Hence, together with bone marrow (BM)-derived fibroblasts and portal fibroblasts (PF), HSC-derived myofibroblasts compose the stromal population of cancer-associated myofibroblasts (CAFs) that contribute actively to HCC development and progression [60]. Of note, CAFs show a markedly altered phenotype compared to normal fibroblasts [61, 62]. Normal fibroblasts may suppress tumor growth by contact inhibition [62], whereas CAFs promote an immune-tolerant tumor environment by interaction with monocytes and lymphocytes [63]. Indeed, CAFs inhibit lymphocyte tumor infiltration, increase the activity of immunosuppressive regulatory T cells, and induce apoptosis in monocytes [64, 65]. Furthermore, CAFs were reported to impair antitumor functions of T cells via activation of neutrophils [66]. CAFs may further promote hepatocarcinogenesis by downregulation of tumor-suppressive microRNAs [67, 68]. CAF activity has also been associated with tumor angiogenesis. CAFs have been shown to secrete vascular endothelial growth factor (VEGF) and angiopoietin 1 or 2 [69–71]. The cross talk between CAFs and cancer cells is crucial for HCC biology. The secretion of laminin 5 (LAMA5) [72] and IL-1 $\beta$  [73] by CAFs has been shown to promote HCC migration, and on the other hand, highly metastatic HCC cells were found to be able to convert normal fibroblasts to CAFs, which in turn promote cancer progression by secretion of proinflammatory cytokines [74]. Several studies further suggest an association of CAFs and CSCs that are thought to promote tumor development and to mediate therapeutic resistance. CAFs have been reported to recruit CSCs and to

drive their self-renewal [75, 76]. Moreover, CAFs have been observed to increase expression of keratin 19 by paracrine interactions [77], a marker for hepatic stem cells that has been observed to be correlated with poor prognosis [78]. In summary, CAFs are key drivers in hepatic carcinogenesis by increasing angiogenesis, inflammation, and proliferation and attenuating immune surveillance [60] (Fig. 15.2). CAFs correlate with HCC tumor stage and progression, tumor recurrence after surgery, as well as overall prognosis [79–81].

Lymphatic vessels function as a tissue drainage and immunological control system. They are highly enriched in the liver, carrying approximately 25–50% of the thoracic duct's lymph flow [82]. For a long time, lymphatic vessels were considered to affect carcinogenesis only by providing the structural pathway for metastatic spread of tumor cells. However, recent observations indicate a functional role of the lymphatic endothelium also in the hepatocytes' immunogenic microenvironment, which is affecting the development of chronic liver disease and hepatocarcinogenesis [83]. Thus, lymphatic endothelial cells (LECs) guide immune cell migration by lining the inner surface of lymphatic capillaries and regulate the expression of adhesion molecules and cytokines [84, 85]. Moreover, by secretion of immunosuppressive cytokines (i.e., TGF- $\beta$ ) and the overexpression of co-inhibitory checkpoint



**Fig. 15.2** Cancer-associated fibroblasts (CAFs) characterize the stromal tumor microenvironment and promote hepatocarcinogenesis, tumor progression and treatment resistance. Tumor microenvironment in HCC is predominantly characterized by cancer-associated fibroblasts (CAFs) that contribute actively to tumor development, progression and metastatic spread. Interacting with the immune cells and secreting angiogenic factors, these cells reduce immune surveillance and drive tumor angiogenesis. Moreover, CAFs promote cancer cell proliferation by paracrine interactions as well as production of prooncogenic cytokines (e.g. TGF- $\beta$ ). CAFs are also reported to recruit cancer stem cells, hereby affecting tumor maintenance, heterogeneity and treatment resistance. Finally, CAFs are responsible for the alteration of liver extracellular matrix by production and secretion of Laminin 5 and Integrin  $\beta$ 1 that further promote HCC cell invasion and migration

proteins (i.e., PD-L1), LECs suppress a maturation and proliferation of circulating immune cells [84–86]. LECs further mediate CD4<sup>+</sup> and CD8<sup>+</sup> T-cell tolerance by expression of self-antigens in the presence of inhibitory ligands [87].

Lymphangiogenesis is increased in liver fibrosis and cirrhosis and positively correlate with portal venous pressure and disease severity [88–90]. The enhanced interstitial flow and increased number of LECs is accompanied by increased cytokine production and immune cell recruitment to the inflammatory environment present in almost all chronic liver diseases [91]. The primarily immunosuppressive functions of LECs hereby contribute to an immunotolerant microenvironment favoring HCC development [83, 92]. Moreover, expression of chemokines by LECs may facilitate lymphogenic metastatic tumor spread [84]. Vascular endothelial growth factor C (VEGF-C) is an important stimulator of LEC growth and lymphangiogenesis. VEGF-C is enhanced in liver cirrhosis and HCC, and its expression in HCCs correlates with metastasis and poor patients' outcome [93, 94].

### **Epithelial-to-Mesenchymal Transition in HCC**

Epithelial-to-mesenchymal transition (EMT) describes a reversible process, by which epithelial cell types gradually develop mesenchymal characteristics leading to higher motility and invasive properties that are essential in embryogenic development and wound healing but also implicated in hepatic fibrogenesis and carcinogenesis [95, 96]. Thus, while epithelial cells are characterized by polarity and stable morphology, mesenchymal cells lack polarity, show a loose arrangement, and exhibit the capacity of migration [97]. EMT can be divided in three different biological subtypes [98]. While type 1 EMT determines embryonal development and organogenesis, types 2 and 3 EMT affect liver disease progression and can be activated by several proinflammatory cytokines and growth factors present in the inflammatory state of the liver [99].

Type 2 EMT occurs in response to cell injury as a mechanism of tissue repair and may cause fibrosis due to generation of collagen-producing fibroblasts. TGF- $\beta$ , a cytokine increased under condition of chronic inflammation, has been shown to be one of the strongest activators of type 2 EMT that can affect hepatocytes, cholangiocytes, and hepatic stellate cells (HSC) [100]. Quiescent HSCs, the most frequent progenitor cells of collagen-producing fibroblasts [15], are actually regarded as transitional cells that have undergone partial EMT from epithelial cells and may complete transition upon inflammatory signals [101]. Hence, EMT is regarded as one of the most important promoters of liver fibrogenesis in response to chronic inflammation [101].

Type 3 EMT may occur due to genetic and epigenetic changes during malignant transformation of epithelial cells and is implicated in HCC growth and progression [3]. Cells generated by type 3 EMT differ significantly from types 1 and 2 EMT cells and develop properties of invasion and migration as well as escape from apop-



tosis. Weakened or loss of E-cadherin expression, characteristic for development of the mesenchymal unpolarized phenotype, could be revealed in 58% of human HCC patients and correlated with the presence of metastases and patients' survival [102]. Besides proinflammatory cytokines and growth factors, several studies further indicate induction of type 3 EMT by core proteins of HCV itself [103]. Given not only the correlation of EMT with tumor stage but also response to therapy [104], therapeutic targeting of molecular key players in EMT is highly clinically relevant.

## Clinical Perspectives

Considering the implication of stromal and immunogenic cell compounds in HCC development and progression, medical treatments targeting these factors represent promising tools for future medical treatment of advanced HCC. Presently, sorafenib, an oral multikinase inhibitor targeting vascular endothelial growth factor receptor (VEGFR-2/VEGFR-3) and platelet-derived growth factor receptor (PDGFR), produced by the stromal HCC microenvironment already represents the standard of care treatment for patients with advanced HCC [105]. Lenvatinib, another tyrosine kinase inhibitor with multiple targets, has recently been revealed to be noninferior compared to sorafenib according to the REFLECT trial and has lately been approved by the FDA as first-line treatment for unresectable HCC [106]. Moreover, recently therapeutic strategies targeting the immunogenic tumor microenvironment have been demonstrated to be effective as systemic therapy for several cancer types. Consequently, drugs targeting exhausted lymphocytes expressing PD1 and infiltrating the tumor are able to activate T-cell-driven immune response against cancer cells and were approved for melanoma and non-small cell lung cancer treatment [107, 108]. Preliminary results from open-label trials of these drugs in HCC treatment are encouraging. Indeed, nivolumab and pembrolizumab, anti-PD1 monoclonal antibodies, have been demonstrated to be more effective than placebo in patients with advanced unresectable HCC previously treated with sorafenib [109, 110]. For that reason, these compounds were recently approved by FDA as a second-line treatment for advanced HCC. Moreover, currently several randomized controlled trials investigate the effects of other drugs targeting the HCC immunogenic and stromal microenvironment. Thus, aiming to activate tumor-targeting cytotoxic T lymphocytes, a growing number of studies recently worked on ex vivo tumor-antigen-loaded dendritic cells as an approach of cancer immunotherapy by DC vaccination [111–113]. Several other studies are focused on immunotherapy targeting TAMs, aiming to decrease TAM population present in the HCC by elimination, blocking recruitment, or functional reprogramming of TAM polarization [43]. The results of current ongoing clinical studies are expected in the next few years and may revolutionize future HCC medical treatment.



## References

1. Jenne CN, Kubes P. Immune surveillance by the liver. *Nat Immunol.* 2013;14:996. <https://doi.org/10.1038/ni.2691>. <https://www.nature.com/articles/ni.2691#supplementary-information>.
2. Marrone G, Shah VH, Gracia-Sancho J. Sinusoidal communication in liver fibrosis and regeneration. *J Hepatol.* 2016;65(3):608–17. <https://doi.org/10.1016/j.jhep.2016.04.018>.
3. van Zijl F, Zulehner G, Petz M, Schneller D, Kornauth C, Hau M, et al. Epithelial-mesenchymal transition in hepatocellular carcinoma. *Future Oncol.* 2009;5(8):1169–79. <https://doi.org/10.2217/fon.09.91>.
4. Nishida N, Kudo M. Oncogenic signal and tumor microenvironment in hepatocellular carcinoma. *Oncology.* 2017;93(Suppl 1):160–4. <https://doi.org/10.1159/000481246>.
5. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med.* 2008;359(19):1995–2004. <https://doi.org/10.1056/NEJMoa0804525>.
6. Caja L, Dituri F, Mancarella S, Caballero-Diaz D, Moustakas A, Giannelli G, et al. TGF-beta and the tissue microenvironment: relevance in fibrosis and cancer. *Int J Mol Sci.* 2018;19(5) <https://doi.org/10.3390/ijms19051294>.
7. Ringelhan M, Pfister D, O'Connor T, Pikarsky E, Heikenwalder M. The immunology of hepatocellular carcinoma. *Nat Immunol.* 2018;19(3):222–32. <https://doi.org/10.1038/s41590-018-0044-z>.
8. Robinson MW, Harmon C, O'Farrelly C. Liver immunology and its role in inflammation and homeostasis. *Cell Mol Immunol.* 2016;13(3):267–76. <https://doi.org/10.1038/cmi.2016.3>.
9. Llovet JM, Zucman-Rossi J, Pikarsky E, Sangro B, Schwartz M, Sherman M, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers.* 2016;2:16018. <https://doi.org/10.1038/nrdp.2016.18>.
10. Pellicoro A, Ramachandran P, Iredale JP, Fallowfield JA. Liver fibrosis and repair: immune regulation of wound healing in a solid organ. *Nat Rev Immunol.* 2014;14(3):181–94. <https://doi.org/10.1038/nri3623>.
11. Mossanen JC, Krenkel O, Ergen C, Govaere O, Liepelt A, Puengel T, et al. Chemokine (C-C motif) receptor 2-positive monocytes aggravate the early phase of acetaminophen-induced acute liver injury. *Hepatology.* 2016;64(5):1667–82. <https://doi.org/10.1002/hep.28682>.
12. Lemmers A, Moreno C, Gustot T, Marechal R, Degre D, Demetter P, et al. The interleukin-17 pathway is involved in human alcoholic liver disease. *Hepatology.* 2009;49(2):646–57. <https://doi.org/10.1002/hep.22680>.
13. Meng F, Wang K, Aoyama T, Grivennikov SI, Paik Y, Scholten D, et al. Interleukin-17 signaling in inflammatory, Kupffer cells, and hepatic stellate cells exacerbates liver fibrosis in mice. *Gastroenterology.* 2012;143(3):765–76 e3. <https://doi.org/10.1053/j.gastro.2012.05.049>.
14. Tu T, Calabro SR, Lee A, Maczurek AE, Budzinska MA, Warner FJ, et al. Hepatocytes in liver injury: victim, bystander, or accomplice in progressive fibrosis? *J Gastroenterol Hepatol.* 2015;30(12):1696–704. <https://doi.org/10.1111/jgh.13065>.
15. Mederacke I, Hsu CC, Troeger JS, Huebener P, Mu X, Dapito DH, et al. Fate tracing reveals hepatic stellate cells as dominant contributors to liver fibrosis independent of its aetiology. *Nat Commun.* 2013;4:2823. <https://doi.org/10.1038/ncomms3823>.
16. Maeda S, Kamata H, Luo JL, Leffert H, Karin M. IKKbeta couples hepatocyte death to cytokine-driven compensatory proliferation that promotes chemical hepatocarcinogenesis. *Cell.* 2005;121(7):977–90. <https://doi.org/10.1016/j.cell.2005.04.014>.
17. Akira S, Nishio Y, Inoue M, Wang XJ, Wei S, Matsusaka T, et al. Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway. *Cell.* 1994;77(1):63–71.
18. Mackey-Lawrence NM, Petri WA Jr. Leptin and mucosal immunity. *Mucosal Immunol.* 2012;5(5):472–9. <https://doi.org/10.1038/mi.2012.40>.

19. McCartney EM, Helbig KJ, Narayana SK, Eyre NS, Aloia AL, Beard MR. Signal transducer and activator of transcription 3 is a proviral host factor for hepatitis C virus. *Hepatology*. 2013;58(5):1558–68. <https://doi.org/10.1002/hep.26496>.
20. Haybaeck J, Zeller N, Wolf MJ, Weber A, Wagner U, Kurrer MO, et al. A lymphotoxin-driven pathway to hepatocellular carcinoma. *Cancer Cell*. 2009;16(4):295–308. <https://doi.org/10.1016/j.ccr.2009.08.021>.
21. Wolf MJ, Adili A, Piotrowitz K, Abdullah Z, Boege Y, Stemmer K, et al. Metabolic activation of intrahepatic CD8+ T cells and NKT cells causes nonalcoholic steatohepatitis and liver cancer via cross-talk with hepatocytes. *Cancer Cell*. 2014;26(4):549–64. <https://doi.org/10.1016/j.ccell.2014.09.003>.
22. Kang TW, Yevsa T, Woller N, Hoenicke L, Wuestefeld T, Dauch D, et al. Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. *Nature*. 2011;479(7374):547–51. <https://doi.org/10.1038/nature10599>.
23. Ma C, Kesarwala AH, Eggert T, Medina-Echeverez J, Kleiner DE, Jin P, et al. NAFLD causes selective CD4(+) T lymphocyte loss and promotes hepatocarcinogenesis. *Nature*. 2016;531(7593):253–7. <https://doi.org/10.1038/nature16969>.
24. van der Windt DJ, Sud V, Zhang H, Varley PR, Goswami J, Yazdani HO, et al. Neutrophil extracellular traps promote inflammation and development of hepatocellular carcinoma in nonalcoholic steatohepatitis. *Hepatology*. 2018;68(4):1347–60. <https://doi.org/10.1002/hep.29914>.
25. Jiang R, Tang J, Chen Y, Deng L, Ji J, Xie Y, et al. The long noncoding RNA Inc-EGFR stimulates T-regulatory cells differentiation thus promoting hepatocellular carcinoma immune evasion. *Nat Commun*. 2017;8:15129. <https://doi.org/10.1038/ncomms15129>.
26. Li K, Liu H, Guo T. Th17/Treg imbalance is an indicator of liver cirrhosis process and a risk factor for HCC occurrence in HBV patients. *Clin Res Hepatol Gastroenterol*. 2017;41(4):399–407. <https://doi.org/10.1016/j.clinre.2016.12.004>.
27. Read S, Malmstrom V, Powrie F. Cytotoxic T lymphocyte-associated antigen 4 plays an essential role in the function of CD25(+)CD4(+) regulatory cells that control intestinal inflammation. *J Exp Med*. 2000;192(2):295–302.
28. Nishimura H, Nose M, Hiai H, Minato N, Honjo T. Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor. *Immunity*. 1999;11(2):141–51.
29. Carter L, Fouser LA, Jussif J, Fitz L, Deng B, Wood CR, et al. PD-1:PD-L inhibitory pathway affects both CD4(+) and CD8(+) T cells and is overcome by IL-2. *Eur J Immunol*. 2002;32(3):634–43. [https://doi.org/10.1002/1521-4141\(200203\)32:3<634::AID-IMMU634>3.0.CO;2-9](https://doi.org/10.1002/1521-4141(200203)32:3<634::AID-IMMU634>3.0.CO;2-9).
30. Knolle P, Schlaak J, Uhrig A, Kempf P, Meyer zum Buschenfelde KH, Gerken G. Human Kupffer cells secrete IL-10 in response to lipopolysaccharide (LPS) challenge. *J Hepatol*. 1995;22(2):226–9.
31. Hattori E, Okumoto K, Adachi T, Takeda T, Ito J, Sugahara K, et al. Possible contribution of circulating interleukin-10 (IL-10) to anti-tumor immunity and prognosis in patients with unresectable hepatocellular carcinoma. *Hepatol Res*. 2003;27(4):309–14.
32. Moore KW, de Waal Malefyt R, Coffman RL, O'Garra A. Interleukin-10 and the interleukin-10 receptor. *Annu Rev Immunol*. 2001;19:683–765. <https://doi.org/10.1146/annurev.immunol.19.1.683>.
33. Kehrl JH, Wakefield LM, Roberts AB, Jakowlew S, Alvarez-Mon M, Derynck R, et al. Production of transforming growth factor beta by human T lymphocytes and its potential role in the regulation of T cell growth. *J Exp Med*. 1986;163(5):1037–50.
34. Espevik T, Waage A, Faxvaag A, Shalaby MR. Regulation of interleukin-2 and interleukin-6 production from T-cells: involvement of interleukin-1 beta and transforming growth factor-beta. *Cell Immunol*. 1990;126(1):47–56.
35. Smyth MJ, Strobl SL, Young HA, Ortaldo JR, Ochoa AC. Regulation of lymphokine-activated killer activity and pore-forming protein gene expression in human peripheral blood CD8+ T lymphocytes. Inhibition by transforming growth factor-beta. *J Immunol*. 1991;146(10):3289–97.

36. Othman MS, Aref AM, Mohamed AA, Ibrahim WA. Serum levels of Interleukin-6 and Interleukin-10 as biomarkers for hepatocellular carcinoma in Egyptian patients. *ISRN Hepatol.* 2013;2013:412317. <https://doi.org/10.1155/2013/412317>.
37. El-Emshaty HM, Nasif WA, Mohamed IE. Serum cytokine of IL-10 and IL-12 in chronic liver disease: the immune and inflammatory response. *Dis Markers.* 2015;2015:707254. <https://doi.org/10.1155/2015/707254>.
38. Thomson AW, Knolle PA. Antigen-presenting cell function in the tolerogenic liver environment. *Nat Rev Immunol.* 2010;10(11):753–66. <https://doi.org/10.1038/nri2858>.
39. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity.* 2013;39(1):1–10. <https://doi.org/10.1016/j.immuni.2013.07.012>.
40. Nakajima T, Mizushima N, Kanai K. Relationship between natural killer activity and development of hepatocellular carcinoma in patients with cirrhosis of the liver. *Jpn J Clin Oncol.* 1987;17(4):327–32.
41. Garnelo M, Tan A, Her Z, Yeong J, Lim CJ, Chen J, et al. Interaction between tumour-infiltrating B cells and T cells controls the progression of hepatocellular carcinoma. *Gut.* 2017;66(2):342–51. <https://doi.org/10.1136/gutjnl-2015-310814>.
42. Condeelis J, Pollard JW. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell.* 2006;124(2):263–6. <https://doi.org/10.1016/j.cell.2006.01.007>.
43. Degroote H, Van Dierendonck A, Geerts A, Van Vlierberghe H, Devisscher L. Preclinical and clinical therapeutic strategies affecting tumor-associated macrophages in hepatocellular carcinoma. *J Immunol Res.* 2018;2018:7819520. <https://doi.org/10.1155/2018/7819520>.
44. Wu K, Kryczek I, Chen L, Zou W, Welling TH. Kupffer cell suppression of CD8+ T cells in human hepatocellular carcinoma is mediated by B7-H1/programmed death-1 interactions. *Cancer Res.* 2009;69(20):8067–75. <https://doi.org/10.1158/0008-5472.CAN-09-0901>.
45. Tacke F. Targeting hepatic macrophages to treat liver diseases. *J Hepatol.* 2017;66(6):1300–12. <https://doi.org/10.1016/j.jhep.2017.02.026>.
46. Yeung OW, Lo CM, Ling CC, Qi X, Geng W, Li CX, et al. Alternatively activated (M2) macrophages promote tumour growth and invasiveness in hepatocellular carcinoma. *J Hepatol.* 2015;62(3):607–16. <https://doi.org/10.1016/j.jhep.2014.10.029>.
47. Ding T, Xu J, Wang F, Shi M, Zhang Y, Li SP, et al. High tumor-infiltrating macrophage density predicts poor prognosis in patients with primary hepatocellular carcinoma after resection. *Hum Pathol.* 2009;40(3):381–9. <https://doi.org/10.1016/j.humpath.2008.08.011>.
48. Banchereau J, Steinman RM. Dendritic cells and the control of immunity. *Nature.* 1998;392(6673):245–52. <https://doi.org/10.1038/32588>.
49. Steinman RM, Hemmi H. Dendritic cells: translating innate to adaptive immunity. *Curr Top Microbiol Immunol.* 2006;311:17–58.
50. Almand B, Resser JR, Lindman B, Nadaf S, Clark JI, Kwon ED, et al. Clinical significance of defective dendritic cell differentiation in cancer. *Clin Cancer Res.* 2000;6(5):1755–66.
51. Steinman RM. Lasker Basic Medical Research Award. Dendritic cells: versatile controllers of the immune system. *Nat Med.* 2007;13(10):1155–9. <https://doi.org/10.1038/nm1643>.
52. Kakumu S, Ito S, Ishikawa T, Mita Y, Tagaya T, Fukuzawa Y, et al. Decreased function of peripheral blood dendritic cells in patients with hepatocellular carcinoma with hepatitis B and C virus infection. *J Gastroenterol Hepatol.* 2000;15(4):431–6.
53. Chen S, Akbar SM, Tanimoto K, Ninomiya T, Iuchi H, Michitaka K, et al. Absence of CD83-positive mature and activated dendritic cells at cancer nodules from patients with hepatocellular carcinoma: relevance to hepatocarcinogenesis. *Cancer Lett.* 2000;148(1):49–57.
54. Ormandy LA, Farber A, Cantz T, Petrykowska S, Wedemeyer H, Horning M, et al. Direct ex vivo analysis of dendritic cells in patients with hepatocellular carcinoma. *World J Gastroenterol.* 2006;12(20):3275–82.
55. Gabrilovich DI, Chen HL, Girgis KR, Cunningham HT, Meny GM, Nadaf S, et al. Production of vascular endothelial growth factor by human tumors inhibits the functional maturation of dendritic cells. *Nat Med.* 1996;2(10):1096–103.
56. Beckebaum S, Zhang X, Chen X, Yu Z, Frilling A, Dworacki G, et al. Increased levels of interleukin-10 in serum from patients with hepatocellular carcinoma correlate with profound

- numerical deficiencies and immature phenotype of circulating dendritic cell subsets. *Clin Cancer Res*. 2004;10(21):7260–9. <https://doi.org/10.1158/1078-0432.CCR-04-0872>.
57. Maier KP. Cirrhosis of the liver as a precancerous condition. *Praxis (Bern 1994)*. 1998;87(44):1462–5.
  58. Singal AG, El-Serag HB. Hepatocellular carcinoma from epidemiology to prevention: translating knowledge into practice. *Clin Gastroenterol Hepatol*. 2015;13(12):2140–51. <https://doi.org/10.1016/j.cgh.2015.08.014>.
  59. El-Serag HB. Hepatocellular carcinoma. *N Engl J Med*. 2011;365(12):1118–27. <https://doi.org/10.1056/NEJMra1001683>.
  60. Affo S, Yu LX, Schwabe RF. The role of cancer-associated fibroblasts and fibrosis in liver cancer. *Annu Rev Pathol*. 2017;12:153–86. <https://doi.org/10.1146/annurev-pathol-052016-100322>.
  61. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646–74. <https://doi.org/10.1016/j.cell.2011.02.013>.
  62. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med*. 2011;17(3):320–9. <https://doi.org/10.1038/nm.2328>.
  63. Ji J, Eggert T, Budhu A, Forgues M, Takai A, Dang H, et al. Hepatic stellate cell and monocyte interaction contributes to poor prognosis in hepatocellular carcinoma. *Hepatology*. 2015;62(2):481–95. <https://doi.org/10.1002/hep.27822>.
  64. Zhao W, Su W, Kuang P, Zhang L, Liu J, Yin Z, et al. The role of hepatic stellate cells in the regulation of T-cell function and the promotion of hepatocellular carcinoma. *Int J Oncol*. 2012;41(2):457–64. <https://doi.org/10.3892/ijo.2012.1497>.
  65. Zhao W, Zhang L, Yin Z, Su W, Ren G, Zhou C, et al. Activated hepatic stellate cells promote hepatocellular carcinoma development in immunocompetent mice. *Int J Cancer*. 2011;129(11):2651–61. <https://doi.org/10.1002/ijc.25920>.
  66. Cheng Y, Li H, Deng Y, Tai Y, Zeng K, Zhang Y, et al. Cancer-associated fibroblasts induce PDL1+ neutrophils through the IL6-STAT3 pathway that foster immune suppression in hepatocellular carcinoma. *Cell Death Dis*. 2018;9(4):422. <https://doi.org/10.1038/s41419-018-0458-4>.
  67. Zhang Z, Li X, Sun W, Yue S, Yang J, Li J, et al. Loss of exosomal miR-320a from cancer-associated fibroblasts contributes to HCC proliferation and metastasis. *Cancer Lett*. 2017;397:33–42. <https://doi.org/10.1016/j.canlet.2017.03.004>.
  68. Wang F, Li L, Piontek K, Sakaguchi M, Selaru FM. Exosome miR-335 as a novel therapeutic strategy in hepatocellular carcinoma. *Hepatology*. 2018;67(3):940–54. <https://doi.org/10.1002/hep.29586>.
  69. Torimura T, Ueno T, Kin M, Harada R, Taniguchi E, Nakamura T, et al. Overexpression of angiopoietin-1 and angiopoietin-2 in hepatocellular carcinoma. *J Hepatol*. 2004;40(5):799–807. <https://doi.org/10.1016/j.jhep.2004.01.027>.
  70. Taura K, De Minicis S, Seki E, Hatano E, Iwaisako K, Osterreicher CH, et al. Hepatic stellate cells secrete angiopoietin 1 that induces angiogenesis in liver fibrosis. *Gastroenterology*. 2008;135(5):1729–38. <https://doi.org/10.1053/j.gastro.2008.07.065>.
  71. Lin N, Chen Z, Lu Y, Li Y, Hu K, Xu R. Role of activated hepatic stellate cells in proliferation and metastasis of hepatocellular carcinoma. *Hepatol Res*. 2015;45(3):326–36. <https://doi.org/10.1111/hepr.12356>.
  72. Santamato A, Fransvea E, Dituri F, Caligiuri A, Quaranta M, Niimi T, et al. Hepatic stellate cells stimulate HCC cell migration via laminin-5 production. *Clin Sci (Lond)*. 2011;121(4):159–68. <https://doi.org/10.1042/CS20110002>.
  73. Okabe H, Beppu T, Ueda M, Hayashi H, Ishiko T, Masuda T, et al. Identification of CXCL5/ENA-78 as a factor involved in the interaction between cholangiocarcinoma cells and cancer-associated fibroblasts. *Int J Cancer*. 2012;131(10):2234–41. <https://doi.org/10.1002/ijc.27496>.
  74. Fang T, Lv H, Lv G, Li T, Wang C, Han Q, et al. Tumor-derived exosomal miR-1247-3p induces cancer-associated fibroblast activation to foster lung metastasis of liver cancer. *Nat Commun*. 2018;9(1):191. <https://doi.org/10.1038/s41467-017-02583-0>.

75. Jiang J, Ye F, Yang X, Zong C, Gao L, Yang Y, et al. Peri-tumor associated fibroblasts promote intrahepatic metastasis of hepatocellular carcinoma by recruiting cancer stem cells. *Cancer Lett.* 2017;404:19–28. <https://doi.org/10.1016/j.canlet.2017.07.006>.
76. Liu C, Liu L, Chen X, Cheng J, Zhang H, Zhang C, et al. LSD1 stimulates cancer-associated fibroblasts to drive Notch3-dependent self-renewal of liver cancer stem-like cells. *Cancer Res.* 2018;78(4):938–49. <https://doi.org/10.1158/0008-5472.CAN-17-1236>.
77. Rhee H, Kim HY, Choi JH, Woo HG, Yoo JE, Nahm JH, et al. Keratin 19 expression in hepatocellular carcinoma is regulated by fibroblast-derived HGF via a MET-ERK1/2-AP1 and SP1 Axis. *Cancer Res.* 2018;78(7):1619–31. <https://doi.org/10.1158/0008-5472.CAN-17-0988>.
78. Kim H, Choi GH, Na DC, Ahn EY, Kim GI, Lee JE, et al. Human hepatocellular carcinomas with “Stemness”-related marker expression: keratin 19 expression and a poor prognosis. *Hepatology.* 2011;54(5):1707–17. <https://doi.org/10.1002/hep.24559>.
79. Coulouarn C, Corlu A, Glaire D, Guenon I, Thorgeirsson SS, Clement B. Hepatocyte-stellate cell cross-talk in the liver engenders a permissive inflammatory microenvironment that drives progression in hepatocellular carcinoma. *Cancer Res.* 2012;72(10):2533–42. <https://doi.org/10.1158/0008-5472.CAN-11-3317>.
80. Ju MJ, Qiu SJ, Fan J, Xiao YS, Gao Q, Zhou J, et al. Peritumoral activated hepatic stellate cells predict poor clinical outcome in hepatocellular carcinoma after curative resection. *Am J Clin Pathol.* 2009;131(4):498–510. <https://doi.org/10.1309/AJCP86PPBNGOHNLL>.
81. Zhang DY, Goossens N, Guo J, Tsai MC, Chou HI, Altunkaynak C, et al. A hepatic stellate cell gene expression signature associated with outcomes in hepatitis C cirrhosis and hepatocellular carcinoma after curative resection. *Gut.* 2016;65(10):1754–64. <https://doi.org/10.1136/gutjnl-2015-309655>.
82. Pupulin LF, Vilgrain V, Ronot M, Becker CD, Breguet R, Terraz S. Hepatic lymphatics: anatomy and related diseases. *Abdom Imaging.* 2015;40(6):1997–2011. <https://doi.org/10.1007/s00261-015-0350-y>.
83. Lund AW, Wagner M, Fankhauser M, Steinskog ES, Broggi MA, Spranger S, et al. Lymphatic vessels regulate immune microenvironments in human and murine melanoma. *J Clin Invest.* 2016;126(9):3389–402. <https://doi.org/10.1172/JCI79434>.
84. Swartz MA. Immunomodulatory roles of lymphatic vessels in cancer progression. *Cancer Immunol Res.* 2014;2(8):701–7. <https://doi.org/10.1158/2326-6066.CIR-14-0115>.
85. Lukacs-Kornek V, Malhotra D, Fletcher AL, Acton SE, Elpek KG, Tayalia P, et al. Regulated release of nitric oxide by nonhematopoietic stroma controls expansion of the activated T cell pool in lymph nodes. *Nat Immunol.* 2011;12(11):1096–104. <https://doi.org/10.1038/ni.2112>.
86. Fletcher AL, Lukacs-Kornek V, Reynoso ED, Pinner SE, Bellemare-Pelletier A, Curry MS, et al. Lymph node fibroblastic reticular cells directly present peripheral tissue antigen under steady-state and inflammatory conditions. *J Exp Med.* 2010;207(4):689–97. <https://doi.org/10.1084/jem.20092642>.
87. Rouhani SJ, Eccles JD, Riccardi P, Peske JD, Tewalt EF, Cohen JN, et al. Roles of lymphatic endothelial cells expressing peripheral tissue antigens in CD4 T-cell tolerance induction. *Nat Commun.* 2015;6:6771. <https://doi.org/10.1038/ncomms7771>.
88. Vollmar B, Wolf B, Siegmund S, Katsen AD, Menger MD. Lymph vessel expansion and function in the development of hepatic fibrosis and cirrhosis. *Am J Pathol.* 1997;151(1):169–75.
89. Yamauchi Y, Michitaka K, Onji M. Morphometric analysis of lymphatic and blood vessels in human chronic viral liver diseases. *Am J Pathol.* 1998;153(4):1131–7. [https://doi.org/10.1016/S0002-9440\(10\)65657-X](https://doi.org/10.1016/S0002-9440(10)65657-X).
90. Yokomori H, Oda M, Kaneko F, Kawachi S, Tanabe M, Yoshimura K, et al. Lymphatic marker podoplanin/D2-40 in human advanced cirrhotic liver--re-evaluations of microlymphatic abnormalities. *BMC Gastroenterol.* 2010;10:131. <https://doi.org/10.1186/1471-230X-10-131>.
91. Limatola E, Filosa S. Exogenous vitellogenesis and micropinocytosis in the lizard, *Podarcis sicula*, treated with follicle-stimulating hormone. *Gen Comp Endocrinol.* 1989;75(2):165–76.



92. Shields JD, Kourtis IC, Tomei AA, Roberts JM, Swartz MA. Induction of lymphoid-like stroma and immune escape by tumors that express the chemokine CCL21. *Science*. 2010;328(5979):749–52. <https://doi.org/10.1126/science.1185837>.
93. Xiang Z, Zeng Z, Tang Z, Fan J, Sun H, Wu W, et al. Increased expression of vascular endothelial growth factor-C and nuclear CXCR4 in hepatocellular carcinoma is correlated with lymph node metastasis and poor outcome. *Cancer J*. 2009;15(6):519–25. <https://doi.org/10.1097/PPO.0b013e3181c6aa6b>.
94. Yamaguchi R, Yano H, Nakashima O, Akiba J, Nishida N, Kurogi M, et al. Expression of vascular endothelial growth factor-C in human hepatocellular carcinoma. *J Gastroenterol Hepatol*. 2006;21(1 Pt 1):152–60. <https://doi.org/10.1111/j.1440-1746.2005.04217.x>.
95. Nieto MA, Huang RY, Jackson RA, Thiery JP. EMT: 2016. *Cell*. 2016;166(1):21–45. <https://doi.org/10.1016/j.cell.2016.06.028>.
96. Dongre A, Weinberg RA. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat Rev Mol Cell Biol*. 2018; <https://doi.org/10.1038/s41580-018-0080-4>.
97. Polyak K, Weinberg RA. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer*. 2009;9(4):265–73. <https://doi.org/10.1038/nrc2620>.
98. Acloque H, Adams MS, Fishwick K, Bronner-Fraser M, Nieto MA. Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *J Clin Invest*. 2009;119(6):1438–49. <https://doi.org/10.1172/JCI38019>.
99. Yan L, Xu F, Dai CL. Relationship between epithelial-to-mesenchymal transition and the inflammatory microenvironment of hepatocellular carcinoma. *J Exp Clin Cancer Res*. 2018;37(1):203. <https://doi.org/10.1186/s13046-018-0887-z>.
100. Zavadil J, Bottinger EP. TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene*. 2005;24(37):5764–74. <https://doi.org/10.1038/sj.onc.1208927>.
101. Choi SS, Diehl AM. Epithelial-to-mesenchymal transitions in the liver. *Hepatology*. 2009;50(6):2007–13. <https://doi.org/10.1002/hep.23196>.
102. Zhai B, Yan HX, Liu SQ, Chen L, Wu MC, Wang HY. Reduced expression of E-cadherin/catenin complex in hepatocellular carcinomas. *World J Gastroenterol*. 2008;14(37):5665–73.
103. Battaglia S, Benzoubir N, Nobilet S, Charneau P, Samuel D, Zignego AL, et al. Liver cancer-derived hepatitis C virus core proteins shift TGF-beta responses from tumor suppression to epithelial-mesenchymal transition. *PLoS One*. 2009;4(2):e4355. <https://doi.org/10.1371/journal.pone.0004355>.
104. Fuchs BC, Fujii T, Dorfman JD, Goodwin JM, Zhu AX, Lanuti M, et al. Epithelial-to-mesenchymal transition and integrin-linked kinase mediate sensitivity to epidermal growth factor receptor inhibition in human hepatoma cells. *Cancer Res*. 2008;68(7):2391–9. <https://doi.org/10.1158/0008-5472.CAN-07-2460>.
105. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med*. 2008;359(4):378–90. <https://doi.org/10.1056/NEJMoa0708857>.
106. Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet*. 2018;391(10126):1163–73. [https://doi.org/10.1016/S0140-6736\(18\)30207-1](https://doi.org/10.1016/S0140-6736(18)30207-1).
107. Sui H, Ma N, Wang Y, Li H, Liu X, Su Y, et al. Anti-PD-1/PD-L1 therapy for non-small-cell lung cancer: toward personalized medicine and combination strategies. *J Immunol Res*. 2018;2018:6984948. <https://doi.org/10.1155/2018/6984948>.
108. Herzberg B, Fisher DE. Metastatic melanoma and immunotherapy. *Clin Immunol*. 2016;172:105–10. <https://doi.org/10.1016/j.clim.2016.07.006>.
109. Killock D. Immunotherapy: Nivolumab keeps HCC in check and opens avenues for check-mate. *Nat Rev Clin Oncol*. 2017;14(7):392. <https://doi.org/10.1038/nrclinonc.2017.70>.

110. Zhu AX, Finn RS, Edeline J, Cattan S, Ogasawara S, Palmer D, et al. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. *Lancet Oncol.* 2018;19(7):940–52. [https://doi.org/10.1016/S1470-2045\(18\)30351-6](https://doi.org/10.1016/S1470-2045(18)30351-6).
111. Shang N, Figini M, Shangguan J, Wang B, Sun C, Pan L, et al. Dendritic cells based immunotherapy. *Am J Cancer Res.* 2017;7(10):2091–102.
112. Palmer DH, Midgley RS, Mirza N, Torr EE, Ahmed F, Steele JC, et al. A phase II study of adoptive immunotherapy using dendritic cells pulsed with tumor lysate in patients with hepatocellular carcinoma. *Hepatology.* 2009;49(1):124–32. <https://doi.org/10.1002/hep.22626>.
113. El Ansary M, Mogawer S, Elhamid SA, Alwakil S, Aboelkasem F, Sabaawy HE, et al. Immunotherapy by autologous dendritic cell vaccine in patients with advanced HCC. *J Cancer Res Clin Oncol.* 2013;139(1):39–48. <https://doi.org/10.1007/s00432-012-1298-8>.



## 8. CURRICULUM VITAE

PERSONAL INFORMATION Antonio SAVIANO was born in Naples (Italy) on 25/02/1987

### WORK EXPERIENCE, EDUCATION AND TRAINING

December 2018 - Present

#### PhD Program

Institut de Recherche sur les maladies virales et hépatiques, INSERM UMR\_S1110 - Université de Strasbourg

- **Director:** Prof. Thomas Baumert, **Co-director:** Prof. Patrick Pessaux
- **Title:** « Physiopathologie du foie à l'échelle de la cellule unique : caractérisation de l'hétérogénéité cellulaire et identification de nouvelles cibles thérapeutiques dans les maladies hépatiques chroniques et le cancer hépatocellulaire »

September 2018 - Present

#### Associate Physician (Praticien attaché)

Service H.G.E., Pole Hépto-digestif, NHC, Hôpitaux Universitaires de Strasbourg, F-Strasbourg

July 2012- July 2017

#### Fellowship and Specialization in Internal Medicine

Catholic University of Rome – “A. Gemelli” Hospital, Department of Internal Medicine

- **Grade:** 50/50 cum laude.
- **Thesis title:** “Retrospective evaluation of contrast enhanced ultrasound for the diagnosis of local lesions in patients affected by chronic liver disease. Validation of CEUS LI-RADS criteria”
- **Special skills acquired:** management of patients referred to inpatient and outpatient Liver Disease Clinic and Emergency department, autonomy in execution of Internistic and General Ultrasound (US), Doppler US of the Liver, Vascular US, basics musculoskeletal US and echocardiography.

June 2013

#### ALS - Advanced Life Support Certification

IRC - Italian Resuscitation Council

October 2005- October 2011

#### Medicine and Surgery Degree (M.D.)

Catholic University of Rome – “A. Gemelli” Hospital, Department of Internal Medicine

- **Grade:** 110/110 cum laude.
- **Thesis title:** “ Surgical resection and radiofrequency ablation for treatment of single hepatocellular carcinoma up to 3 cm complicating liver cirrhosis. A comparative effectiveness analysis in a multicentre series.”

October 2008- October 2011

#### Internship in Internal Medicine Department and Liver Disease Clinic

Catholic University of Rome – “A. Gemelli” Hospital, Department of Internal Medicine

July 2010- September 2010

#### LLP-Erasmus Program, Internship in Internal Medicine and Hepato-Gastroenterology Department

Centre Hospitalier Régional Universitaire (CHRU) Brest - France

Publications in journals  
with impact factor

- Pompili M, **Saviano A**, de Matthaeis N, Cucchetti A, Ardito F, Federico B, et al. Long-term effectiveness of resection and radiofrequency ablation for single hepatocellular carcinoma  $\leq 3$  cm. Results of a multicenter Italian survey. *Journal of Hepatology*. 2013; 59(1):89-97.
- Francica G, **Saviano A**, De Sio I, De Matthaeis N, Brunello F, Cantamessa A, et al. Long-term effectiveness of radiofrequency ablation for solitary small hepatocellular carcinoma: a retrospective analysis of 363 patients. *Digestive and Liver Disease*. 2013; 45(4):336-41.
- Iezzi R, Pompili M, La Torre MF, Campanale MC, Montagna M, **Saviano A**, et al. Radiofrequency ablation plus drug-eluting beads transcatheter arterial chemoembolization for the treatment of single large hepatocellular carcinoma. *Digestive and Liver Disease*. 2015; 47(3):242-8.
- Pompili M, De Matthaeis N, **Saviano A**, De Sio I, Francica G, Brunello F, et al. Single hepatocellular carcinoma smaller than 2 cm: are ethanol injection and radiofrequency ablation equally effective? *Anticancer research*. 2015; 35(1):325-32.
- Ciresa M, De Gaetano A, Pompili M, **Saviano A**, Infante A, Montagna M, et al. Enhancement patterns of intrahepatic mass-forming cholangiocarcinoma at multiphasic computed tomography and magnetic resonance imaging and correlation with clinicopathologic features. *Eur Rev Med Pharmacol Sci*. 2015; 19(15):2786-97.
- Abbate V, Marcantoni M, Giuliani F, Vecchio FM, Gatto I, Mele C, **Saviano A** et al. HepPar1-Positive Circulating Microparticles Are Increased in Subjects with Hepatocellular Carcinoma and Predict Early Recurrence after Liver Resection. *Int J Mol Sci*. 2017 May 12;18(5).
- **Saviano A**, Iezzi R, Giuliani F, Salvatore L, Mele C, Posa A, Ardito F, De Gaetano AM, Pompili M; HepatoCATT Study Group. Liver Resection versus Radiofrequency Ablation plus Transcatheter Arterial Chemoembolization in Cirrhotic Patients with Solitary Large Hepatocellular Carcinoma. *J Vasc Interv Radiol*. 2017 Jul 19.
- Ponziani FR, Spinelli I, Rinninella E, Cerrito L, **Saviano A** et al. Reverse time-dependent effect of alphafetoprotein and disease control on survival of patients with Barcelona Clinic Liver Cancer stage C hepatocellular carcinoma. *World J Hepatol*. 2017 Dec 28; 9(36):1322-1331.
- **Saviano A**, Baumert TF. Mortality from liver cirrhosis and HCC in the DAA era: success in viral control is darkened by raise of metabolic disease. *Hepatobiliary Surgery and Nutrition*, 2019 (in press)
- Turon-Lagot V, **Saviano A**, Schuster C, Verrier ÉR. Hepatitis D virus: viral cycle and new therapeutic approaches. *Virologie*. 2019 Jun 1;23(3):149-159. French.
- Hamdane N, Jühling F, Crouchet E, El Saghire H, Thumann C, Oudot MA, Bandiera S, **Saviano A** et al. HCV-Induced Epigenetic Changes Associated With Liver Cancer Risk Persist After Sustained Virologic Response. *Gastroenterology*. 2019 Jun;156(8):2313-2329.e7.
- **Saviano A**, Roehlen N, Virzi A, Roca Suarez AA, Hoshida Y, Lupberger J, Baumert TF. Stromal and Immune Drivers of Hepatocarcinogenesis. *Hepatocellular carcinoma*. Springer 2019.
- Aizarani N, **Saviano A**, Sagar, Mailly L, Durand S, Herman JS, Pessaux P et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature*. 2019 Aug;572(7768):199-204.
- Wakabayashi T, Ouhmich F, Gonzalez-Cabrera C, Felli E, **Saviano A**, Agnus V, Savadjiev P, Baumert TF, Pessaux P, Marescaux J, Gallix B. Radiomics in hepatocellular carcinoma: a quantitative review. *Hepatol Int*. 2019 Aug 31.
- **Felli E\***, **Saviano A\***, Tripon S, Baumert T, Pessaux P, Baumert TF. Letter to the Editor: Abdominal Surgery in Idiopathic Noncirrhotic Portal Hypertension: Is Preemptive TIPS Reducing Postoperative Complications? *Hepatology* 2019, Oct 9. **\*These authors contributed equally to this work.**

Selection of Posters and  
oral presentations

- AASLD - The Liver Meeting 2011: Pompili M, de Matthaeis N, **Saviano A**, Brunello F, Cantamessa A, Giorgio A, et al., editors. Assessment of the risk of local recurrence after radiofrequency ablation for small ( $\leq 3.0$  cm) hepatocellular carcinoma (hcc): analysis of a multicentre series of 363 patients. *Hepatology* 2011; 54 (4), 1360A-1361A (**selected as poster of distinction**).
- AASLD - The Liver Meeting 2011: de Matthaeis N, Pompili M, **Saviano A**, Brunello F, Cantamessa A, Giorgio A, et al., editors. Radiofrequency ablation for single small ( $\leq 3.0$  cm) hepatocellular carcinoma (hcc): long term overall and disease free survival rates in a multicentre series of 363 patients. Preliminary results. *Hepatology* 2011; 54 (4), 1401A
- EASL - International Liver Congress 2012: Pompili M, **Saviano A**, de Matthaeis N, Giulianti F, Ardito F, Federico B, et al. 107 Liver resection and radiofrequency ablation in the treatment of single small ( $\leq 3$  cm) hepatocellular carcinoma (hcc): long term effectiveness in a multicenter series. *Journal of Hepatology*. 2012;56:s47-s48.
- AASLD - The Liver Meeting 2013: Rinninella E, Ponziani FR, Biolato M, Siciliano M, D'Aversa F, Laterza L, **Saviano A**, et al., editors. Personalized Therapy in a Multidisciplinary Team Improves Survival in Patient with HCC: The Hepatocatt Experience. *Hepatology* 2013 ; 58, 1215A-1216A
- EASL – International Liver Congress 2013: de Matthaeis N, **Saviano A**, De Sio I, Francica G, Brunello F, Cantamessa A, et al. 254 Single hepatocellular carcinoma smaller than 2 cm: are percutaneous ethanol injection and radiofrequency ablation equally effective? *Journal of Hepatology*. 2013(58):S108-S9.
- AASLD - The Liver Meeting 2015: **Saviano A**, Iezzi R, Giulianti F, Salvatore L, Mele C, Posa A, et al. Surgical resection versus radiofrequency ablation plus drug-eluting beads transcatheter arterial chemoembolization for the treatment of solitary large hepatocellular carcinoma. A single center experience. *Hepatology* 2015 ; 62, 427A
- AASLD – The Liver Meeting 2017: Ponziani FR, Spinelli I, Rinninella E, **Saviano A**, Annicchiarico BE, Siciliano M, Riccardi L, et al. Prognostic Factors for Survival of BCLC-C Stage Hepatocellular Carcinoma Patients According to Previous Treatments: A Real-Life Experience. *Gastroenterology* 2017 152 (5), S1180
- AASLD – The Liver Meeting 2017: Jühling F, Hamdane N, Thumann C, Oudot M, El Saghire H, Bandiera S, **Saviano A** et al. Hepatitis C virus-induced epigenetic changes associated with hepatocarcinogenesis persist following viral cure. *Hepatology* 2017; 66, 120A-120A
- EASL – International Liver Congress 2018: **Saviano A**, Riccardi L, Zocco MA, Baumert T, Rapaccini GL, Pompili M, Gasbarrini A. Contrast-enhanced ultrasound for non-invasive diagnosis of hepatocellular carcinoma: a comparison between CEUS LI-RADS and ESCULAP criteria in a large high-risk cohort of patients (**selected as best poster liver tumor section**). *Journal of Hepatology*. 2018; 68, S417-S418
- EASL – International Liver Congress 2018: Duong FHT, **Saviano A**, Robinet E, Dentinger M, Bou About G, Thumann C et al. *In vivo* knockdown of Plk1 using a novel breakable mesoporous silica particle siRNA carrier reduces tumor growth in an orthotopic xenograft mouse model of hepatocellular carcinoma. *Journal of Hepatology*. 2018; 68, S662
- AASLD - The Liver Meeting 2018: Aizarani N, **Saviano A**, Mailly L, Sagar, Durand S, Pessaux P et al. A Human Liver Cell Atlas Based on Single-Cell RNA-Sequencing Uncovers Cell Heterogeneity in Human Liver and Reveals Cancer-Specific and Environment-Specific Gene Signatures. *Hepatology* 2018; 68, 876A-876A (**selected as poster of distinction**).
- AASLD - The Liver Meeting 2018: Juhling F, Verrier E, **Saviano A**, El Saghire H, Heydmann L, Fauvelle C et al., Single Cell RNA-Seq Analysis of Hepatitis B Virus Host-Interactions in Hepatocellular Carcinoma. *Hepatology* 2018; 68, 322A-322A
- AASLD - The Liver Meeting 2018: Hamadane N, Juhling F, Thumann C, Li S, El Saghire H, Crouchet E, Bandiera S, **Saviano A** et al. Targeting HCV-Induced Epigenetic Reprogramming for HCC Chemoprevention Post Cure. *Hepatology* 2018; 68, 125A-125A

## Liver pathophysiology at the single cell level: characterization of cellular heterogeneity and identification of novel therapeutic targets for chronic liver diseases and hepatocellular carcinoma

### Résumé

Le carcinome hépatocellulaire (CHC) est parmi les principales causes de mortalité dans le monde et les traitements disponibles sont insuffisants. Ceci est dû à la connaissance limitée de la complexité biologique et du microenvironnement hépatiques en situation normale et pathologique. Pour répondre à ces besoins, nous avons développé un protocole de séquençage d'ARN sur cellule unique (scRNA-seq) à partir de tissus primaires de foie humain. Nous avons assemblé un atlas de cellules du foie humain et comparé le profil scRNA-seq du foie normal au profil du CHC. L'atlas a révélé l'hétérogénéité au sein des principales populations de cellules hépatiques, la zonation transcriptomique des cellules endothéliales et l'existence de progéniteurs épithéliaux dans le foie adulte capable de se différencier à la fois en cholangiocytes et en hépatocytes. L'analyse par scRNA-seq du CHC a dévoilé l'hétérogénéité marquée de cette tumeur, les modifications de son microenvironnement cellule par cellule et les interactions entre les cellules tumorales et le virus de l'hépatite B en découvrant des voies et des facteurs moteurs de la cancérogenèse hépatique jusqu'à présent inconnus.

**Mots clés :** Carcinome hépatocellulaire, hépatocarcinogénèse, séquençage ARN sur cellule unique, virus de l'hépatite B, microenvironnement, interactions virus-hôte.

### Résumé en anglais

Hepatocellular carcinoma (HCC) is a leading cause of death worldwide and the current treatments are unsatisfactory. One reason is the limited knowledge on the complexity and microenvironment of healthy and diseased liver. To address this gap, we have developed a single cell RNA sequencing (scRNA-seq) pipeline for primary human liver tissues. We assembled an atlas of human liver cells and compared the scRNA-seq profile of normal liver and HCC. The atlas revealed the heterogeneity within the main populations of liver cells, the transcriptomic zonation of endothelial cells and the existence of an epithelial progenitor in the adult liver capable of differentiating into both cholangiocytes and hepatocytes. ScRNA-seq analysis uncovered the marked cell heterogeneity of HCC, its microenvironment changes at single-cell level and the interactions between tumor cells and hepatitis B virus discovering previously unknown pathways and drivers of hepatocarcinogenesis.

**Keywords:** hepatocellular carcinoma, hepatocarcinogenesis, single-cell RNA-sequencing, hepatitis B virus, microenvironment, virus-host interactions.