Influence of Progenitor-Derived Regeneration Markers on Hepatitis C Virus-Related Cirrhosis Outcome (ANRS CO12 CirVir Cohort)

Dominique Wendum, Richard Layese, Nathalie Ganne-Carrié, Valérie Bourcier, Fatiha Merabtene, Carole Cagnot, Emmanuel Sauce, Nathalie Barget, Pierre Bedossa, Benoit Terris, et al.

Influence of progenitor-derived regeneration markers on HCV-related cirrhosis outcome (ANRS CO12 CirVir cohort).

Dominique Wendum (1,2), Richard Layese (3,4), Nathalie Ganne-Carrié (5,6,7), Valérie Bourcier (5), Fatiha Merabtene (2,8), Carole Cagnot (9), Emmanuel Sauce (10), Nathalie Barget (10), Pierre Bedossa (11), Benoit Terris (12), Janick Selves (13), Paulette Bioulac-Sage (14,15,16), Nathalie Sturm (17), Christophe Sattonnet (18), Pierre Nahon (5,6,7), Françoise Roudot- Thoraval (3,4), Marianne Ziol (6, 7, 10) for the ANRS CO 12 CirVir group.

(1) APHP, Hôpital St Antoine, Anatomie Pathologique, Paris, France;

(2) Sorbonne Universités, UPMC Univ Paris 06, INSERM UMRS_938 Centre de Recherche St Antoine (CRSA), F-75012 Paris, France;

(3) APHP, Hôpital Henri Mondor, Unité de recherche clinique, Créteil, France,

(4) Université Paris Est (UPEC), IMRB, A-TVB DHU, CEpiA EA 7376 (Clinical Epidemiology and Aging Unit), Créteil, France

(5) APHP, Hôpital Jean Verdier, Service d'hépatologie, Bondy, France

(6) Université Paris 13, Sorbonne Paris-Cité, Bobigny, France

(7) INSERM UMR 1162, Paris, France

(8) Sorbonne Universités, UMS 30 LUMIC plateforme d'histomorphologie St Antoine, F-75012 Paris, France

(9) ANRS, Paris

(10) APHP, Hôpital Jean Verdier, Anatomie Pathologique, Bondy, France

(11) APHP, Hôpital Beaujon, Département de Pathologie, Clichy, France

(12) APHP, Hôpital Cochin, Anatomie Pathologique, Paris, France

(13) IUCT-Oncopole Toulouse, Departement d'Anatomie Pathologique, Toulouse, France

(14) Pathology Department, Pellegrin Hospital, CHU Bordeaux, F-33076 Bordeaux, France

(15) Inserm, UMR1053 Bordeaux Research in Translational Oncology BaRITOn, F-33000 Bordeaux France

(16) Univ. Bordeaux, UMR1053 Bordeaux Research in Translational Oncology BaRITOn, F-33000 Bordeaux France

(17) CHU Grenoble, Département de Pathologie, Grenoble, France

(18) Selarl Diag, Nice, France

For the ANRS CO12 CirVir study:

Pierre Nahon[1], Patrick Marcellin[2], Dominique Guyader[3], Stanislas Pol[4], Hélène Fontaine[4], Dominique Larrey[5], Victor De Lédinghen[6], Denis Ouzan[7], Fabien Zoulim[8], Dominique Roulot[9], Albert Tran[10], Jean-Pierre Bronowicki[11], Jean-Pierre Zarski[12], Vincent Leroy[12], Ghassan Riachi[13], Paul Calès[14], Jean-Marie Péron[15], Laurent Alric[16], Marc Bourlière[17], Philippe Mathurin[18], Sebastien Dharancy[18], Jean-Frédéric Blanc[19], Armand Abergel[20], Lawrence Serfaty[21], Ariane Mallat[22], Jean-Didier Grangé[23], Pierre Attali[24], Yannick Bacq[25], Claire Wartelle[26], Thông Dao[27], Yves Benhamou[28], Christophe Pilette[29], Christine Silvain[30], Christos Christidis[31], Dominique Capron[32], Gérard Thiefin[33], Sophie Hillaire[34], Vincent Di Martino[35].

[1]AP-HP, Hôpital Jean Verdier, Service d'Hépatologie, Bondy, Université Paris 13, Bobigny et INSERM U1162, Université Paris 5, Paris; [2]AP-HP, Hôpital Beaujon, Service d'Hépatologie, Clichy; [3]CHU Pontchaillou, Service d'Hépatologie, Rennes; [4]AP-HP, Hôpital Cochin, Département d'Hépatologie et INSERM UMS20, Institut Pasteur, Université Paris Descartes, Paris; [5]Hôpital Saint Eloi, Service d'Hépatologie, Montpellier; [6]Hôpital Haut-Lévêque, Service d'Hépatologie, Bordeaux; [7]Institut Arnaud Tzanck, Service d'Hépatologie, St Laurent du Var; [8]Hôpital Hôtel Dieu, Service d'Hépatologie, Lyon; [9]AP-HP, Hôpital Avicenne, Service d'Hépatologie, Bobigny; [10]CHU de Nice, Service d'Hépatologie, et INSERM U1065, Université de Nice-Sophia-Antipolis, Nice; [11]Hôpital Brabois, Service d'Hépatologie, Vandoeuvre-les-Nancy; [12]Hôpital Michallon, Service d'Hépatologie, Grenoble; [13]Hôpital Charles-Nicolle, Service d'Hépatologie, Rouen; [14]CHU d'Angers, Service d'Hépatologie, Angers; [15]Hôpital Purpan, Service d'Hépatologie, Toulouse; [16]CHU Toulouse, Service de Médecine Interne-Pôle Digestif UMR 152, Toulouse; [17]Hôpital Saint Joseph, Service d'Hépatologie, Marseille; [18]Hôpital Claude Huriez, Service d'Hépatologie, Lille; [19]Hôpital St André, Service d'Hépatologie, Bordeaux; [20]Hôpital Hôtel Dieu, Service d'Hépatologie, Clermont-Ferrand; [21]AP-HP, Hôpital Saint-Antoine, Service d'Hépatologie, Paris; [22]AP-HP, Hôpital Henri Mondor, Service d'Hépatologie, Créteil; [23]AP-HP, Hôpital Tenon, Service d'Hépatologie, Paris; [24]AP-HP, Hôpital Paul Brousse, Service d'Hépatologie, Villejuif; [25]Hôpital Trousseau, Unité d'Hépatologie, CHRU de Tours; [26]Hôpital d'Aix-En-Provence, Service d'Hépatologie, Aix-En-Provence; [27]Hôpital de la Côte de Nacre, Service d'Hépatologie, Caen; [28]AP-HP, Groupe Hospitalier de La Pitié-Salpêtrière, Service d'Hépatologie, Paris; [29]CHU Le Mans, Service d'Hépatologie, Le Mans; [30]CHU de Poitiers, Service d'Hépatologie, Poitiers; [31]Institut Mutualiste Montsouris, Service d'Hépatologie, Paris; [32]Hôpital Amiens Nord, Service d'Hépatologie, Amiens; [33]Hôpital Robert Debré, Service d'Hépatologie, Reims; [34]Hôpital Foch, Service d'Hépatologie, Suresnes; [35]Hôpital Jean Minjoz, Service d'Hépatologie, Besançon. FRANCE.

Corresponding author contact information:

Marianne Ziol, Service d'anatomie pathologique, Hopital Jean Verdier APHP, Av du 14 juillet, 93143 Bondy, France

Phone:33148026678

Fax:33148026877

E mail: marianne.ziol@aphp.fr

Grant Support: ANRS (France REcherche Nord & sud Sida-HIV Hépatites-FRENSH).

Key-words: EpCAM; Keratin7; intermediate hepatobiliary cells, hepatocellular carcinoma, liver decompensation

Acknowledgements: This work is dedicated to the memory of Prof. Jean-Claude Trinchet.

ABSTRACT

Background and aims: Progenitor-derived regeneration give rise to the aberrant expression of biliary markers such as cytokeratin 7 (K7) and epithelial cell adhesion molecule (EpCAM) in hepatocytes.. We aimed to describe the expression of these molecules in patients with compensated HCV-related cirrhosis and to investigate its potential influence on cirrhosis complications.

Methods: Among patients with Child-Pugh A uncomplicated HCV-related cirrhosis enrolled in the prospective ANRS CO12 CirVir cohort, we selected individuals with a liver biopsy collected within 2 years before inclusion in the study. K7 and EpCAM immunostaining identified intermediate hepatobiliary cells. The influence of the biliary markers expression in hepatocytes on decompensation events and the occurrence of hepatocellular carcinoma (HCC) was studied using a multivariate Cox proportional hazards regression model.

Results: Among the 337 patients eligible for the study (men, 67%; median age, 52 y), 198 (58.8%) patients had biopsies with K7 positive hepatocytes including extensive staining in 40 (11.9%) and 203 patients had EpCAM positive hepatocytes (60.6%). During follow-up (median: 54.2 months), 47 patients (14%) experienced a decompensation event, and HCC was diagnosed in 37 patients (11%). Extensive K7 staining was independently associated with the occurrence of a decompensation event (HR: 3.00, 95% CI: 1.30; 6.89, p=0.010). EpCAM expression was independently associated with HCC occurrence (HR: 2.37, 95% CI: 1.07; 5.23, p=0.033) along with age and a low prothrombin ratio.

Conclusion: Progenitor-derived regeneration depicted by K7 and EpCAM immunostaining of hepatocytes in liver biopsies of patients with compensated HCV-related cirrhosis marks a cirrhosis stage more prone to develop complications.

Abbreviations: K, cytokeratin; HCC, hepatocellular carcinoma; IHC, intermediate hepatobiliary cells; LB, liver biopsy; HCV, hepatitis C virus; SVR: sustained virological response

Introduction:

In Western countries, chronic infection with hepatitis C virus (HCV) is the leading viral cause of cirrhosis (1). Two large longitudinal prospective cohorts clarified the natural history of patients with HCV-related advanced liver fibrosis and reported high frequencies of life-threatening complications (2, 3). The prediction of hepatocellular carcinoma (HCC) risk and the occurrence of complications in patients with early stage cirrhosis remains an ongoing challenge, requiring improvements in current stratification risk. In the era of highly efficacious therapies for HCV (4, 5), it becomes essential to identify patients with HCV-related cirrhosis who are at higher risk of developing progressive disease and HCC taking into account virosuppression. The study of cirrhosis at the cellular and molecular level might predict major liver-related endpoints, including HCC development, liver disease progression, liver transplantation or death, beyond clinically available prognostic indicators but could also highlight specific points of dysregulation that can guide therapeutic or preventive interventions. The liver possesses a marked capacity to regenerate following injury, and this regeneration usually occurs due to the division of mature hepatocytes. However, in severe and/or chronic liver diseases, regeneration can result from progenitor cells located along the intralobular bile ductule and along the canals of Hering (6). This “abnormal” regeneration gives rise to a specific "transitional" compartment, known as intermediate hepatobiliary cells, which are characterized as small hepatocytes that express cytokeratin 7 (K7) —a marker of the biliary epithelium— while mature hepatocytes are negative for K7(7). This hepatocyte regeneration pathway expand the stem cell compartment, and makes these cells more prone to carcinogenesis as suggested by experimental models (8, 9) and in vitro studies (10, 11).

In a retrospective monocentric study of 150 patients with HCV-related cirrhosis and active viral replication, we described that the aberrant expression of K7 biliary markers in hepatocytes was related to HCC occurrence (12). The influence of EpCAM, another molecule expressed in early hepatocytes derived from the progenitor cell compartment (13), lost in as few as two cell divisions (14) on cirrhosis prognosis, had not been investigated.

We aimed to describe the progenitor-derived regeneration, depicted by K7 and EpCAM expression in hepatocytes, in a prospective multicentric cohort comprising patients with compensated HCV-related cirrhosis, and to investigate the potential influence of this aberrant expression on decompensation events and on the occurrence of hepatocellular carcinoma.

METHODS

This is an ancillary study derived from the CirVir cohort (3) with its own funding from the ANRS. This protocol was approved by the Ethics Committee (Comité de Protection des Personnes, Aulnay-sous-Bois, France) and conformed to the ethical guidelines of the 1975 Declaration of Helsinki. All the patients gave written informed consent before participating in the study. The full CirVir protocol is available on the ANRS website (http://anrs.fr).

Patients

Among patients with biopsy-proven compensated (Child-Pugh A) and uncomplicated virus-related cirrhosis who were enrolled in the French multicenter ANRS CO12 CirVir cohort between March 2006 and July 2012 and who were prospectively followed up for the development of HCC (3), only patients with HCV (i.e., patients positive for HCV antibodies at any level of viral replication) either with or without HBV coinfection were selected for this ancillary study. The second inclusion criterion for this ancillary study was an available liver biopsy (LB) performed within 2 years before study enrollment. The location, date, and identification number of the LB that confirmed a diagnosis of cirrhosis was recorded upon enrollment. When LBs were performed within 2 years before study enrollment, pathologists at the different clinical centers were asked to send unstained slides or blocks of the liver biopsy sample for the purposes of the present study.

Baseline data available in CirVir eCRF

Assessment of ANRS CO12 CirVir cohort included the following clinical and biological parameters: gender, age, body mass index, diabetes, arterial hypertension, dyslipidemia, metabolic syndrome, past and ongoing alcohol or tobacco consumption, prothrombin ratio, bilirubin serum level, albumin serum level, platelets count, liver

enzymes (AST, ALT, GGT), AFP serum level, virological data (ARN VHC, HIV co infection, current or previous antiviral treatments), esophageal varices, and a Doppler ultrasonography to rule out complications at baseline, including HCC. Patient information was recorded in a computerized database by a clinical research associate who was specifically dedicated to the ANRS CO12 CirVir cohort at each center.

Follow-up and HCC diagnosis

Follow-up was scheduled according to French guidelines (Haute Autorité de Santé). Doppler US examination was performed every 6 months for HCC screening. In cases where focal liver lesions were detected by US, a diagnostic procedure using contrast-enhanced imaging (CT-scan or MRI), serum alpha-fetoprotein (AFP) assay and/or guided biopsy were performed according to 2005 AASLD guidelines (15), updated in 2011 (16). An HCC diagnosis was established either by histological examination performed by an experienced pathologist or by using probability-based non-invasive criteria (mainly dynamic imaging showing early arterial hypervascularization and portal washout) according to the different periods of time (before and after 2011). When an HCC diagnosis was established, treatment was decided using a multidisciplinary approach according to either AASLD or EASL guidelines for HCC (15-17) Reports of imaging techniques showing liver focal lesions were secondarily reviewed by the two senior hepatologists from institution 1 (VB and PN).

All events occurring during follow-up (liver-related or not) were recorded based on information obtained from the medical files of patients from each center. In particular, all episodes and the severity of liver decompensation encompassing ascites, hepatic encephalopathy and gastro-intestinal bleeding were described; these episodes were managed according to international recommendations and outcome (18, 19). All extra-hepatic events occurring during follow-up were also recorded. Likely causes of death

were established. Patients who underwent liver transplantation were censored (or considered to be loss of follow-up) at the date of transplantation for the analysis. All treatments that included antiviral therapy were recorded at inclusion, and any modification during follow-up was noted. According to the time of enrollment, patients received varying antiviral treatments approved by either national and/or international guidelines as described earlier (20), and the subsequent sustained virological response (SVR) was recorded when achieved. The main end points of the study were the occurrence of HCC and/or an initial first decompensation event. The time frame was defined as the time from the date of inclusion until the first decompensation and/or HCC detection. Follow-up for this ancillary study ended at the date of either the first decompensation event or a diagnosis of HCC, at the date of liver transplantation, at the date of death due to causes other than HCC and decompensation or at the last recorded visit through December 31, 2015.

Histology and immunohistochemistry

LB specimens were previously formalin-fixed and paraffin embedded. Dewaxed and rehydrated paraffin sections (3-micrometers thick) were stained with hematein-eosin (HE) and picrosirius red. Immunostaining was performed using a standardized immunoperoxidase method with an anti-K7 monoclonal antibody (DakoCytomation®, anti-K7 antibody OV-TL 12/30, 1/50 dilution with prior antigen retrieval, EDTA [pH 9]) and an anti-EpCAM monoclonal antibody (Novocastra-Leica® NCL-MOC31, 1/25 dilution with prior antigen retrieval, 10 mM citrate buffer [pH 6]) with an automated staining system (Leica® Novolink-polymer detection system, Dako® autostainer).

HE stained liver biopsy samples were assessed for steatosis (grade 0:<5%; grade 1: 5 to 33%; grade 2: 33 to 66%; grade 3: >66%), activity according to the METAVIR scoring system and ballooning (grade 0: absent; grade 1: size of ballooned

hepatocytes less than 2-fold that of a normal hepatocyte; grade 2: size more than 2-fold) as described in the SAF score (21). The presence of an alcoholic hepatitis, which was defined by the association of liver cell ballooning, polymorphonuclear infiltrates and Mallory bodies (22), was also reported. We considered that a steatohepatitis was associated with the viral-related cirrhosis when we observed both steatosis involving more than 5% of hepatocytes and a clear ballooning grade of 1 or 2. As previously described (12), hepatocytes were considered to express K7 when at least 2 foci of more than 5 polygonal-shaped hepatocytes measuring from 6 to 40 microns in diameter were detected. K7 staining was considered as extensive when one or several cirrhotic nodules showed positive staining in more than 30% of hepatocytes. Hepatocytes were considered positive for EpCAM when a focus of more than 3 positive contiguous hepatocytes was detected. HE and immunostained sections were all independently assessed by the 2 pathologists (DW, MZ). In cases of disagreement, sections were re-examined using a multipipe microscope and a consensus was reached, which was used for the final analysis.

To further characterize K7 and EpCAM expression, we also studied the liver biopsies of 39 patients who had sequential biopsies (the first one at the time of HCV-related cirrhosis diagnosis and the second at the time of HCC diagnosis), of 28 patients with HCV chronic hepatitis under the cirrhosis stage, of 16 obese patients with normal liver, and 12 patients with steatohepatitis without cirrhosis.

Statistical analyses

Descriptive results are presented as the means ± standard deviation (SD) or medians [interquartile range (IQR)] for continuous variables and as numbers (percentages) for categorical variables. Baseline characteristics were compared among groups of patients with or without biliary markers using either Student's t-test or Mann-

Whitney/Kruskall-Wallis test for continuous variables. Categorical variables were compared using the χ^2 test or Fisher's exact test when necessary. Curves of the cumulative incidence of HCC and decompensation as well as the survival curves were constructed using the Kaplan-Meier method and compared according to the expression of biliary markers using the log-rank test. The influence of all potential features (including the expression of biliary markers) on the occurrence of HCC, decompensation events, and overall and specific mortality was assessed using a Cox proportional hazards regression model in the univariate analyses. In the multivariate analysis, the final model was built in several steps: 1) all parameters with a p-value <0.20 in univariate analysis were included in a model, 2) from this model, we excluded one by one each non significant parameter with a p-value>0.05 level to finally obtain a final multivariate model only containing significant factors independently associated with the studied outcome. All the predictors were assessed based on their baseline levels with the exception of SVR, which was included as a time-varying covariate in all Cox models to account for the fact that SVR may occur at different time points. Fixed SVR values were used for patients who never experienced SVR (SVR=0) and patients who achieved SVR at the time of their enrollment (SVR=1). Non-SVR patients at enrollment who achieved SVR status during follow-up were switched from non-SVR to SVR status. The final treatment that led to undetectable HCV RNA was set as the time point for switching the SVR values from 0 to 1. No re-infection or relapse as defined by detectable HCV RNA in a patient who previously achieved SVR was observed during the follow-up.

All statistical analyses were performed using Stata 13.0 (StataCorp, College Station, TX). A p-value ≤ 0.05 was considered statistically significant.

RESULTS

Patient selection:

As shown in figure 1, among the 1354 patients with compensated HCV-related cirrhosis who were prospectively included in the ANRS CO12 CirVir study from 2006 to 2012 (3), 853 were excluded because the LB was performed more than two years before patient's enrollment, and 157 were further excluded because the LB sample had not been sent by the pathology laboratory. Seven patients were excluded because of failed K7 and EpCAM immunostainings. Analyses were performed in the remaining 337 patients. The selected patients with an available LB for immunostaining had baseline characteristics similar to those observed in the whole cohort except for age (median age was 52 in the selected patients vs 55 in the whole cohort, $p<0.001$).

Liver biopsy analysis and baseline characteristics (tables 1 and 2):

The LBs consisted of 1 to 12 fragments (mean 3) with a mean total size of 18 mm (range: 3-44). Cirrhosis was confirmed in all the biopsy samples. Necro-inflammatory activity was moderate or severe in 156 patients (46.3%). Steatohepatitis was observed in 141 patients (41.8%) and alcoholic hepatitis in 7 (2.1%). No histological cholestasis was observed. Focal K7 staining was detected in 158 patients (46.9%), and extensive K7 staining (Fig. 2) in 40 patients (11.9%). EpCAM was localized to the membrane of periseptal hepatocytes in 203 patients (60.6%). The Kappa interobserver agreement for the K7 and EpCAM assessment was 0.74 (95% CI: 0.67; 0.81) and 0.74 (95% CI: 0.65; 0.82), respectively. The extent of EpCAM staining was variable from one focus to extensive staining of several rows of hepatocytes at the borders of the cirrhotic nodules (Fig. 2). K7 (either focal or extensive) and EpCAM staining in the hepatocytes were highly correlated ($p<0.001$), and both stains were often localized to the same cirrhotic nodules and in the same areas of the cirrhotic nodules. However, as shown in

the serial sections of the biopsies in figure 2, both stains were not strictly colocalized at the cellular level since EpCAM-stained hepatocytes tended to surround a K7-stained hepatocyte, and some EpCAM hepatocytes present within the first rows of some cirrhotic nodules lost their EpCAM expression but then acquired CK7. On the other hand, K7- or EpCAM-positive hepatocytes could also be detected independently in a non-contiguous manner. Moreover, K7 foci frequently colocalized with periportal lymphocytic infiltrate, while EpCAM staining did not. K7 and EpCAM were frequently observed in steatohepatitis foci, but also independently of steatohepatitis foci. The detection of K7- and EpCAM-stained hepatocytes in LBs according to the baseline clinical, biological and histological characteristics of patients is described in tables 1 and 2, respectively. The presence of K7 and the presence of EpCAM were significantly related to higher AFP levels, a lower platelet count, higher GGT levels, lower albumin levels and a reduced prothrombin ratio. However, only K7 staining was associated with higher ALT levels. Regarding clinical characteristics, K7 and EpCAM were not related to age, alcohol intake (past or at inclusion), alcoholic hepatitis, diabetes or BMI. Only female gender was slightly associated with K7 staining. Regarding histological criteria, concomitant K7 and EpCAM expression was strongly associated with steatohepatitis, but only CK7 was strongly related to the activity grade. Characteristics of patients according to the presence of a histological steatohepatitis are given in supplementary Table 3. Patients with steatohepatitis were younger, with metabolic syndrome, higher BMI and ALT and AST levels.

In 39 patients (13 from this cohort study) with sequential biopsies (mean time between biopsies =5 years), EpCAM and K7 status (negative or positive) were constant in 31 (79%) and 28 patients (72%), respectively, while either disappearance or emergence of these markers occurred in other cases. No K7 nor EpCAM was observed in the 16

normal liver, in the 28 HCV chronic hepatitis with F1,2 and 3 METAVIR fibrosis stage, and in 11 out of 12 non alcoholic steatohepatitis with F1,2, and 3 FLIP fibrosis stages.

Influence of histological characteristics on liver decompensation (table 3)

During follow-up (median 54.2 months [IQR: 35.2-69.3]), 47 patients (14.0%) experienced at least one liver decompensation (median time of occurrence, 33.2 months [IQR: 14.0-56.5]). Among them, 39 had no SVR at the time of decompensation and 8 had obtained SVR.

The clinical, biological and histological baseline characteristics associated with the occurrence of liver decompensation are described in table 3. Based on the $p<0.20$ cutoff from the univariate analysis, the biological and histological variables potentially related to the occurrence of liver decompensation were albumin level ≤ 35 g/L, total bilirubin > 17 µmol/L, platelet count < 100×10^3 /mm^3, GGT > 2ULN, prothrombin ratio ≤ 80%, higher AFP levels, diabetes, esophageal varices, absence of SVR, histological alcoholic hepatitis, and expression of K7 in the LB. EpCAM expression was not associated with a higher risk of liver decompensation. Based on the multivariate analysis, total bilirubin level >17 µm/L, albumin level ≤35 g/L, absence of SVR, and K7 staining (extensive: HR: 3.00, 95% CI: 1.30; 6.89, p=0.010) remained independently related to the occurrence of a liver decompensation event. Seven patients who achieved SVR experienced a decompensation event. As shown in figure 3, the incidence of decompensation during a 3-year follow-up was 19.6% in patients with extensive K7 staining compared to 3.9% in patients without K7 staining. When non SVR patients were selected (n=304, 28 decompensation events), multivariate analysis showed results similar to those observed in the whole cohort, with albumin, total bilirubin and K7 being related to decompensation (supplementary table 4). In patients

with SVR at endpoint (n=187, 8 decompensation events), no clinical, biological or histological variables were related to decompensation.

Influence of histological characteristics on HCC occurrence (table 4)

HCC was diagnosed in 37 patients (11.0%), and the median time of occurrence was 33.4 months [IQR: 25.8-45.5]). Among them, 28 HCC occurred in patients without SVR, and 9 in patients who achieved SVR. Baseline biological and histological variables with $p<0.20$ in the univariate analysis that were potentially related to the occurrence of HCC were age, a lower prothrombin ratio, lower platelet counts, lower albumin levels, higher AFP levels, higher GGT levels, the absence of SVR and the presence of EpCAM-stained hepatocytes in the LB. K7 expression was not associated with a higher risk of HCC. The multivariate analysis showed that age, prothrombin ratio and EpCAM-positive hepatocytes in the LB (HR: 2.37; 95% CI: 1.07; 5.23, $p=0.033$) remained independently related to HCC occurrence. As shown in figure 3, the detection of EpCAM-stained hepatocytes upon study enrollment was significantly associated with HCC occurrence. After 50 months of follow-up, the cumulative incidence of HCC reached 15.1% in EpCAM-positive patients vs 7.4% in EpCAM-negative patients. We further investigated the influence of K7 on HCC risk in patients with and without SVR during follow-up and found that the presence of K7 staining (focal and/or extensive) was significantly related to HCC occurrence in patients without SVR ($p=0.048$), but it was not related to HCC in patients with SVR (supplementary table 5).

Influence of histological characteristics on survival (supplementary figures 1 and 2):

During follow-up, 36 patients died, and 19 deaths were liver-related (8 decompensation or liver failure, 5 HCC, 1 cholangiocarcinoma, 5 variceal bleeding).

Five patients died from severe non-spontaneous bacterial peritonitis (SBP) infection. Eight patients have been transplanted. Based on variables with $p<0.20$ in the univariate analysis, overall survival was potentially related to age, albumin level, AFP levels, ALT levels, steatosis, the absence of steatohepatitis, the presence of alcoholic hepatitis, the absence of SVR and K7 staining (focal or extensive). The multivariate analysis showed that K7 staining (focal: HR 3.49, 95% CI: 1.47; 8.29, p=0.005), absence of SVR (HR: 2.63, 95% CI: 1.10; 6.25, p=0.029), and absence of steatohepatitis (HR 2.70, 95% CI: 1.30; 5.88, p=0.009) were independent variables related to overall death. These results are detailed in supplementary table 1. Liver-related deaths including HCC and bacterial infection were potentially related to older age, BMI, AFP levels, platelet counts, albumin level, prothrombin ratio, the absence of steatohepatitis, histological alcoholic hepatitis, the absence of SVR, EpCAM and K7 staining based on the univariate analysis ($p<0.20$). The multivariate analysis showed that obesity, the absence of SVR, lower platelet counts, the absence of steatohepatitis and K7 staining (either focal or extensive) were independently related to liver-related deaths (Supplementary table 2).

Discussion:

In this large prospective national cohort of patients with compensated HCV-related cirrhosis, the expression of progenitor-derived regeneration markers such as K7 and EpCAM in hepatocytes was observed in 58.8% and 60.6% of 337 LBs, respectively. Normal liver, HCV chronic hepatitis and non alcoholic steatohepatitis without cirrhosis were always negative for these markers. Expression of both markers was related to the presence of steatohepatitis and liver dysfunction biological markers. We found that an extensive K7 staining, which was observed in 40 patients (11.9%), was independently associated with the occurrence of a decompensation event (HR: 3.00 [95% CI: 1.30; 6.89], p=0.010), while EpCAM expression was independently associated with HCC occurrence (HR: 2.37 [95% CI: 1.07; 5.23], p=0.033).

We previously described the aberrant expression of biliary markers and reported a correlation between K7 expression and HCC occurrence. To the best of our knowledge, no other longitudinal study investigating phenotypic changes of hepatocytes and their influence on cirrhosis complications has been published. The goal of the present study was to validate our previous observation on a larger and prospective cohort with LBs processed at different centers. Surprisingly, the expression of K7, when assessed according to the same criteria, could predict decompensation but not HCC occurrence. The explanation for this apparent discrepancy may rely on the major changes of antiviral treatment efficacy. In our previous study, all 150 subjects presented active ongoing HCV infection, and patients who cleared HCV infection during follow-up were excluded from that study. In the present study, which reflects the dramatic changes in the treatment efficacy and tolerance in cirrhotic patients, 191 of 331 patients (57.7%) achieved SVR. Given that a sustained virological response significantly decreases HCC occurrence (4, 20, 23, 24),

it is not surprising that the baseline biomarkers observed in the first study were no longer predictive in the present prospective cohort. In accordance with the first study, the presence of K7 staining (focal or extensive) remained significantly related to HCC occurrence in patients without SVR (p=0.048). In the previous study, we did not investigate the influence of K7 on decompensation and did not perform EpCAM immunostaining. The combined results of both studies suggest that the presence of K7 foci at baseline reflects an ongoing active progenitor-derived regeneration that favors carcinogenesis in patients with persistent viral infection but not in patients who cleared infection. Whether clearing infection leads to a reversion of the aberrant K7 phenotype has not been investigated by sequential biopsies, but D'Ambrosio R et al.(25) showed reversal of several parameters such as zonation and ductular proliferation in a sequential pre- and post-treatment paired biopsy study. Moreover, the presence of K7 foci was related to ALT levels and to the activity score. We could hypothesize that periportal lymphocytic infiltrate, induces, after hepatocytes apoptosis, a signal for abnormal regeneration from progenitor cells in the same zone, or a transdifferentiation of mature hepatocytes. The aberrant phenotype induced by necrotic and inflammatory activity of viral hepatitis would be reversed by viral eradication, and could consequently lowers the ability of K7 expression to predict a higher risk of HCC occurrence. However, sequential biopsies, that were not performed in this study, would have been required to prove this hypothesis

A striking finding of our study was the strong relationship between extensively K7-stained cirrhotic nodules present in 11.9% of the patients' liver biopsies and the occurrence of decompensation events and liver-related death. This link was independent of SVR, and 7 patients who achieved SVR experienced a decompensation event. This suggests that extensive K7 staining in compensated

cirrhosis could reflect an advanced and irreversible stage with a definitively altered phenotype of some cirrhotic nodules that are prone to decompensation. In addition, neither the duration of HCV infection (the date of the HCV contamination was available in 160 patients) nor the delay from HCV infection diagnosis to study enrollment (available in 331 patients) were related to K7 staining. Therefore, this altered phenotype is related to the prognosis independently of the duration of HCV infection. Of note, extensive K7 staining was not related to bilirubin level (table 1) or histological cholestasis, the latter of which is known to induce a “metaplastic” response in hepatocyte as K7 expression is thought to protect hepatocytes from bile salt toxicity (26, 27). Given the location of K7 staining far from cholangiocytes in cases of extensive staining, we could hypothesize that changes in K7 expression could result from the transdifferentiation of mature hepatocytes into intermediate hepatocytes as described in experimental models (28, 29). Therefore, an early cirrhosis stage reflecting abnormal regeneration would be represented by intermediate hepatocytes associating with cholangiocytes and expressing EpCAM and K7, and a later stage would additionally involve biliary transdifferentiation of mature hepatocytes with K7 extensive staining at a distance from the borders of the cirrhotic nodules. On the other hand, biliary transdifferentiation of hepatocytes could also depend on prior EpCAM expression, since centrinodular K7 extensive staining is most often surrounded by extensive EpCAM staining.

EpCAM has been described as a useful marker to identify early hepatocytes derived from the stem and progenitor cell lineages and presents in buds that repopulate regions of parenchymal extinction in the cirrhotic liver (14). We found contiguous EpCAM-positive hepatocytes at the border of some cirrhotic nodules in 60.6% of the biopsies. Although EpCAM- and K7-positive hepatocytes were not strictly colocalized,

the expression of both proteins was observed in the same patients and frequently in the same cirrhotic nodules. This is in line with the hypothesis that cirrhotic nodules or buds are newly generated from the stem cell pathway of regeneration. The expression of both proteins were related to biological markers of liver dysfunction such as AFP levels, low platelet count and albumin, but K7 alone was related to ALT levels and histological activity score while EpCAM was not related to either of these inflammatory activity markers. This marker of abnormal regeneration would indicate a higher risk of carcinogenesis related to the proliferation of progenitor cells independent of activity score and of virological response.

In the search of a clinically pertinent physiopathological subclassification of the different stages of cirrhosis (30), efforts have focused on fibrosis qualification and quantification with significant correlations at clinical endpoints (31). However, the amount of fibrosis alone does not reflect the cause of liver dysfunction, and we highlight here that K7 extensive staining is an early phenotypic marker associated with liver dysfunction. Moreover, quantification of fibrosis is limited by sample size, and most studies excluded a large number of patients due to the small size or fragmentation of the biopsy samples (32). In our study, the LBs were small (mean size 16 mm), and we did not exclude any samples. When patients were stratified on biopsy size (under or over 16 mm), the influence of K7 and EpCAM on decompensation and HCC occurrence were unchanged.

Steatohepatitis, which is defined in HCV infected patients by the association of ballooning and steatosis involving more than 5% of hepatocytes, was observed in 41.8% of patients. This is in line with findings of other HCV cohorts, and the high frequency of steatohepatitis in individuals with HCV-related cirrhosis compared to the expected frequency in the general population is probably explained by an accelerated

fibrosis progression in HCV-infected individuals with steatohepatitis (33-35). However, we showed ~~here~~ that steatohepatitis associated with HCV-related cirrhosis did not have any influence on liver decompensation nor HCC occurrence ~~risk~~. Surprisingly, absence of steatohepatitis was significantly related to overall and to liver related death. We could hypothesize that steatohepatitis is an initial contributor to the progression to cirrhosis, that alone is not involved in the further development of cirrhosis complications. Moreover, to explain the paradoxical survival "advantage" of steatohepatitis, we could hypothesize that steatohepatitis, which was observed in younger patients, reflect an earlier stage of cirrhosis. It was not related to statin or Metformin treatment.

In conclusion, we showed that phenotypic changes of hepatocytes at the cellular level, which were observed in the biopsies of patients with compensated HCV cirrhosis, are related to a higher risk of developing cirrhosis complications and liver-related death independent of SVR. Thus, aberrant phenotypes that reflect abnormal progenitor-derived regeneration, and/or transdifferentiation of hepatocytes in biopsies could represent a practical read out of the cirrhosis stage at risk for complications in patients who do not achieve SVR, eventhough the interobserver agreement of K7 and EpCAM was not perfect (Kappa 0.74, fair to good). Its prognostic value in patients with SVR remains to be determined. The demonstration of the link between phenotypic alterations of hepatocytes, exclusively observed in cirrhosis and occurrence of complications could also open the way for the search of peripheral blood markers, of circulating EpCAM or CK7 positive hepatocyte, and for anti EpCAM vaccination (36).

References

1. El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma. Gastroenterology 2012;142:1264-1273.e1261.
2. Di Bisceglie AM, Shiffman ML, Everson GT, Lindsay KL, Everhart JE, Wright EC, Lee WM, et al. Prolonged therapy of advanced chronic hepatitis C with low-dose peginterferon. N Engl J Med 2008;359:2429-2441.
3. Trinchet JC, Bourcier V, Chaffaut C, Ait Ahmed M, Allam S, Marcellin P, Guyader D, et al. Complications and competing risks of death in compensated viral cirrhosis (ANRS CO12 CirVir prospective cohort). Hepatology 2015;62:737-750.
4. van der Meer AJ, Feld JJ, Hofer H, Almasio PL, Calvaruso V, Fernandez-Rodriguez CM, Aleman S, et al. Risk of cirrhosis-related complications in patients with advanced fibrosis following hepatitis C virus eradication. J Hepatol 2016.
5. Ganne-Carrie N, Layese R, Bourcier V, Cagnot C, Marcellin P, Guyader D, Pol S, et al. Nomogram for individualized prediction of hepatocellular carcinoma occurrence in hepatitis C virus cirrhosis (ANRS CO12 CirVir). Hepatology 2016;64:1136-1147.
6. Roskams T. Different types of liver progenitor cells and their niches. J Hepatol 2006;45:1-4.
7. Roskams TA, Theise ND, Balabaud C, Bhagat G, Bhathal PS, Bioulac-Sage P, Brunt EM, et al. Nomenclature of the finer branches of the biliary tree: canals, ductules, and ductular reactions in human livers. Hepatology 2004;39:1739-1745.
8. Thorgeirsson SS, Factor VM, Snyderwine EG. Transgenic mouse models in carcinogenesis research and testing. Toxicol Lett 2000;112-113:553-555.
9. Libbrecht L, Meerman L, Kuipers F, Roskams T, Desmet V, Jansen P. Liver pathology and hepatocarcinogenesis in a long-term mouse model of erythropoietic protoporphyria. J Pathol 2003;199:191-200.
10. Ma S, Chan KW, Hu L, Lee TK, Wo JY, Ng IO, Zheng BJ, et al. Identification and characterization of tumorigenic liver cancer stem/progenitor cells. Gastroenterology 2007;132:2542-2556.
11. Cerec V, Glaise D, Garnier D, Morosan S, Turlin B, Drenou B, Gripon P, et al. Transdifferentiation of hepatocyte-like cells from the human hepatoma HepaRG cell line through bipotent progenitor. Hepatology 2007;45:957-967.
12. Ziol M, Nault JC, Aout M, Barget N, Tepper M, Martin A, Trinchet JC, et al. Intermediate hepatobiliary cells predict an increased risk of hepatocarcinogenesis in patients with hepatitis C virus-related cirrhosis. Gastroenterology 2010;139:335-343.e332.
13. Turner R, Lozoya O, Wang Y, Cardinale V, Gaudio E, Alpini G, Mendel G, et al. Human hepatic stem cell and maturational liver lineage biology. Hepatology 2011;53:1035-1045.
14. Stueck AE, Wanless IR. Hepatocyte buds derived from progenitor cells repopulate regions of parenchymal extinction in human cirrhosis. Hepatology 2015;61:1696-1707.
15. Bruix J, Sherman M. Management of hepatocellular carcinoma. Hepatology 2005;42:1208-1236.
16. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatology 2011;53:1020-1022.

17. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2012;56:908-943.
18. EASL clinical practice guidelines on the management of ascites, spontaneous bacterial peritonitis, and hepatorenal syndrome in cirrhosis. J Hepatol 2010;53:397-417.
19. de Franchis R. Revising consensus in portal hypertension: report of the Baveno V consensus workshop on methodology of diagnosis and therapy in portal hypertension. J Hepatol 2010;53:762-768.
20. Nahon P, Bourcier V, Layese R, Audureau E, Cagnot C, Marcellin P, Guyader D, et al. Eradication of Hepatitis C Virus Infection in Patients With Cirrhosis Reduces Risk of Liver and Non-Liver Complications. Gastroenterology 2016.
21. Bedossa P, Poitou C, Veyrie N, Bouillot JL, Basdevant A, Paradis V, Tordjman J, et al. Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. Hepatology 2012;56:1751-1759.
22. Altamirano J, Miquel R, Katoonizadeh A, Abraldes JG, Duarte-Rojo A, Louvet A, Augustin S, et al. A histologic scoring system for prognosis of patients with alcoholic hepatitis. Gastroenterology 2014;146:1231-1239.e1231-1236.
23. van der Meer AJ, Veldt BJ, Feld JJ, Wedemeyer H, Dufour JF, Lammert F, Duarte-Rojo A, et al. Association between sustained virological response and all-cause mortality among patients with chronic hepatitis C and advanced hepatic fibrosis. Jama 2012;308:2584-2593.
24. Morgan RL, Baack B, Smith BD, Yartel A, Pitasi M, Falck-Ytter Y. Eradication of hepatitis C virus infection and the development of hepatocellular carcinoma: a meta-analysis of observational studies. Ann Intern Med 2013;158:329-337.
25. D'Ambrosio R, Aghemo A, Rumi MG, Ronchi G, Donato MF, Paradis V, Colombo M, et al. A morphometric and immunohistochemical study to assess the benefit of a sustained virological response in hepatitis C virus patients with cirrhosis. Hepatology 2012;56:532-543.
26. Van Eyken P, Sciot R, Desmet VJ. A cytokeratin immunohistochemical study of cholestatic liver disease: evidence that hepatocytes can express 'bile duct-type' cytokeratins. Histopathology 1989;15:125-135.
27. Yabushita K, Yamamoto K, Ibuki N, Okano N, Matsumura S, Okamoto R, Shimada N, et al. Aberrant expression of cytokeratin 7 as a histological marker of progression in primary biliary cirrhosis. Liver 2001;21:50-55.
28. Yanger K, Zong Y, Maggs LR, Shapira SN, Maddipati R, Aiello NM, Thung SN, et al. Robust cellular reprogramming occurs spontaneously during liver regeneration. Genes Dev 2013;27:719-724.
29. Sekiya S, Suzuki A. Hepatocytes, rather than cholangiocytes, can be the major source of primitive ductules in the chronically injured mouse liver. Am J Pathol 2014;184:1468-1478.
30. Garcia-Tsao G, Friedman S, Iredale J, Pinzani M. Now there are many (stages) where before there was one: In search of a pathophysiological classification of cirrhosis. Hepatology 2010;51:1445-1449.
31. Nagula S, Jain D, Groszmann RJ, Garcia-Tsao G. Histological-hemodynamic correlation in cirrhosis-a histological classification of the severity of cirrhosis. J Hepatol 2006;44:111-117.
32. Tsochatzis E, Bruno S, Isgro G, Hall A, Theocharidou E, Manousou P, Dhillon AP, et al. Collagen proportionate area is superior to other histological methods for sub-classifying cirrhosis and determining prognosis. J Hepatol 2014;60:948-954.

33. Everhart JE, Lok AS, Kim HY, Morgan TR, Lindsay KL, Chung RT, Bonkovsky HL, et al. Weight-related effects on disease progression in the hepatitis C antiviral long-term treatment against cirrhosis trial. Gastroenterology 2009;137:549-557.
34. Ghany MG, Lok AS, Everhart JE, Everson GT, Lee WM, Curto TM, Wright EC, et al. Predicting clinical and histologic outcomes based on standard laboratory tests in advanced chronic hepatitis C. Gastroenterology 2010;138:136-146.
35. Dyal HK, Aguilar M, Bhuket T, Liu B, Holt EW, Torres S, Cheung R, et al. Concurrent Obesity, Diabetes, and Steatosis Increase Risk of Advanced Fibrosis Among HCV Patients: A Systematic Review. Dig Dis Sci 2015;60:2813-2824.
36 Choi YJ, Park SJ, Park YS, Park HS, Yang KM, Heo K. EpCAM peptide-primed dendritic cell vaccination confers significant anti-tumor immunity in hepatocellular carcinoma cells. PLoS One. 2018 Jan 3;13(1):e0190638.

Figure legends

Figure 1: Patient's selection flow chart

Figure 2: Representative microphotograph of a biopsy showing cirrhotic nodules positive for K7 and EpCAM staining. At least one nodule showed extensive K7 staining of hepatocytes (black arrow, a). Ductular reaction (stars, a) or isolated cholangiocytes were not taken into account. More than 30% of the hepatocytes presented cytoplasmic and/or submembranous K7 immunostaining (b, arrows). EpCAM stained the membrane of contiguous hepatocytes (c and d, arrows) preferentially at the border of the cirrhotic nodules when detected. EpCAM also stained some ductules (stars, c). EpCAM- and K7-stained intermediate hepatocytes were observed in the same cirrhotic nodules as shown in this example with serial sections (white arrows in a and c). However, at the cellular level, both stainings (b, d) were not stackable, and EpCAM-positive areas encompassed K7 stained hepatocytes

Figure 3: Cumulative incidence of liver decompensation (higher panel) and HCC occurrence (lower panel) during follow-up according to K7 (A) and EpCAM (B) staining at enrollment.

A, top panel: Extensive staining by K7 (red curve) was highly associated with liver decompensation. At the 3-year follow-up, the cumulative incidence of liver decompensation was 19.6% for patients with extensive K7 staining versus 3.9% for patients without K7 staining. B, top panel: EpCAM expression was not associated with the occurrence of liver decompensation.

A, bottom panel: K7 expression was not related to HCC occurrence. B, bottom panel: the HCC incidence rate was significantly higher in patients with EpCAM staining compared to patients without EpCAM staining (p=0.039). At the 3-year follow-up, the

incidence of HCC was 9.8% in patients with EpCAM staining vs 2.7% in patients without EpCAM staining.

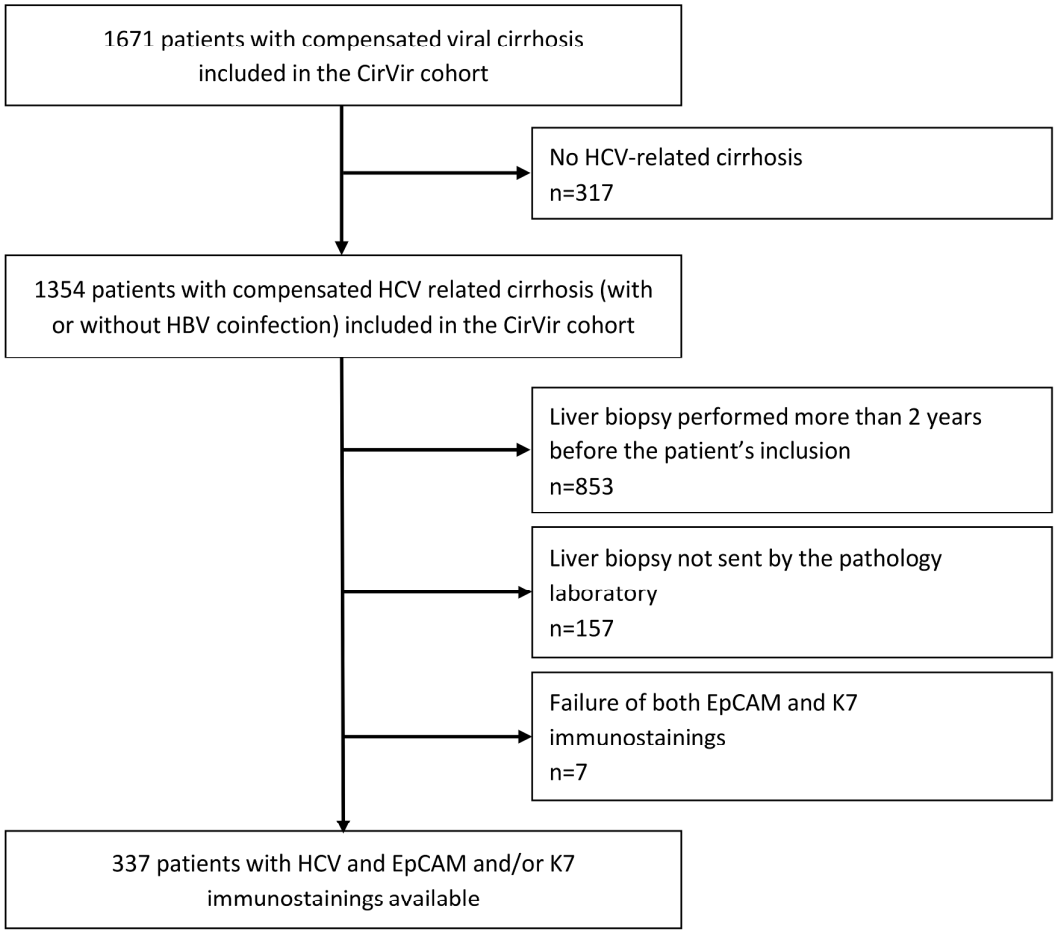


Figure 1

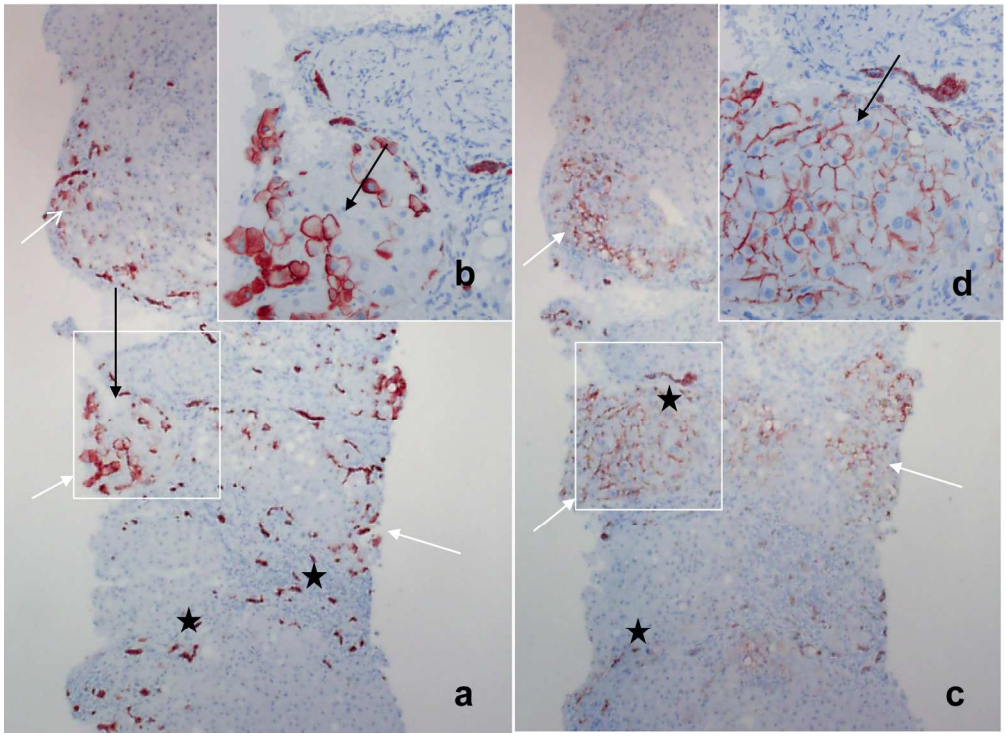
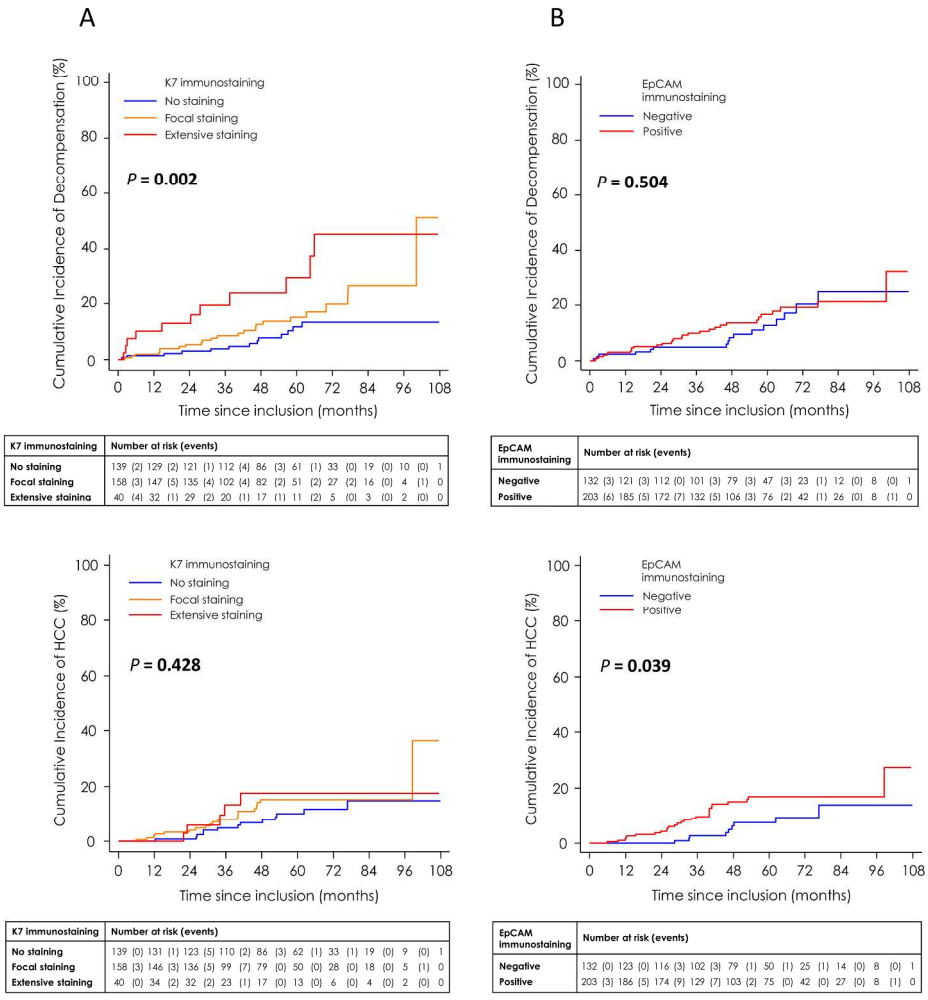


Figure 2: Representative microphotograph of a biopsy showing cirrhotic nodules positive for K7 and EpCAM staining.

177x129mm (300 x 300 DPI)



K7 immunostaining	Number at risk (events)
No staining	139 (2) 129 (2) 121 (1) 112 (4) 86 (3) 61 (1) 33 (0) 19 (0) 10 (0) 1
Focal staining	158 (3) 147 (5) 135 (4) 102 (4) 82 (2) 51 (2) 27 (2) 16 (0) 4 (1) 0
Extensive staining	40 (4) 32 (1) 29 (2) 20 (1) 17 (1) 11 (2) 5 (0) 3 (0) 2 (0) 0

EpCAM immunostaining	Number at risk (events)
Negative	132 (3) 121 (3) 112 (0) 101 (3) 79 (3) 47 (3) 23 (1) 12 (0) 8 (0) 1
Positive	203 (6) 185 (5) 172 (7) 132 (5) 106 (3) 76 (2) 42 (1) 26 (0) 8 (1) 0

K7 immunostaining	Number at risk (events)
No staining	139 (0) 131 (1) 123 (5) 110 (2) 86 (3) 62 (1) 33 (1) 19 (0) 9 (0) 1
Focal staining	158 (3) 146 (3) 136 (5) 99 (7) 79 (0) 50 (0) 28 (0) 18 (0) 5 (1) 0
Extensive staining	40 (0) 34 (2) 32 (2) 23 (1) 17 (0) 13 (0) 6 (0) 4 (0) 2 (0) 0

EpCAM immunostaining	Number at risk (events)
Negative	132 (0) 123 (0) 116 (3) 102 (3) 79 (1) 50 (1) 25 (1) 14 (0) 8 (0) 1
Positive	203 (3) 186 (5) 174 (9) 129 (7) 103 (2) 75 (0) 42 (0) 27 (0) 8 (1) 0

Figure 3: Cumulative incidence of liver decompensation (higher panel) and HCC occurrence (lower panel) during follow-up according to K7 (A) and EpCAM (B) staining at enrollment.

274x290mm (300 x 300 DPI)

Table 1: Patient characteristics at enrollment according to K7 immunostaining status

Variables	All patients n=337	Patients with K7 assessment, n	No K7 staining n=140	Focal K7 staining n=157	Extensive K7 staining n=40	P–value
Male gender	**226 (67.1)**	**337**	**104 (74.8)**	**99 (62.7)**	**23 (57.5)**	**0.033**
Age (years)	52.0 [46.5 – 59.3]	337	52.4 [45.6 – 60.1]	51.7 [47.1-58.3]	49.1 [45.7 – 56.9]	0.512
Associated liver disease						
Excessive alcohol intake	73 (21.7)	337	30 (21.6)	32 (20.3)	11 (27.5)	0.610
Metabolic syndrome	80 (23.7)	337	32 (23)	39 (24.7)	9 (22.5)	0.927
HDV	2 (0.6)	337	1 (0.7)	0	1 (2.5)	-
HBV	7 (2.1)	337	2 (1.4)	3 (1.9)	2 (5.0)	0.704
Past alcohol intake > 30g/day	106 (32.7)	324	43 (31.4)	50 (33.1)	13 (36.1)	0.857
Alcohol intake at inclusion > 10g/day	27 (8.5)	316	11 (8.7)	12 (8.0)	4 (10.3)	0.884
BMI (kg/m^2)	25.8 [23.2 – 28.7]	302	25.9 [23.2 – 29.4]	25.9 [23.5 – 28.2]	25.6 [22.3 – 28.1]	0.714
Diabetes	52 (15.4)	337	18 (13.0)	27 (17.1)	7 (17.5)	0.571
Esophageal varices	65 (25.3)	257	33 (30.6)	24 (20.3)	8 (25.8)	0.210
AFP level (ng/mL)	**7.0 [4.2 – 14.0]**	**332**	**5.0 [3.2 – 8.3]**	**9.7 [5.6 – 17.0]**	**13.5 [6.0 – 35.0]**	**<0.001**
Plt ($10^3/mm^3$)	**126 [96 – 167]**	**335**	**136 [102 – 174]**	**124 [94 – 173]**	**110 [81 – 136]**	**0.003**
ALT (UI/L)	**84 [46 – 126]**	**337**	**70 [39 – 113]**	**94 [50 – 143]**	**84 [52 – 113]**	**0.020**
GGT (UI/L)	**103 [61 – 206]**	**337**	**85 [48 – 158]**	**106 [62 – 244]**	**145 [95 – 224]**	**<0.001**
Albumin (g/L)	**41.0 [37.5 – 44.2]**	**334**	**42.0 [38.0 – 45.0]**	**40.2 [37.1 – 44.2]**	**38.8 [35.8 – 43.0]**	**0.014**
Prothrombin ratio (%)	88 [79 – 96]	327	90 [82 – 98]	87 [78 – 94]	86 [79 – 93]	0.050
Total bilirubin (µmol/L)	11.0 [8.0 – 16.0]	337	10.0 [8.0 – 15.0]	12.0 [8.0 – 18.0]	11.0 [9.0 – 15.5]	0.427
Positive HCV viral load	267 (79.2)	337	105 (75.5)	128 (81.0)	34 (85.0)	0.322
HCV genotype		325				0.069
1	210 (64.6)		93 (69.4)	94 (61.8)	23 (59.0)	
2	10 (3.1)		5 (3.7)	3 (2.0)	2 (5.1)	
3	54 (16.6)		21 (15.7)	30 (19.7)	3 (7.7)	
4	49 (15.1)		14 (10.5)	24 (15.8)	11 (28.2)	
5	1 (0.3)		1 (0.7)	0	0	
6	1 (0.3)		0	1 (0.7)	0	
Past or ongoing antiviral treatment	282 (83.7)	337	114 (82.0)	136 (86.1)	32 (80.0)	0.511
IFN-based treatment before biopsy	118 (41.8)	282	54 (47.0)	51 (38.9)	13 (36.1)	0.337
Steatosis > 5%	248 (73.6)	337	94 (67.6)	121 (76.6)	33 (82.5)	0.086
Activity		**337**				**0.001**
0	10 (3.0)		8 (5.8)	2 (1.3)	0	
1	171 (50.7)		85 (61.2)	72 (45.5)	14 (35.0)	
2	146 (43.3)		44 (31.6)	78 (49.4)	24 (60.0)	
3	10 (3.0)		2 (1.4)	6 (3.8)	2 (5.0)	
Ballooning	**149 (44.2)**	**337**	**41 (29.5)**	**81 (51.3)**	**27 (67.5)**	**<0.001**
Steatohepatitis	**141 (41.8)**	**337**	**38 (27.3)**	**77 (48.7)**	**26 (65.0)**	**<0.001**
Alcoholic hepatitis	7 (2.1)	334	1 (0.7)	3 (1.9)	3 (7.5)	0.055
EpCAM positive	**203 (60.6)**	**335**	**60 (43.2)**	**110 (69.6)**	**33 (86.8)**	**<0.001**

K7: cytokeratin 7; HDV: hepatitis delta virus; HBV: hepatitis B Virus; BMI: body mass index; AFP: alpha-fetoprotein; Plt: platelet; ALT: alanine aminotransferase; BMI: body mass index; GGT: gamma-glutamyl transferase

Table 2: Patient characteristics at enrollment according to EpCAM immunostaining status

Variables	All patients n=337	Patients with EpCAM assessment, n	EpCAM negative n=132	EpCAM positive n= 203	P–value
Male gender	226 (67.1)	335	88 (66.7)	137 (67.5)	0.876
Age (years)	52.0 [46.5 – 59.3]	335	51.5 [46.3 – 59.6]	52.2 [46.5 – 59.1]	0.863
Associated liver disease					
Excessive alcohol intake	73 (21.7)	335	25 (18.9)	47 (23.2)	0.359
Metabolic syndrome	80 (23.7)	335	32 (24.2)	47 (23.2)	0.818
HDV	2 (0.6)	335	2 (1.5)	0	-
HBV	7 (2.1)	335	4 (3.0)	3 (1.5)	0.440
Past alcohol intake > 30g/day	106 (32.7)	322	36 (28.6)	69 (35.2)	0.215
Alcohol intake at inclusion > 10g/day	27 (8.5)	314	10 (8.4)	17 (8.7)	0.923
BMI (kg/m^2)	25.8 [23.2 – 28.7]	300	25.4 [22.9 – 28.4]	26.0 [23.5 – 29.0]	0.253
Diabetes	52 (15.4)	335	19 (14.4)	33 (16.3)	0.646
Esophageal varices	65 (25.3)	255	23 (25.6)	41 (24.9)	0.901
AFP level (ng/mL)	**7.0 [4.2 – 14.0]**	**330**	**5.1 [3.7 – 10.0]**	**9.0 [5.0 – 17.3]**	**<0.001**
Plt ($10^3/mm^3$)	**126 [96 – 167]**	**333**	**140 [100 – 177]**	**122 [94 – 158]**	**0.003**
ALT (UI/L)	84 [46 – 126]	335	81 [43 – 119]	86 [51 – 131]	0.142
GGT (UI/L)	**103 [61 – 206]**	**335**	**93 [56 – 162]**	**115 [66 – 229]**	**0.022**
Albumin (g/L)	**41.0 [37.5 – 44.2]**	**332**	**42.0 [38 – 45]**	**40 [37 – 44]**	**0.006**
Prothrombin ratio (%)	**88 [79 – 96]**	**325**	**90 [83 – 100]**	**87 [78 – 94]**	**0.014**
Total bilirubin (µmol /L)	11.0 [8.0 – 16.0]	335	11.0 [8.0 – 14.5]	11.0 [8.0 – 17.0]	0.129
HCV genotype		323			0.086
1	210 (64.6)		91 (72.2)	118 (59.9)	
2	10 (3.1)		3 (2.4)	7 (3.5)	
3	54 (16.6)		19 (15.1)	35 (17.8)	
4	49 (15.1)		12 (9.5)	36 (18.3)	
5	1 (0.3)		1 (0.8)	0	
6	1 (0.3)		0	1 (0.5)	
Past or ongoing antiviral treatment	282 (83.7)	335	112 (84.8)	168 (82.8)	0.614
IFN-based treatment before biopsy	118	280	45 (42.5)	73 (42.0)	0.935
Steatosis > 5%	**248 (73.6)**	**335**	**86 (65.1)**	**161 (79.3)**	**0.004**
Activity		335			0.242
0	10 (3.0)		5 (3.8)	5 (2.5)	
1	171 (50.7)		68 (51.5)	101 (49.8)	
2	146 (43.3)		58 (43.9)	88 (43.3)	
3	10 (3.0)		1 (0.8)	9 (4.4)	
Ballooning	**149 (44.2)**	**335**	**42 (31.8)**	**106 (52.2)**	**<0.001**
Steatohepatitis	**141 (41.8)**	**335**	**39 (29.5)**	**101 (49.7)**	**<0.001**
Alcoholic hepatitis	7 (2.1)	334	2 (1.5)	5 (2.5)	0.708
K7		**335**			**<0.001**
No staining	139 (41.2)		79 (59.8)	60 (29.6)	
Focal staining	158 (46.9)		48 (36.4)	110 (54.2)	
Extensive staining	40 (11.9)		5 (3.8)	33 (16.2)	

BMI: body mass index; HDV: hepatitis Delta Virus; HBV: hepatitis B Virus; AFP: alpha-fetoprotein; Plt: platelet; ALT: alanine aminotransferase; BMI: body mass index; GGT: gamma-glutamyl transferase; K7: cytokeratin 7

Table 3: Results of the univariate and multivariate cox regression analyses for the risk of occurrence of liver decompensation during follow-up

	Univariate analysis			Multivariate analysis		
	HR	95% CI	P-value	HR	95% CI	P-value
Male gender	0.99	[0.54 ; 1.79]	0.960			
Age (years)	0.99	[0.96 ; 1.02]	0.667			
Past alcohol intake > 30g/day	1.20	[0.64 ; 2.24]	0.565			
BMI (kg/m²)						
<25	Ref					
[25 ; 30[	0.59	[0.29 ; 1.21]	0.151			
≥30	1.02	[0.49 ; 2.12]	0.958			
Diabetes	1.95	[1.03 ; 3.71]	0.040			
EV	1.69	[0.87 ; 3.31]	0.122			
AFP level (ng/mL)*	1.76	[0.94 ; 3.28]	0.075			
Plt (10^3/mm^3)			0.010			
<100	3.40	[1.51 ; 7.64]	0.003			
[100 ; 150]	1.97	[0.85 ; 4.56]	0.115			
>150	Ref					
ALT (UI/L)			0.271			
≤N	Ref					
]N ; 2N]	2.14	[0.85 ; 5.44]	0.108			
>2N	1.67	[0.68 ; 4.09]	0.262			
GGT (UI/L)			0.030			
≤N	Ref					
]N ; 2N]	3.09	[0.68 ; 13.95]	0.143			
>2N	5.33	[1.28 ; 22.24]	0.022			
Albumin ≤ 35 g/L	3.42	[1.43 ; 6.75]	<0.001	2.42	[1.20 ; 4.91]	0.014
Prothrombin ratio ≤ 80 %	1.61	[0.89 ; 2.93]	0.116			
Total bilirubin > 17 µmol/L	2.90	[1.60 ; 5.25]	<0.001	2.40	[1.31 ; 4.39]	0.004
SVR**	0.28	[0.12 ; 0.64]	0.003	0.33	[0.14 ; 0.76]	0.010
Steatosis > 5%	1.29	[0.64 ; 2.59]	0.478			
Activity			0.294			
0	Ref					
1	0.60	[0.14 ; 2.59]	0.498			
2	0.96	[0.23 ; 4.06]	0.952			
3	NA	NA	NA			
Ballooning	0.93	[0.52 ; 1.66]	0.809			
Steatohepatitis	0.77	[0.43 ; 1.39]	0.392			
Alcoholic hepatitis	5.6	[1.73 ; 18.31]	0.004			
EpCAM positive	1.23	[0.67 ; 2.26]	0.505			
K7			0.003			0.032
No staining	Ref			Ref		
Focal staining	1.85	[0.93 ; 3.65]	0.078	1.46	[0.73 ; 2.93]	0.280
Extensive staining	4.08	[1.82 ; 9.12]	0.001	3.00	[1.30 ; 6.89]	0.010

BMI: body mass index; AFP: alpha-fetoprotein; Plt: platelet; ALT: alanine aminotransferase; BMI: body mass index; GGT: gamma-glutamyl transferase; SVR: sustained virological response; NA: not applicable; K7: cytokeratin 7

* AFP is expressed in log10

** Included as a time-varying covariate

Table 4: Results of the univariate and multivariate cox regression analyses for the risk of occurrence of hepatocellular carcinoma during follow-up

	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Male gender	1.17	[0.59 ; 2.33]	0.653			
Age (years)	1.03	[0.99 ; 1.06]	0.088	1.04	[1.00 ; 1.07]	0.029
Past alcohol intake > 30g/day	1.38	[0.70 ; 2.75]	0.353			
BMI (kg/m²)			0.390			
<25	Ref					
[25 ; 30[	1.10	[0.48 ; 2.55]	0.816			
≥30	1.79	[0.74 ; 4.33]	0.194			
Diabetes	0.99	[0.41 ; 2.37]	0.980			
EV	0.71	[0.29 ; 1.72]	0.447			
AFP level (ng/mL) *	2.03	[1.04 ; 3.97]	0.039			
Plt (10^3/mm^3)			0.120			
<100	2.33	[0.99 ; 5.45]	0.051			
[100 ; 150]	1.40	[0.58 ; 3.37]	0.458			
>150	Ref					
ALT (IU/L)			0.833			
≤N	Ref					
]N ; 2N]	0.76	[0.29 ; 1.98]	0.581			
>2N	0.94	[0.41 ; 2.14]	0.884			
GGT (IU/L)			0.191			
≤N	Ref					
]N ; 2N]	2.77	[0.60 ; 12.73]	0.189			
>2N	3.65	[0.86 ; 15.46]	0.079			
Albumin ≤ 35 g/L	2.17	[0.90 ; 5.26]	0.085			
Prothrombin ratio ≤ 80 %	2.34	[1.21 ; 4.51]	0.011	2.42	[1.24 ; 4.72]	0.010
Total bilirubin > 17 µmol/L	1.14	[0.52 ; 2.49]	0.744			
SVR**	0.52	[0.24 ; 1.14]	0.100			
Steatosis > 5%	1.04	[0.49 ; 2.20]	0.923			
Activity			0.945			
0	Ref					
1	0.63	[0.15 ; 2.72]	0.539			
2	0.66	[0.15 ; 2.87]	0.580			
3	NA	NA	NA			
Ballooning	0.99	[0.52 ; 1.89]	0.977			
Steatohepatitis	0.96	[0.50 ; 1.85]	0.911			
Alcoholic hepatitis	1.83	[0.25 ; 14.42]	0.550			
EpCAM positive	2.17	[1.02 ; 4.62]	0.044	2.37	[1.07 ; 5.23]	0.033
K7			0.434			
No staining	Ref					
Focal staining	1.52	[0.75 ; 3.07]	0.249			
Extensive staining	1.69	[0.60 ; 4.75]	0.317			

BMI: body mass index; AFP: alpha- fetoprotein; Plt: platelet; ALT: alanine aminotransferase; GGT: gamma-glutamyl transpeptidase; EV: esophageal varices; SVR: sustained virological response; NA: not applicable; K7: cytokeratin 7

* AFP is expressed in log10

** Included as a time-varying covariate