



HAL
open science

Identification des nouveaux déterminants génétiques prédisposant aux carcinomes hépatocellulaires

Jie Yang

► **To cite this version:**

Jie Yang. Identification des nouveaux déterminants génétiques prédisposant aux carcinomes hépatocellulaires. Cancer. Université de Paris, 2019. Français. NNT : 2019UNIP7156 . tel-03146692

HAL Id: tel-03146692

<https://theses.hal.science/tel-03146692>

Submitted on 19 Feb 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.

Thèse de doctorat
de l'Université Sorbonne Paris Cité
Préparée à l'Université Paris Diderot
Ecole doctoral : Hématologie – Oncogénèse – Biothérapie ED561
Inserm-UMR1138, Equipe28, Génétique fonctionnelle des tumeurs solides

Identification of new genetic determinants that predispose to hepatocellular carcinomas

Par

Jie YANG

Thèse de doctorat de Oncogénèse

Dirigée par Pr. Jessica ZUCMAN-ROSSI

Dr. Jean-Charles NAULT

Présentée et soutenue publiquement à Paris le 05 Septembre 2019

Composition du Jury :

Rapporteur : Ariane MALLAT/ PUPH/ Université Paris-Est

Rapporteur : Luca VALENTI/ Professeur/ Université de Milan

Examinateur : Judith FAVIER/ DR/ Université Paris Descartes

Examinateur : Nathalie CHARNAUX/ PUPH/ Université Paris-13

Examinateur : Jean SOULIER/ PUPH/ Université Paris Diderot

Directrice de thèse : Jessica ZUCMAN-ROSSI/ PUPH/ Inserm-UMRS-1138

Co-directeur de thèse : Jean-Charles NAULT/ HDR/ Inserm-UMRS-1138

Title: Identification of new genetic determinants that predispose to hepatocellular carcinomas

Abstract: Hepatocellular carcinoma (HCC) is among the most prevalent and lethal cancers worldwide. In the setting of chronic liver disease (CLD), genetic determinants play an important role in the development of HCC. Until now, few single nucleotide polymorphisms (SNPs) have been reproducibly associated with HCC. Our project aims to identify genetic markers that predispose to HCC.

By case-control study, we investigated the association between SNPs previously identified in GWAS and HCC development taken into account the various etiology and severity of underlying CLDs. Among the 7 SNPs from GWAS of HCC, only *STAT4* rs7574865 was re-validated as genetic risk factor of HBV-related HCC in Europeans. We have also validated that *PNPLA3* rs738409 and *TM6SF2* rs58542926, two SNPs closely associated with the risk of alcoholic- and non-alcoholic fatty liver disease-related cirrhosis, were both independent risk markers of HCC in patient with ALD. The association were further confirmed in a prospective manner with a cohort of patients with alcohol-related cirrhosis. In addition, *PNPLA3* rs738409 was also associated with an increased risk of HCC on non-fibrotic livers.

In addition, we discovered that a newly identified SNP of *HSD17B13* (rs72613567), that are protective from alcoholic and non-alcoholic liver disease, protects also against the progression from ALD to ALD-related HCC. We also observed that *HSD17B13* rs72613567 could even mitigate the HCC risk in patients with at-risk G allele of *PNPLA3*.

In conclusion, *PNPLA3* rs738409, *TM6SF2* rs58542926 and *HSD17B13* rs58542926 are three genetic risk markers of ALD-related HCC. Regarding the functions of these three genes, it indicates that lipid metabolism play an important role in the liver carcinogenesis.

Keyword: single nucleotide polymorphisms, *PNPLA3*, *TM6SF2*, *HSD17B13*, cirrhosis, alcoholic liver disease, liver cancer

Titre : Identification des nouveaux déterminants génétiques prédisposant aux carcinomes hépatocellulaires

Résumé : Le carcinome hépatocellulaire (CHC) est l'un des cancers les plus répandus et les plus meurtriers au monde. Dans le contexte des maladies chroniques du foie, les déterminants génétiques jouent un rôle important dans le développement du CHC. Jusqu'à présent, peu de polymorphismes mononucléotidiques (SNPs) ont été associés de manière reproductible au CHC. Le but de mon projet est d'identifier des marqueurs génétiques prédictif de la survenue du CHC.

Par l'étude de cas/témoins, nous avons testé l'association entre une sélection de SNPs et la survenue du CHC, en fonction de l'étiologie et la gravité de l'hépatopathie. Parmi les 7 SNPs liés aux risques du CHC dans des GWAS, aucun des 6 autres SNPs n'était relié au risque de CHC, à l'exception de *STAT4* rs7574865 qui était validé comme un facteur de risque CHC développé sur hépatite B chronique chez les Européens. Nous avons également validé que *PNPLA3* rs738409 et *TM6SF2* rs58542926, deux SNPs étroitement associés au risque de cirrhose liée à la stéatose hépatique alcoolique et non alcoolique, étaient tous marqueurs de risque indépendants du CHC chez les patients avec une maladie alcoolique du foie. L'association ensuite était confirmée de manière prospective dans une cohorte de patients avec cirrhose d'origine alcoolique. En outre, *PNPLA3* rs738409 était également un facteur de risque de CHC développé sur foie non-cirrhotique (F0-F1).

Enfin, nous avons découvert qu'un nouvel SNP de *HSD17B13* (rs72613567), qui protège contre les maladies du foie alcooliques et non alcooliques, protège également contre la survenue du CHC chez les patients avec une maladie alcoolique du foie. Nous avons également observé que le polymorphisme *HSD17B13* rs72613567 pourrait atténuer le risque de CHC chez les patients présentant un allèle à risque de *PNPLA3*.

En conclusion, les polymorphismes *PNPLA3* rs738409, *TM6SF2* rs58542926 et *HSD17B13* rs58542926 sont trois marqueurs de risque génétiques du CHC développé sur maladie alcoolique du foie. Les fonctions de ces trois gènes suggèrent que le métabolisme des lipides joue un rôle important dans la carcinogenèse du foie.

Mots clefs : polymorphisme, *PNPLA3*, *TM6SF2*, *HSD17B13*, cirrhose, maladie alcoolique du foie, cancer du foie

Index

| | |
|---|-----------|
| Abbreviations | 3 |
| List of figures and tables | 6 |
| Introduction | 7 |
| I. Hepatocellular carcinoma | 8 |
| 1. Epidemiology | 8 |
| 1.1 Hepatitis B virus | 9 |
| 1.2 Hepatitis C virus | 10 |
| 1.3 Chronic alcohol intake | 11 |
| 1.4 Non-alcoholic fatty liver disease | 12 |
| 1.5 Other etiologies | 14 |
| 2. Nature history and risk factors of liver carcinogenesis | 15 |
| 2.1 Nature history of liver carcinogenesis | 15 |
| 2.2 Risk factors of HCC development | 15 |
| 3. Prevention and surveillance | 16 |
| 4. Diagnosis and treatment..... | 19 |
| 4.1 Invasive and non-invasive diagnostic strategies..... | 19 |
| 4.2 The Barcelona Clinic Liver Cancer (BCLC) staging..... | 20 |
| 4.3 Curative treatments | 21 |
| 4.4 Non-curative treatments..... | 22 |
| II. Pathogenesis and molecular mechanism..... | 23 |
| 1. Somatic genetic alterations of HCC | 23 |
| 2. Molecular classification of HCC..... | 26 |
| III. Genetic variants associated with risk of chronic liver disease and HCC..... | 27 |
| 1. HCC susceptibility and single nucleotide polymorphism | 27 |
| 2. Polymorphisms and predisposition to chronic liver diseases..... | 29 |
| 2.1 Genome-wide association study | 29 |
| 2.2 Candidate-gene study | 35 |
| 3. Genetic associations in HCC..... | 37 |
| 3.1 Genome-wide association study | 37 |
| 3.2 Candidate-gene study..... | 41 |

| | |
|---|------------|
| Results | 45 |
| Context and aim of the project | 46 |
| Article 1: “<i>PNPLA3</i> and <i>TM6SF2</i> variants as risk factors of hepatocellular carcinoma across various etiologies and severity of underlying liver diseases” | 47 |
| Article 2: “A <i>HSD17B13</i> variant protects from hepatocellular carcinoma development in alcoholic liver disease” | 68 |
| Discussion | 82 |
| 1. Interaction between environmental and genetic risk factors of HCC | 83 |
| 1.1 Validation of SNPs identified in GWAS of HCC..... | 83 |
| 1.2 Association between <i>PNPLA3/TM6SF2</i> variants and increased risk of alcohol-related HCC | 84 |
| 1.3 Biological function of <i>PNPLA3/TM6SF2</i> on lipid metabolism and liver carcinogenesis | 88 |
| 1.4 Protective effect of <i>HSD17B13</i> variant on alcohol-related HCC risk and potential mechanism of action | 93 |
| 2. Potential clinical applications | 98 |
| 2.1 Identification of patients at risk of HCC development | 98 |
| 2.2 Drug development for chemoprevention | 102 |
| Conclusions and perspectives | 104 |
| References | 105 |

Abbreviations

ABCG: ATP binding cassette subfamily G member
AFB1: Aflatoxin B1
AFP: Alpha fetoprotein
AH: Alcoholic hepatitis
ALC: Alcoholic liver cirrhosis
ALD: Alcoholic liver disease
ALT: Alanine transaminase
APC: Adenomatous polyposis coli
APOC3: Apolipoprotein C3
ASIR: Age-standardized incidence rate
AST: Aspartate aminotransferase
AUC: Area under the curve
BCLC: Barcelona Clinic Liver Cancer
BMI: Body mass index
CAV3: Caveolin-3
CDK14: Cyclin dependent kinase 14
CHB: Chronic hepatitis B
CLD: Chronic liver disease
CT: Computed tomography
CTNNB1: Catenin (Cadherin-Associated Protein), Beta 1
DAA: Direct-acting antiviral agent
DEPDC5: DEP domain containing 5
EHMT2: Euchromatic histone-lysine-methyltransferase 2
FDFT1: Farnesyl diphosphate farnesyl transferase 1
FDX1: Ferredoxin 1
GCKR: Glucokinase regulator
GRIK1: Glutamate ionotropic receptor kainate type subunit 1
GSTM1: Glutathione S-transferase mu 1
GWAS: Genome-wide association study
HCA: Hepatocellular adenoma

HCC: Hepatocellular carcinoma
HSCs: Hepatic stellate cells
HSD: Hydroxysteroid
HSD17B13: 17-beta-hydroxysteroid dehydrogenase 13
INTS10: Integrator complex subunit 10
KIF1B: kinesin family member 1B
LR: Liver resection
LT: Liver transplantation
LXR: Liver X receptor
LYPLAL1: Lysophospholipase-like 1
MBOAT: Membrane bound O-acyl transferase
MDT: Multidisciplinary treatment
MICA: MHC class I polypeptide related sequence A
MLL: Mixed lineage leukemia protein
MTHFR: Methylene tetrahydrofolate reductase
mTOR: Mechanistic target of rapamycin
MTTP: Microsomal triglyceride transfer protein
NAFL: Nonalcoholic fatty liver
NAFLD: Non-alcoholic fatty liver disease
NAS: NAFLD activity score
NASH: Nonalcoholic steatohepatitis
NCAN: Neurocan
OR: odds ratio
PDGFR: Platelet-derived growth factor receptors
PNPLA3: Patatin-like phospholipase domain-containing 3
PTEN: Protein tyrosine phosphatase
PUFAs: Polyunsaturated fatty acids
RAF: Radiofrequency ablation
RB1: Retinoblastoma 1
RPS6KA3: Ribosomal protein S6 kinase A3
SNP: Single nucleotide polymorphism
SREBP-1c: Sterol regulatory element-binding proteins 1c
STAT4: Signal transducer and activator of transcription 4

SVR: Sustained virological response
T2D: Type 2 diabetes
TACE: Transarterial chemoembolization
TCF19: transcription factor 19
TERT: Telomerase reverse transcriptase
TG: Triglyceride
TLL1: Tolloid like 1
TM6SF2: Transmembrane 6 Superfamily Member 2
TNF: Tumor necrosis factor
TP53: Tumor protein p53
UBE2L3: Ubiquitin-conjugating enzyme E2L 3
US: Ultrasonography
VEGF: Vascular endothelial growth factor
VLDL: Very low density lipoprotein

List of figures and tables

Figures:

| | |
|---|-----|
| Figure 1 : Estimated age-standardized incidence rates of liver cancer worldwide in both sexes..... | 8 |
| Figure 2 : Modified BCLC staging system and treatment strategy | 21 |
| Figure 3 : Genetic alterations of Hepatocellular carcinoma. | 24 |
| Figure 4 : Molecular classification of Hepatocellular carcinoma..... | 26 |
| Figure 5 : The possible mechanism of <i>PNPLA3</i> involved in the pathogenesis of liver disease progression. | 91 |
| Figure 6 : The possible mechanism of <i>TM6SF2</i> involved in lipid metabolism..... | 92 |
| Figure 7 : Risk of ALD-related HCC according to the combination of at-risk G allele of <i>PNPLA3</i> rs738409 and the protective TA allele of <i>HSD17B13</i> rs72613567..... | 95 |
| Figure 8 : Proposed mechanisms by which <i>HSD17B13</i> is involved in hepatic lipid droplet homeostasis..... | 96 |
| Figure 9 : HCC risk prediction with genetic variants of <i>PNPLA3</i> and <i>TM6SF2</i> | 100 |

Tables:

| | |
|---|----|
| Table 1: Trends in the epidemiology of hepatocellular carcinoma and impact of non-alcoholic fatty liver disease according to geographic regions. | 13 |
| Table 2: Genetic association between <i>PNPLA3</i> rs738409 (G) and alcohol-related cirrhosis..... | 33 |
| Table 3: SNPs identified in GWAS of chronic liver diseases. | 34 |
| Table 4: SNPs identified in GWAS associated with the risk of HCC..... | 37 |
| Table 5: Genetic association between <i>PNPLA3</i> rs738409 (G) and HCC risk according to the type of underlying liver disease | 42 |
| Table 6: SNPs associated with HCC in candidate-gene studies or meta-analysis. | 44 |

Introduction

I. Hepatocellular carcinoma

1. Epidemiology

Primary liver cancer ranks now as the sixth commonly diagnosed cancer and the fourth leading cause of cancer death worldwide (Bray et al., 2018). More than 80% of primary liver cancer are currently hepatocellular carcinoma (HCC). Typically, HCC occurs frequently in middle-aged and older (≥ 45 years) adult and is two to four times more frequent in males than in females (Torre et al., 2015). Besides the differences in age and gender, the incidence of HCC varies among different regions. The estimated age-standardized incidence rates (ASIR) per 100,000 people are highest in less developed regions, particularly in sub-Saharan Africa (14.1), Eastern (17.7), Micronesia (15.2) and South-Eastern Asia (13.3) (Bray et al., 2018) (**Figure 1**). The mortality rates in these regions are 3- to 7-folds higher than other regions (Bertuccio et al., 2017). 50% of the newly diagnosed cases worldwide arise in China and account for more than 300,000 cancer death (Bray et al., 2018; McGuire, 2016). Even if there is a substantial decline in some Asia countries, HCC incidence continues to increase in some developed countries, like United State and Australia, making HCC a global healthy problem (Altekruse, Henley, Cucinelli, & McGlynn, 2014; Bertuccio et al., 2017; Siegel, Miller, & Jemal, 2018; Wallace, Preen, Short, Adams, & Jeffrey, 2019).

Estimated age-standardized incidence rates (World) in 2018, liver, both sexes, all ages

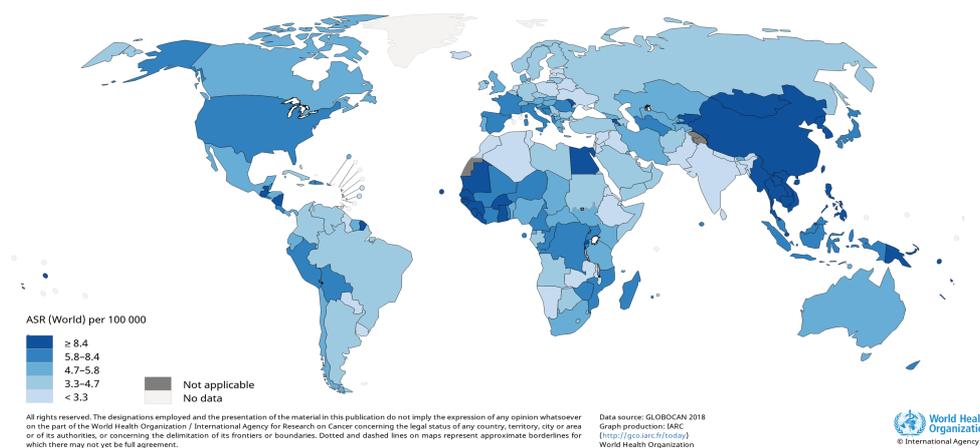


Figure 1 : Estimated age-standardized incidence rates of liver cancer worldwide in both sexes.

Data from GLOBOCAN 2018.

HCC arises mostly in patients with chronic liver disease (CLD) induced by various etiologies and 80 to 90% cases occur in the liver of cirrhosis (Colombo et al., 1991). The major causes of cirrhosis are chronic hepatitis B, chronic hepatitis C and other upcoming etiologies such as alcoholic liver disease, non-alcoholic fatty liver disease (NAFLD) due to metabolic syndrome, exposure to aflatoxin B1 and hemochromatosis. Due to economic factors and different lifestyles, the distribution of these etiologies differs among regions, which explains also the region heterogeneity of HCC incidence.

1.1 Hepatitis B virus

Hepatitis B virus (HBV) accounts for 54% of total HCC cases worldwide (El-Serag, 2012; European Association For The Study Of The, European Organisation For, & Treatment Of, 2012; Mittal & El-Serag, 2013). HBV infects almost one third of global population and the prevalence of chronic HBV carriers is 5% (350–400 million people) (McMahon, Alberts, Wainwright, Bulkow, & Lanier, 1990). The geographical distribution of chronic HBV infection is quite distinct. It is highly prevalent in the African countries (6.1%) and Western Pacific Regions (6.2%), with perinatal transmission and horizontal spread among children as major modes of infection. The prevalence is quite lower (0.3%-1.5%) in North America, Western Europe and Australia where sexual and parenteral exposure as major ways of transmission (Schweitzer, Horn, Mikolajczyk, Krause, & Ott, 2015).

According to retrospective studies, chronic HBV carriers have 5- to 100- folds increased risk of HCC compared to healthy population. In Asian population, HCC incidence is 0.6 per 100 in subjects with chronic hepatitis B without cirrhosis and 3.7 per 100 in subjects with compensated cirrhosis; the corresponding 5-year HCC cumulative incidences are 3% and 17%, respectively. While in Caucasian population, HCC incidence rate is 0.3 per 100 in subjects with chronic HBV without cirrhosis, and 2.2 per 100 in subjects with compensated cirrhosis; the corresponding 5-year HCC cumulative incidences are 1% and 10%, respectively (El-Serag, 2012). Although patients with cirrhosis have the highest risk of HCC, there is a subset of HBV-related HCC arising in the setting of non-cirrhotic liver (Chayanupatkul et al., 2017; J. D. Yang et al., 2011).

Natural history of hepatitis B, including HBV replication, HBV genotypes, HBV mutations and duration of infection, is also related with HCC risk in HBV carriers. Patients with higher levels of HBV replication, determined by positive tests for HBeAg and high levels of HBV DNA load, have a higher HCC risk (C. J. Chen, Yang, Iloeje, & Group, 2009; J. D. Chen et al., 2010; H. I. Yang et al., 2002). HBV genotypes are geographically distinct and have a different impact on clinical outcomes. Compared to genotype B, genotype C in Asia is more closely associated with cirrhosis and HCC (Kao, Chen, Lai, & Chen, 2002). While in Western Europe and North America, carriers with genotype D have a higher incidence of HCC than those with genotype A (C. L. Lin & Kao, 2015; Sanchez-Tapias, Costa, Mas, Bruguera, & Rodes, 2002). Some studies reported that mutations in HBV genome (ex. BCP T1762/A1764 and pre-core G1896A) were also associated with prevalence of HCC independent of HBV genotype and viral load (Orito et al., 2001).

Host factors (male gender, earlier age of infection or longer duration of chronic HBV infections, Asian or African ethnicity, family history of HCC, exposure to aflatoxin B1, alcohol or tobacco, and coinfection with HCV or HDV) are also linked with risk of HCC (C. J. Chen et al., 2006; Iloeje et al., 2006; Kushner, Serper, & Kaplan, 2015). From year 1980s, many countries with common HBV infection began to vaccinate all newborns. This health policy has effectively decreased the rate of HBV infection and HBV-related HCC. However, infected patients before 1980s are now reaching at age of higher risk of HCC. In addition, due to suboptimal vaccination rate and high mother-to-child transmission rate, there is still a persist burden of HBV-related HCC in the coming decades (Jutavijittum et al., 2016).

1.2 Hepatitis C virus

Hepatitis C virus is the second cause of HCC, accounting for >25% of cases. The global prevalence of HCV infection is near three percent (170 millions), mostly in East Asia, North Africa, and Middle East regions. (Szabo, Lotz, Paska, Kiss, & Schaff, 2003; Zaltron, Spinetti, Biasi, Baiguera, & Castelli, 2012). Different from HBV, people are infected mainly via blood-borne transmission. In United State and south Europe, due to intravenous drug use, abundant young people were infected with HCV during 1960s

and 1970s. In Japan, HCV was much more common than HBV because of commercial blood banks, especially in 1950s (Tanaka et al., 2002).

The majority of the infected population is asymptomatic in the acute phase (Loomba et al., 2011). However, 70% to 80% patients with HCV infection turn into the chronic infection phase, which may typically induce liver fibrosis or cirrhosis within 20 to 30 years with an estimated rate between 17% and 55%. (Deutsch, Papadopoulos, Hadziyannis, & Koskinas, 2013; Ward, 2013). It is estimated that 2-6% of all patients with HCV-related cirrhosis will develop HCC over 10 years. Compared with HCV negative people, HCV-infected patient have 17-fold increased risk of HCC (Donato et al., 2002). In endemic regions, from 50% to 70% HCC cases are due to HCV infection (Bezemer et al., 2012; Omata, Yoshida, & Shiratori, 2005).

Besides viral genotype (HCV 1b), host factors (advanced stage of fibrosis, coinfection with HBV or HIV, alcohol consumption, sustained virology response, gender, age of patients) are risk factors of HCC in HCV infected patients (Moosavy et al., 2017; Vandenbulcke et al., 2016). Until now there is not yet efficient vaccination to prevent HCV infection. With the advent of highly effective and well-tolerated direct-acting antiviral agents (DAA), more HCV patients seen in clinical practice will likely be in sustained virological response (SVR). Access to HCV treatment is increasing, but remains limited. Without vaccine prevention, limited access to treatment and the annual number of patients with cirrhosis, HCC will roughly increase and is estimated to reach the highest level in 2020s (Modin et al., 2019).

1.3 Chronic alcohol intake

Long-term excessive use of alcohol (>50-70g/day) is a well-established cause of cirrhosis and HCC. The prevalence of alcohol consumption is much higher in developed western countries, where alcohol-related HCC accounts for 32%-45% of HCC cases (T. R. Morgan, Mandayam, & Jamal, 2004).

In liver, ethanol is metabolized into acetaldehyde, a carcinogen that acts at epigenetic level, affecting methyl transfer, promoting DNA hypomethylation and disturbing the expression of tumor suppressor genes and oncogenes (Varela-Rey, Woodhoo,

Martinez-Chantar, Mato, & Lu, 2013). Moreover, oxidative stress, lipid peroxidation and chronic inflammation are also involved in alcohol-related liver carcinogenesis. The multistep process from alcoholic liver disease to HCC consists a series of histological changes from steatosis, alcoholic hepatitis, fibrosis and cirrhosis. Globally, cirrhosis occurs in 10%-15% of cases. The prevalence of HCC in patients with cirrhosis is 15%-20% with an annual incidence of 2%-4% (Mancebo et al., 2013). However, HCC could also develop in patients with high alcohol intake but without cirrhosis (Trevisani, Frigerio, Santi, Grignaschi, & Bernardi, 2010).

Alcohol consumption can increase the risk of HCC in combination with other risk factors, including chronic infection with HCV or HBV, smoking and metabolic syndrome (obesity, diabetes mellitus, dyslipidemia and arterial hypertension) (C. W. Lin et al., 2013; Marrero, Fontana, et al., 2005; Nahon & Nault, 2017). Alcohol abuse is increasing in many countries, suggesting that alcohol will continue to be a common cause of HCC throughout the world (Bruix, Gores, & Mazzaferro, 2014).

1.4 Non-alcoholic fatty liver disease

In addition to alcohol, the dramatic rise in non-alcoholic fatty liver disease (NAFLD) has also been associated with increased risk of HCC in developed countries (J. W. Park et al., 2015). In year 1980, NAFLD was firstly described as a spectrum of liver diseases with accumulation of excess triglyceride (TG) in hepatocyte in patients with metabolic syndrome without excessive alcohol consumption (Ludwig, Viggiano, McGill, & Oh, 1980). At the histological level, NAFLD ranges from to nonalcoholic fatty liver (NAFL) with simple hepatic fat accumulation, to nonalcoholic steatohepatitis (NASH) with steatosis, fibrosis and inflammation. At least 20-30% of patients with benign condition (NAFL) progress to NASH and finally cirrhosis in 10-20% of these cases (Vernon, Baranova, & Younossi, 2011).

The incidence of NAFLD and NAFLD-related HCC to continue to increase globally along with the increasing prevalence of obesity and diabetes mellitus (Pais et al., 2017). Currently, it is estimated that there are 415 million people worldwide suffering from type 2 diabetes (T2D) and there are 650 million people with obesity (Ogurtsova et al.,

2017; H. Sung et al., 2019). In addition, T2D and obesity are also independent risk factors of HCC (El-Serag, Hampel, & Javadi, 2006; Larsson & Wolk, 2007).

The Bridge study revealed that NAFLD accounted for nearly 10% and 12% of all HCC cases in Europe and North America, respectively. In contrast, the incidence of NAFLD-related HCC is only 1% to 6% in Asia, where HBV or HCV are still dominant causes of HCC (J. W. Park et al., 2015). Because of epidemiological and methodological bias, the NAFLD-related HCC incidence, together with the burden of NAFLD, is somehow underestimated (Degasperi & Colombo, 2016). In Eastern countries, due to lifestyle changes and a dramatically decline of viral infection, NAFLD is becoming a major cause of liver disease. It is likely that the increasing NAFLD incidence will translate into a significant increase in NAFLD-related HCC in the coming decades. **Table 1** summarizes the current epidemiology of HCC and the potential impact of NAFLD on HCC incidence across geographic regions in the future (Bertot & Adams, 2019).

| Region | Asia | Europe | Australia | Africa | North America |
|-----------------|---|---|---|--------|---|
| HCC incidence * | 35-99 | 10-20 | 12 | 8.9 | <5 |
| Main etiologies | HBV | HCV,alcohol | HCV,alcohol | HBV | HCV,alcohol |
| Trends | Reducing in ereas with vaccination programs | Currently increasing though likely to fall due to effective HCV treatment | Currently increasing though likely to fall due to effective HCV treatment | Stable | Currently increasing though likely to fall due to effective HCV treatment |
| Impact of NAFLD | Minor but emerging | Minor but emerging | Minor but emerging | Minor | Major and impacting incidence |

Table 1: Trends in the epidemiology of hepatocellular carcinoma and impact of non-alcoholic fatty liver disease according to geographic regions.

* Age-adjusted incidence rates (per 100 000 person), HCV: hepatitis C virus, HCC: hepatocellular carcinoma

Liver cirrhosis is also the strongest risk factor of HCC in patients with NAFLD (Kawamura et al., 2012). The 5-year incidence of HCC in NAFLD patients with cirrhosis is 11.3% (Oda, Uto, Mawatari, & Ido, 2015). However, HCC on non-cirrhotic liver is more common in patient with NAFLD than other etiologies (Sanyal, Poklepovic, Moyneur, & Barghout, 2010). Patients with HCC developed on non-cirrhotic liver are predominantly males of older age and with a higher prevalence of metabolic features (diabetes, obesity, hypertension, coronary artery disease and dyslipidemia) compared to HCC cases with cirrhosis (Guzman et al., 2008; Yasui et al., 2011). In addition,

combination with other risk factors, such as alcohol consumption, will also increase the risk of HCC (Ascha et al., 2010).

Although the incidence of HCC in patients with NAFLD-related cirrhosis is lower than that in patients with HBV- or HCV-related cirrhosis, the dramatic increase of NAFLD in general population suggests that NAFLD will be the leading cause of HCC in the future.

1.5 Other etiologies

Aflatoxin B1 (AFB1) is a metabolite produced by *Aspergillus* fungus. In sub-Saharan Africa, Eastern Asia and parts of South America, a number of staple foods, like maize and ground nuts, are frequently contaminated by AFB1 due to the humidity. The metabolites of AFB1 are mutagenic and carcinogenic. Chronic intake of AFB1 contaminated foods could induce specific DNA mutations in tissue (mutational signature number 24 characterized by a high rate of C>A mutations and a somatic TP53 mutation encoding p.Arg249Ser (Schulze et al., 2015). Abundant evidences show that AFB1 is a strong risk factor of HCC. In addition, AFB1 acts synergistically with HBV to foster HCC development. The geographical distribution of these two etiologies largely overlapped. Series of actions in endemic areas are taken to reduce the exposure of AFB1 and have already shown its efficiency to limit the occurrence of AFB1-associated HCC.

There are also some rare genetic liver diseases that could cause HCC. Classic hemochromatosis is an autosomal recessive disorder due to homozygous mutations C282Y/C282Y of *HFE*. The incidence of hemochromatosis is 1/1000 and HCC is reported in 7.5-30% of these cases (Niederrau et al., 1996; Nowak, Giger, & Krayenbuehl, 2018). Early diagnosis and phlebotomy could reverse liver damage and reduce HCC risk (Fargion et al., 1994). Wilson's disease, alpha 1 antitrypsin deficiency, familial intrahepatic cholestasis and hepatic glycogen storage are also reported to be linked with development of HCC. Finally, there are some HCCs developed in normal liver without any known etiology (Dorfman et al., 2007).

2. Nature history and risk factors of liver carcinogenesis

2.1 Nature history of liver carcinogenesis

Liver carcinogenesis is a dynamic process. HCC develops mostly in chronic liver disease, and liver carcinogenesis is strongly related to the occurrence of cirrhosis. Due to long term stimulation by viral or chemical factors (alcohol or bioactive lipids), liver frequently harbors chronic inflammation (Bishayee, 2014). Meanwhile, three processes have to be distinguished, namely, stimulation of fibroblasts, maturation and deposition of fibers, and catabolism of fibers. (Popper & Uenfriend, 1970). It could induce changes in hepatic architecture and damages of liver function. In most cases, progression from fibrosis to cirrhosis occurs after an interval of 15–20 years influenced by both genetic and environmental factors.

As mentioned, 80 to 90% of HCC arise in the setting of cirrhosis. The micro-environmental composition and organization of cirrhosis can serve as prognostic factors in HCC pathogenesis (Novikova, Khromova, & Kopnin, 2017). This environment promotes the evolutionary process of dysplasia, from low-grade to high-grade dysplasia, early HCC, progressed HCC and advanced HCC. A combination of histological characteristics (cellular density, pseudo-glandular formation, steatosis, presence of portal spaces, stromal invasion, nuclear atypia, isolated artery, etc) makes it possible to differentiate premalignant lesions from early HCC. During tumor growth, there is often portal vein thrombosis, as a feature of advanced stage which is associated with poor prognosis.

2.2 Risk factors of HCC development

Host factors, including gender, age and severity of underlying liver disease, have an impact on the progression from chronic liver disease to HCC. In all populations, the high male-to-female ratios of incidence indicate that gender is a strong risk factor of HCC development. The first explanation is the different exposure of environmental risk between males and females. Normally, males are more likely to have HBV and HCV infection, to consume more alcohol, to smoke more cigarette and to have metabolic syndrome. However, some epidemiological studies have shown that even

after adjustment for gender difference, males still have more risk of HCC than females (El-Serag, 2012; B. K. Park et al., 2007). It indicates that sex related physio-biology somehow contributes to the risk of liver carcinogenesis. In contrast, due to the use of oral contraceptive, females are more likely to develop a benign liver tumor, hepatocellular adenoma (HCA). A small fraction of patients with HCA, especially those with *CTNNB1* exon3 mutation, trend to have malignant transformation. Moreover, risk of HCC increases progressively with age (Fattovich, Stroffolini, Zagni, & Donato, 2004) and with presence of liver failure or portal hypertension that reflect the severity of cirrhosis. The risk of HCC is also higher in patients with a family history of cancer (Hassan et al., 2009).

3. Prevention and surveillance

Regarding the close association between the environmental risk factors and the long natural history of liver carcinogenesis, prevention is the most effective way to decrease HCC incidence and mortality.

HBV vaccination is the most important primary prevention aiming to eradicate HBV in uninfected population, especially in newborns. Since the first introduction of vaccine in the 1982, over 1 billion HBV vaccine have been used worldwide. Compared with children born before implementation of vaccination program, those born after have reduced HCC incidence (Chang et al., 1997; Gwack et al., 2011; Wichajarn, Kosalaraksa, & Wiangnon, 2008). Implementation of blood safety strategies, safe injection practices in drug uses and safer sex practices help to prevent transmission of HBV in uninfected adults. The effective antiviral treatment is the second prevention aiming to delay the progression of chronic hepatitis B and to reduce HCC risk in patients who are already infected. Current treatments to suppress replication of hepatitis B include pegylated interferon and nucleotide (side) (NUC) analogues (tenofovir or entecavir). A meta-analysis showed that the use of newer NUCs reduced to 2.5% in the treatment group, compared with 11.7% in the placebo group (J. J. Sung, Tsoi, Wong, Li, & Chan, 2008). However, occurrence of HCC couldn't be completely abrogated in patients treated by NUC.

There is not yet HCV vaccination but research is ongoing. Sustained virologic response (SVR) in those with chronic HCV has an impact on HCC development. Patients achieving SVR, regardless the fibrosis stage, have a relative risk of 0.24 compared to those did not (R. L. Morgan et al., 2013). The direct-acting antiviral agents (DAAs), achieving almost complete SVR, is associated with reduced risk of HCC. However, risk remains in patients with cirrhosis, highlighting the importance of HCC surveillance (Pinero et al., 2019).

Efforts to reduce the exposure of AFB1 have been made in multiple regions, by switching from maize to less AFB1 susceptible crops, public health education and government regulation. Reduced HCC risk has been observed particular in China. In Qidong City, there was an estimated overall 65% reduction in HCC after the government initiated switch from maize to rice (J. G. Chen & Kensler, 2014). But in many rural areas, AFB1 exposure is still problematic and enhanced efforts to reduce the exposure is needed.

Considering HCC risk induced by alcoholic and non-alcoholic fatty liver disease, alcohol withdrawal, healthy diet and regular exercises are recommended to prevent the progression of these chronic liver diseases. However, the effect on reducing HCC risk remains to be evaluated.

HCC is often diagnosed at advanced stages with tumor portal thrombosis or metastasis or liver failure. Advanced HCC has limited treatment options and the outcome is generally poor. However, HCC has a prolonged subclinical growth period during which surveillance of at-risk population is extremely important for an early detection and curative therapy to reduce the mortality. In a randomized controlled study, patients with HBV infection were randomly allocated in screening group and control group. HCC-related mortality was significantly lower in the screening group (83.2 per 100 000) than in the control group (131.5 per 100 000), which indicated that the surveillance reduced HCC mortality by 37% (B. H. Zhang, Yang, & Tang, 2004).

Whether a patient enters a surveillance program is mainly determined by the level of HCC risk. According to AASLD and EASL guidelines, patients with cirrhosis

irrespective of the etiology, and patients with CHB on non-cirrhotic liver and with a HCC incidence more than 0.2%/year, are recommended to be enrolled into a surveillance program. The most accurate and widely used test for HCC surveillance is ultrasonography (US). Sensitivity of the test is between 65% and 80%, and specificity is more than 90% when it is done by experienced operator (A. Singal et al., 2009). The most common used serum marker is α -fetoprotein (AFP). With cutoff of 20ng/ml, sensitivity of the test is around 60%, and specificity is 80% (Lok et al., 2010). Combined use of AFP and US increases detection rate by 6-8%, but increases also the cost and false-positive rate. Therefore, the role of AFP in HCC surveillance is still controversial (B. Zhang & Yang, 1999). HCC surveillance every six month is recommended. A 6-month interval is more effective in terms of early detection than a 12-month interval, and more cost-effective than a 3-month interval (Santi et al., 2010; Trinchet et al., 2011).

However, the surveillance is far away from perfection in clinical practice. First, current HCC surveillance is underused. In the surveillance process, the identification of cirrhosis in general patients with liver disease is already underused, leaving a lot of patients unconscious of the presence of cirrhosis. Then in patients with cirrhosis, only 20% undergo surveillance, which sometimes are inconsistent. Future interventions must target these failure points and make the surveillance process more effective. Second, the current surveillance test is suboptimal. In patients with cirrhosis, the performance of US has a sensitivity of 84% for HCC at any stage while a sensitivity of 47% for early stage HCC (Tzartzeva et al., 2018). So, new serum-based or imaging-based markers, which are accurate for the early HCC detection, are required. Last but not least, the targeted population for HCC surveillance is far more accurate. HCC risk varies widely between patients. The surveillance is not recommended in patients with NAFLD without cirrhosis. Regarding the recent changes in the treatment paradigm for viral hepatitis and the growing incidence of NAFLD, further efforts are needed to identify reliably biomarkers to identify at risk subgroups among patients with non-cirrhotic NAFLD. Improvement of risk stratification could also improve the cost-effectiveness.

Therefore, a more accurate model is urgently needed to indicate which level of risk should trigger surveillance. And specific strategies, including new biomarkers, are needed to achieve better effectiveness in HCC surveillance.

4. Diagnosis and treatment

4.1 Invasive and non-invasive diagnostic strategies

Diagnosis of HCC is mainly based on imaging and biopsy.

Liver ultrasonography is the most frequently used tool for HCC screening. Normally HCC nodules show hypoechoic appearance, but small HCC nodules (<10mm) could be hyperechoic. But none of these appearances is specific (European Association for the Study of the Liver. Electronic address & European Association for the Study of the, 2018). Therefore, HCC is never diagnosed by US alone. Once a liver nodule is suspected in cirrhotic liver (>10mm), contrast-enhanced CT or MRI is required in routine.

The typical criteria for non-invasive diagnosis of HCC by imaging is hyperarterialization of the nodule at the arterial phase and hypo-density at the venous phase, the so-called “wash out”(Marrero, Hussain, et al., 2005). These criteria have a specificity close to 100% for the diagnosis of HCC (Forner et al., 2008). However, the non-invasive criteria are only validated in cirrhotic liver. For liver nodules occurring in non-cirrhotic liver or for those absent of a conclusive imaging pattern, tumor and non-tumor liver biopsy is strongly recommended. The sensitivity of liver biopsy for the diagnosis of HCC can vary between 70% and 90%, depending on biopsy location, differentiation, sample size and expertise. (Forner et al., 2008). A negative biopsy does not rule out HCC, but requires an enhanced surveillance or a second liver biopsy. Another challenge of biopsy is to distinguish HCC from high-grade dysplastic nodule in cirrhotic liver. In addition to morphological features, positive staining for glypican 3, heat shock protein 70 and glutamin synthetase could help to increase the diagnosis accuracy in this setting (Di Tommaso et al., 2007).

4.2 The Barcelona Clinic Liver Cancer (BCLC) staging

Once HCC is diagnosed, staging is crucial to assess prognosis and to guide treatment. The Barcelona Clinic Liver Cancer (BCLC) system is the most commonly used staging system and has been widely validated. It is based on tumor burden (number and size of nodules, vascular invasion status, extrahepatic spread), liver function (Child-Pugh's class, bilirubin, albumin, clinically relevant portal hypertension, ascites) and performance status (the Eastern Cooperative Oncology Group [ECOG] classification and presence of symptoms) (Llovet, Bru, & Bruix, 1999). As shown in **Figure 2**, the BCLC staging system classifies HCC patients in five groups: very early-stage HCC (BCLC 0): patients with single nodule less than 2cm without vascular invasion/satellites, with good health status (ECOG 0) and well-preserved liver function (Child-Pugh A); early-stage HCC (BCLC A): patients with single nodule >2cm (whatever the size) or 3 nodules <3 cm, good health status and preserved liver function; intermediate HCC (BCLC B): patients with large multinodular tumors without symptoms, without vascular invasion or extrahepatic spread; advanced HCC (BCLC C): patients with extrahepatic spread tumor and/or with macrovascular invasion; end-stage HCC (BCLC D): patients with poor liver function and/or poor performance status (ECOG>2) (European Association for the Study of the Liver. Electronic address & European Association for the Study of the, 2018). This staging system also allocates the best treatment based on the current evidences of the literature. Considering the complexity of HCC and the various available treatments, multidisciplinary treatment (MDT) involving hepatologists, surgeons, radiologists, pathologists and oncologists is strongly recommended.

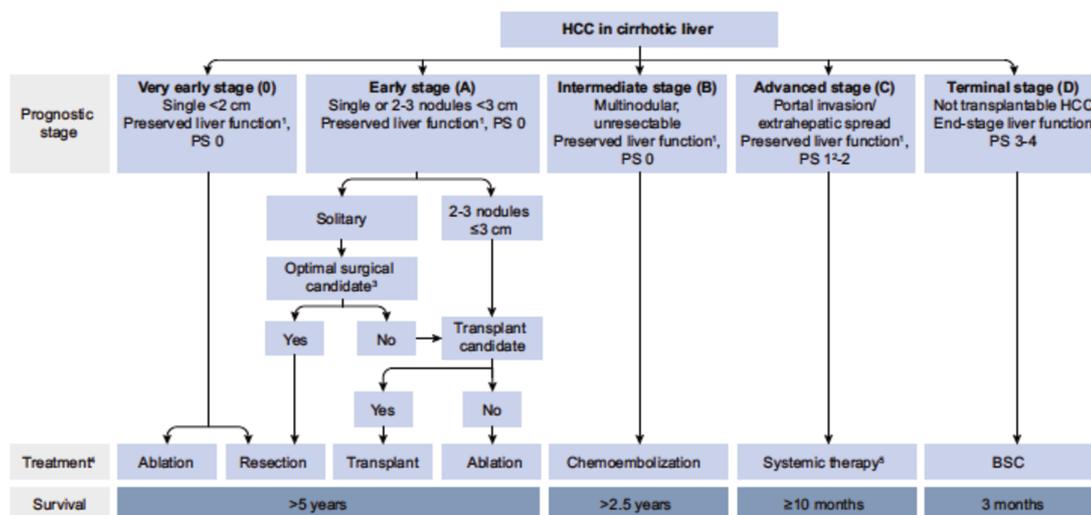


Figure 2 : Modified BCLC staging system and treatment strategy

(European Association for the Study of the Liver. Electronic address & European Association for the Study of the, 2018).

4.3 Curative treatments

Liver resection (LR) is the main choice for single HCC of any size arising on non-cirrhotic liver, when sufficient hepatic function is preserved. Resection of HCC in cirrhotic liver should be based on multi-parametric composite assessment to limit post-operative liver failure and to achieve optimal outcome. The ideal candidate in patients with cirrhosis should have a solitary tumor, well-preserved liver function, limited portal hypertension, and good performance status. Perioperative mortality of LR in patients with cirrhosis should be less than 3%. (Han et al., 2015). With the above criteria, LR is applied in 5-10% cases. Patients without clinical significant portal hypertension and normal bilirubin achieve 70% survival at 5 years, whereas survival is 50% or less when both adverse factors are present (Ishizawa et al., 2008). However, there are 70% tumor recurrence at five years after resection (Hasegawa et al., 2013).

Liver transplantation (LT) is the optimal treatment for early stage HCC on cirrhosis, since it may achieve complete cure of both HCC and the underlying liver disease. Milan criteria (single tumor ≤ 5 cm or multiple tumors ≤ 3 nodules ≤ 3 cm in size, without vascular invasion and metastasis) are the main benchmark for selecting of the best candidates for LT. The five-year survival is 65%-80% in patients meeting Milan criteria. 5-year recurrence rate is much lower (15% vs 70%) compared to LR. LT beyond the Milan

criteria is associated with high risk of recurrence and reduced long-term survival, even if extension of Milan Criteria could be proposed in some cases (Herrero et al., 2008; Mazzaferro et al., 2009; Yao et al., 2001). Moreover, after successful down-staging within Milan criteria, LT could be considered. The biggest limitation of LT is the shortage of donor. It makes the organ allocation among cancer and non-cancer candidates, and the waiting time of cancer candidates is more crucial. Living donor is an alternative strategy for organ shortage but is still limited by societal constraints and scarcity of suitable donors in Europe.

Local ablation (radiofrequency, microwave, laser and cryoablation) induces tumor necrosis and is a treatment option for single tumor less than 3cm in size and tumors which are not suitable for resection. Among them, radiofrequency ablation (RFA) is currently the first-line percutaneous treatment to reach complete tumor ablation with good and long term outcomes (Germani et al., 2010). Accurate imaging-guidance and technique improvements help to achieve complete ablation, challenging resection as first-line therapy in patients with small HCC (Cucchetti et al., 2013; Hasegawa et al., 2013). In patients with HCC nodule larger than 3cm, traditional monopolar ablation has a lower efficiency compared to LR.

4.4 Non-curative treatments

Another optional approach is transarterial chemoembolization (TACE) which is based on the predominantly arterial vascularization of HCC. The goal of this treatment is to induce a strong cytotoxic and ischemic effect on the tumor. The common drugs used during TACE are doxorubicin, epirubicin, cisplatin or miriplatin (Lencioni, de Baere, Soulen, Rilling, & Geschwind, 2016). According to guideline, patients with compensated liver function with asymptomatic multifocal or large hepatocellular carcinoma without metastasis or portal tumor thrombosis, who are not amenable to resection, are also optimal candidates for TACE. Half of the patients have objective response and improved survival. With appropriate patient selection and optimal treatment delivery, the current median survival could exceed 30–40 months (European Association for the Study of the Liver. Electronic address & European Association for the Study of the, 2018).

More than half of the patients are diagnosed with advanced HCC with the presence of tumor portal thrombosis and/or metastasis. Sorafenib, a multikinase inhibitor targeting *RAS*, *RAF*, *VEGFR* and *PDGFR* improved overall survival versus placebo (10.7 vs. 7.9 months) in these patients (Llovet et al., 2008). However, the mechanism of action is not well understood and there is not yet widely accepted clinical or molecular biomarker helping decision making or outcome prediction. It takes ten year for approval of a second-line agent, regorafenib, for patients tolerating and progressing on sorafenib (Bruix et al., 2017). Last year, another drug, cabozantinib was also recommended as a second-line treatment. At the same time, lenvatinib, another multikinase inhibitor has showed no inferiority as first-line treatment compared to sorafenib (Kudo et al., 2018).

Patients with HCC of terminal stage defined by liver failure or poor performance status (BCLC staged D) have very limited life expectancy of about 3–4 months (Llovet et al., 1999). The principle of management in these patients is symptomatic treatment in order to provide best supportive care.

II. Pathogenesis and molecular mechanism

1. Somatic genetic alterations of HCC

During liver carcinogenesis, hepatocytes accumulate genetic and epigenetic alterations involved in different process (loss of cell cycle control, escape from senescence control, resistance to cell death, phenotypic plasticity, metabolic changes) that favor cell proliferation and malignant transformation (**Figure 3**).

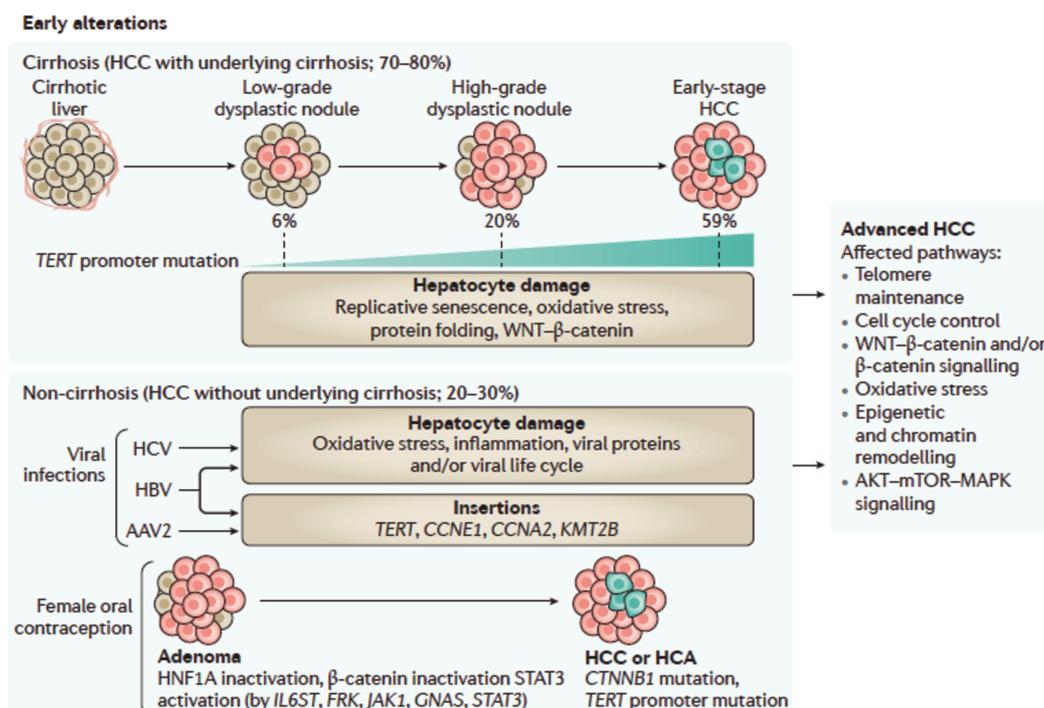


Figure 3 : Genetic alterations of Hepatocellular carcinoma.

(Llovet et al., 2016)

Telomerase reverse transcriptase (TERT) encodes a catalytic subunit of telomerase complex, which is essential for telomere elongation and maintenance preventing chromosome erosion at each cell division (Gunes & Rudolph, 2013). Somatic mutation in *TERT* promoter is the earliest recurrent genomic alteration occurring during the multistep process of liver carcinogenesis on cirrhosis. The rate of mutations increases progressively during liver carcinogenesis, from 6% in low-grade dysplasia nodule and 19% in high-grade nodule to 61% in early HCC. In normal hepatocyte, *TERT* is not expressed. However, *TERT* promoter mutations (hot spot mutation at 126 and 144 nucleotides upstream the initial translation codon) induce re-expression of the telomerase, telomere synthesis which may promote cellular abnormal proliferation. Integration of HBV DNA in *TERT* promoter is another mechanism of telomerase reactivation in HBV-related HCC (Nault et al., 2013).

In a normal liver, hepatocyte predominantly exists in quiescent phase (G0). Cyclin-dependent kinases, P53 and retinoblastoma 1 (RB1) pathways play important role in preventing quiescent hepatocytes to proliferate (Y. Lin et al., 1996). In HCC, inactivating mutations in *TP53* are common (20-52%), especially in HBV-related or

AFB1-related HCC. Strikingly, *TP53* mutation is exclusive with *CTNNB1* (*catenin* (*Cadherin-Associated Protein*), *Beta 1*) mutations. Besides, *RBI* (3-8%), *CDKN2A* (2-12%) and *ATM* (2-6%) mutations are also recurrent in HCC (Ahn et al., 2014; Hsu et al., 1991).

The Wnt/B-catenin pathway plays important roles in cell proliferation, survival, embryonic development, polarization and specification of cell fate. B-catenin encoded by *CTNNB1* is a key actor of the Wnt/ β -catenin pathway. One of the common somatic genetic alterations in HCC is activating mutations of *CTNNB1* (11-37%). Activating mutation of *CTNNB1* or inactivating mutation (5-15%) in *AXIN* or *APC* which block the degradation of β -catenin could both lead to the uncontrolled activation of the pathway (de La Coste et al., 1998; Schulze et al., 2015).

The AKT/mTOR pathway play important role in modulating cell metabolism and proliferation. Nearly 20-50% HCC has the activation of AKT/mTOR pathway. However, the mutation is quite rare with inactivating mutations in *PTEN* (1-3%) and *TSC1* or *TSC2* (3-8%). RAS/RAF/mitogen-activated protein kinase pathways are involved in proliferation, differentiation, migration and survival. In 2-9% cases of HCC, inactivating mutation in *RPS6KA3* was found, which lead to loss of negative control loop and finally lead to the activation of the pathway (Guichard et al., 2012; Zucman-Rossi, Villanueva, Nault, & Llovet, 2015).

Oxidative stress is a key player in development of chronic liver disease and the progression to HCC development. When hepatocytes are exposed to oxidative stress, the pathway of NRF2/KEAP1 is triggered in response. In HCC, there is 5-15% activating mutations in *NRF2* and *KEAP1* reported (Guichard et al., 2012). And in-vitro model showed that constitutive activation of the pathway promoted cell survival and inhibited cell death, revealing that *NRF2* was a critical oncogene for HCC progression and development (Zavattari et al., 2015).

Alterations in chromatin remodeling and histone methyltransferase genes have emerged as a major dysregulated pathway in HCC. Among them, there are inactivating mutations of *ARID1A* (13%) and *ARID2* (7%) as well as mutations of the *MLL* gene

family (mutations of *MLL* (3-4%), *MLL2* (2-3%), *MLL3* (3-6%) and *MLL4* (2-3%)), which codes for histone methyl transferases (Zucman-Rossi et al., 2015).

2. Molecular classification of HCC

Regarding the intertumor and intratumor heterogeneity of HCC, development of new screening and stratification programs to refine diagnosis and to improve patients' outcomes becomes crucial. Among multiple technologies and the related multi-omics dataset, the gene expressing profile acts as a mediator between most other datasets (Aben et al., 2018). Therefore, molecular classification of HCC is based on transcriptomic profiling.

In year 2004, Snorri Thorgeisson analyzed the gene expression data from rat fetal hepatoblasts and adult hepatocytes with HCC from human and mouse models. He proposed the first classification of HCC with two main groups: cluster A as "proliferative" HCC linked with poor prognosis and cluster B as "non-proliferative" HCC (J. S. Lee et al., 2006).

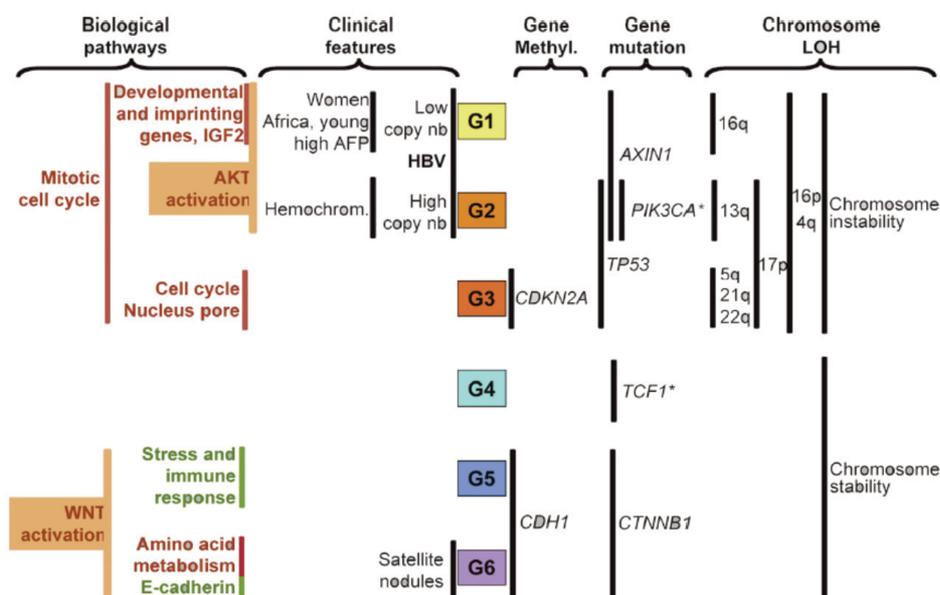


Figure 4 : Molecular classification of Hepatocellular carcinoma.

(Boyault et al., 2007)

In year 2007, our laboratory described a transcriptomic classification of HCC in 6 subgroups (G1-G6) associated with clinical and genetic features (Figure 4) (Boyault

et al., 2007). Generally, G1-G2-G3 was associated with chromosomal instability, in contrast, tumors classified in G4-G5-G6 were chromosomally stable. HCCs in G1 and G2 groups were more frequently HBV-related, and harbored AKT pathway activation and *AXIN1* mutation. The difference between them was that G1 group displays a high AFP level and was associated low copy number of HBV while G2 group was associated with high copy number of HBV and *TP53* mutation. G3 group was characterized by overexpression of genes of the cell cycle and was associated with a specific histological subtype of HCC, the so-called macrotrabecular massive HCC. HCC in G4 group was well differentiated HCC and was occasionally mutated in *HNF1A*. Subgroups of G5 and G6 were closely associated with activating mutation of *CTNNB1* which led the activation of the WNT/ β -catenin pathway. In addition, G6 was characterized by satellite nodules and under-expression of E-cadherin.

Several molecular classifications of HCC have been proposed in literature, such as S1/S2/S3 system, TGFB, MET, AKT/MTOR (Chiang et al., 2008; Coulouarn, Factor, & Thorgeirsson, 2008; Hoshida et al., 2009; Kaposi-Novak et al., 2006). Undoubtedly, these studies lead to crucial progress in our understanding of the molecular mechanisms of HCC. However, challenges lie on refinement of these classification and translating the findings to clinical practice.

III. Genetic variants associated with risk of chronic liver disease and HCC

1. HCC susceptibility and single nucleotide polymorphism

HCC is a genetic disease not only in terms of somatic genetic alterations but also in terms of the heterogeneity of the genetic determinants of susceptibility to develop HCC. The heterogeneity of HCC susceptibility exists at two levels: among different ethnicity and in the same ethnicity. People living in the same region, for example, in America, Asians and Pacific Islanders have higher age-standardized HCC rates than Hispanics, African-Americans and Whites at all ages and among both genders. And individuals with the same etiology and the same ethnicity have also different risk to develop HCC. There is a relative percentage of HCC cases diagnosed without any known predisposing

risk factor and without any chronic liver disease (CLD) (El-Serag & Mason, 2000). Moreover, familial history of HCC is also a risk factor of HCC development in patients with chronic HBV and HCV infection (Hassan et al., 2009).

All these facts suggest a role of individual genetic predisposition in the susceptibility of HCC. Single nucleotide polymorphism (SNP) is the most abundant form of genetic variations, defined as a single base change in coding, non-coding and intergenic regions. The identification of genetic variants influencing the risk of HCC is crucial for identifying the subgroup of patients at high risk of HCC development.

For decades, the contribution of SNPs in the development of HCC has been investigated. Numerous candidate-gene studies were undertaken by selecting genes based on the known functions relevant to liver tumorigenesis, for example, inflammatory genes, oncogenes and CLD-associated genes. Some polymorphisms associated with viral hepatitis, NAFLD and fibrosis progression have also been suggested as risk factors of HCC. Unfortunately, some of these studies suffer from major methodological drawbacks because of their case-control, retrospective and monocentric design. Prospective cohort study with sufficient number of homogeneous populations which may overcome these pitfalls is still scarce.

In contrast to candidate-gene study, genome-wide association study (GWAS) proposes to compare the distributions of millions of SNPs in HCC patients and control population without a prior hypothesis. This agnostic approach allows to reveal previously unsuspected variants associated with risk of HCC and, in the meantime, to unravel new pathway possibly involved in liver carcinogenesis. Over the past few years, this high-throughput genotyping technology has revealed SNPs potentially associated with HCC risk. However, several drawbacks exist in GWAS study performed in HCC patients. Due to differences in ethnicity or control population, some of the reported SNPs are rarely reproducible in another GWAS. Thus, well-designed studies with homogeneously selected patients are needed for the validation of these SNPs.

HCC is a complex polygenic disease. According to “common disease, common variants” hypothesis, each variant at each gene will have a small additive or

multiplicative effect on the risk of HCC. Until now, the SNPs identified so far explain only a small proportion of the overall variability in HCC susceptibility. In order to better predispose HCC, more genetic markers are needed.

2. Polymorphisms and predisposition to chronic liver diseases

2.1 Genome-wide association study

Chronic hepatitis B: As what is mentioned above, nearly 95% of population with HBV infection will spontaneously eliminate the virus, and the remaining will establish chronic infection and progress to chronic B hepatitis (CHB). Among them, one fourth of chronic carriers will progress to cirrhosis and 5% with cirrhosis may develop HCC (F. S. Wang, 2003). Owing these epidemiologic facts, genetic variations play a role in the predisposition of infection, occurrence of chronic hepatitis B, liver cirrhosis and even HCC. The first GWAS about CHB was reported in 2009. In Japanese, rs3077 in *HLA-DPA1* and rs9277535 in *HLA-DPB1* were significantly associated with persistent infection with hepatitis B virus (Kamatani et al., 2009). Then 7 other GWAS validated these two variants in other Asian populations, along with the discovery of other novel variants within the HLA Class II Genes. In clinical study, strong T cell responses involving both HLA class II-restricted CD4⁺ helper T cells and HLA class I-restricted CD8⁺ cytotoxic T cells were present in subjects who spontaneously clear HBV infection (Chisari & Ferrari, 1995; Ganem & Prince, 2004). It indicates that identified genetic variants in HLA gene may play a role in the immune response controlling HBV clearance and disease pathogenesis. In Korean population, the allelic frequency of SNPs in non-HLA class II genes, rs652888 in the *euchromatic histone-lysine-methyltransferase 2 (EHMT2)* and rs1419881 in *transcription factor 19 (TCF19)*, were found significantly different in patients with CHB and healthy control (Y. J. Kim et al., 2013). In Chinese population, four novel loci, rs3130542 near *HLA-C* (an MHC class I gene), rs4821116 in the *ubiquitin-conjugating enzyme E2L 3 (UBE2L3)* gene, rs7000921 in *integrator complex subunit 10 (INTS10)* gene and rs2296651 in *SLC10A1* gene were separately reported in GWAS associated with the clearance of HBV infection (Hu et al., 2013; Y. Li et al., 2016). The only GWAS in non-Asian population identified that rs2724432, locating upstream of *ferredoxin 1 (FDX1)*, was associated with disease

progression in patients with CHB, by comparing inactive HBV carriers with LC plus HCC groups in Arab population (A. Al-Qahtani et al., 2013).

Chronic hepatitis C: GWAS are conducted to identify the host genetic variants that influence treatment efficacy or clinical course in HCV infection. Between 2009 and 2010, four independent groups from the US, Australia, Europe, and Japan performed GWAS in HCV infected individuals and identified several SNPs of *IL28B* gene, associated with both spontaneous and interferon-induced HCV clearance (Ge et al., 2009; Rauch et al., 2010; Suppiah et al., 2009; Tanaka et al., 2009). Regarding disease progression, several GWAS have identified that SNPs, including rs16851720 and rs4374383 in genes linked with apoptosis, rs910049 and rs3135363 in MHC region were separately associated with liver fibrosis progression (Patin et al., 2012; Urabe et al., 2013). A GWAS in HIV/HCV-coinfected patients revealed a strong association between hepatic fibrosis progression and rs61183828, who is linked with a cluster of SNPs close to two genes involved in cell signaling and cell structure (*caveolin-3*, *CAV3*) and HCV replication (*RAD18*) (Ulveling et al., 2016).

Non-alcoholic fatty liver disease: NAFLD is a complex disease with a substantial heritable component. The first GWAS of NAFLD was conducted in 2008 in a multi-ethnic population. It identified that rs738409 in *patatin-like phospholipase domain-containing 3* (*PNPLA3*) was significantly associated with the presence of liver steatosis (Romeo et al., 2008). Considering the complexity of NAFLD, a GWAS in Japanese population adjusted the well-known confounders as age, gender, and BMI, and reported that *PNPLA3* rs738409 was strongly associated with the development and fibrosis progression of NAFLD ($P = 6.89 \times 10^{-14}$, OR = 2.05) (Kitamoto et al., 2013). Until now, numerous GWAS have confirmed its strong association with hepatic steatosis as well as hepatic fibrosis. Usually, a systematic meta-analysis regarding all the possible studies provide the highest level of evidence. *PNPLA3* rs738409 remains the most commonly validated in the clinical course of NAFLD, including susceptibility to liver injury and fibrosis progression, in meta-analysis studies across a vast array of ethnicities and in large samples of patients (Sookoian & Pirola, 2011; R. Xu, Tao, Zhang, Deng, & Chen, 2015; L. Zhang et al., 2015). Specifically, GWAS of liver

enzymes has also revealed the association of *PNPLA3* rs738409 with elevated serum alanine transaminase (ALT) levels (Chambers et al., 2011; Yuan et al., 2008).

In 2014, another GWAS about hepatic fat reported a non-synonymous SNP of *transmembrane 6 Superfamily Member 2* (*TM6SF2*, rs58542926) associated with higher levels of ALT, and lower levels of LDL-cholesterol (very low density lipoprotein, VLDL), triglycerides and alkaline phosphatase (Kozlitina et al., 2014). Until now, numerous studies, including meta-analysis studies have already validated the association between this *TM6SF2* SNP and the risk of NAFLD (Goffredo et al., 2016; Pirola & Sookoian, 2015). In addition, a study further revealed that T allele of the *TM6SF2* rs58542926 was associated with increased risk of advanced fibrosis in patients with NAFLD, independent of the *PNPLA3* genotype and other confounding factors (Y. L. Liu, Reeves, et al., 2014).

A GWAS of computed tomography (CT) measured hepatic steatosis identified that additional genetic variants influencing NAFLD localized in or near the genes *neurocan* (*NCAN*-rs2228603), *protein phosphatase 1, regulatory (inhibitor) subunit 3B* (*PPP1R3B*-rs4240624), *glucokinase regulator* (*GCKR*-rs780094), and *lysophospholipase-like 1* (*LYPLAL1*-rs12137855) and *PNPLA3*-rs738409 (Speliotes et al., 2011). Among them, rs2228603 of *NCAN* was found in moderate linkage disequilibrium with rs58542926 in *TM6SF2* (pairwise r^2 : 0.798; <0.0001 in Caucasian). A meta-analysis incorporating a total of 2091 NAFLD cases and 3003 controls from five studies had confirmed the association between *GCKR* rs780094 and NAFLD in both Asian and non-Asian populations (Zain, Mohamed, & Mohamed, 2015). A GWAS of non-Hispanic white woman investigated the association between genetic variants and the severity of NAFLD at histological level. It described some genetic variants conferring to susceptibility of NAFLD and disease progression. Among them, the SNP rs2645424 in *farnesyl diphosphate farnesyl transferase 1* (*FDFT1*) was associated with the NAFLD activity score (NAS). The SNPs associated with lobular inflammation included rs1227756 in *COL13A1*, rs6591182 on chromosome 11, and rs887304 in *EFCAB4B*. And SNPs associated with serum level of ALT included rs2499604 on chromosome 1, rs6487679 on chromosome 12 in *PZP*, rs1421201 on chromosome 18, and rs2710833 on chromosome 4 (Chalasani et al., 2010). These findings call for

further extensive validation in larger cohorts. Very recently, a GWAS using whole-exome sequencing data identified a splicing variant in *17-beta-hydroxysteroid dehydrogenase 13 (HSD17B13)*, rs72613567, associated with reduced level of ALT. Most importantly, this splicing variant of HSD17B13 could also reduce the risk of NAFLD and protect from the progression from steatosis to NAFLD-related cirrhosis (Abul-Husn et al., 2018).

To resume, *PNPLA3* rs738409, *TM6SF2* rs58542926, *HSD17B13* rs72613567 and *GCKR* rs780094 are the most robust variants identified in GWAS contributing to genetic susceptibility of NAFLD.

Alcoholic liver disease: Excess consumption of alcohol led to various health problems, including alcoholic liver disease (ALD). However, only a subset of individuals with high alcohol intake will develop ALD. Individual genetic factors play a role in ALD susceptibility. GWAS of ALD has been studied focusing on the phenotypes ‘alcohol-related cirrhosis’ and ‘severe alcoholic hepatitis’. In a GWAS study, *PNPLA3* rs738409 was confirmed as an important risk locus for alcohol-related cirrhosis at a genome-wide level of significance, meanwhile two SNPs separately in *TM6SF2* (rs58542926) and *membrane bound O-acyl transferase (MBOAT)*, rs626283) genes were also identified as risk loci (Buch et al., 2015). The biological functions of these three genes suggest a role of lipid processing in the pathogenesis of alcohol-related cirrhosis. *PNPLA3* rs738409 has been widely validated as associated with ALD and the progression to alcohol-related cirrhosis. The relative studies are summarized in **Table 2**. A meta-analysis including these studies also provided strong and unequivocal evidence for a significant role of *PNPLA3* rs738409 in the progression of ALD (Salameh et al., 2015). Moreover, GG genotype of *PNPLA3* rs738409 confers significant additional risk of medium-term mortality in patients with severe alcoholic hepatitis (Meffert et al., 2018).

| Author and year | Population | Cases (number) | Control (number) | p-value |
|---------------------|--------------------|----------------|---|----------|
| Tian, 2010 | Mexican | ALC (482) | Alcohol dependent but normal liver function test(305) | 4.70E-05 |
| Seth, 2010 | British | ALC (266) | Heavy drinkers but normal liver function test(182) | 2.00E-06 |
| Trépo, 2011 | Belgian and French | ALD (328) | Healthy controls (330) | 8.00E-03 |
| Stickel, 2011 | German | ALC (210) | Alcohol dependent: no liver injury (439) | 1.18E-05 |
| Nischalke, 2011 | German | ALC (80) | Healthy controls (190) | <0.002 |
| Falleti, 2011 | Italian | ALD (166) | Health controls (428) | <0.0001 |
| Rosendahl, 2012 | Dutch and German | ALD (135) | Healthy controls (2781) | <0.0001 |
| Dutta, 2013 | Indian | ALD (60) | Healthy controls (100) | 0.037 |
| Burza, 2014 | Italian | ALD (84) | Alcohol dependence: no liver injury (300) | 0.021 |
| Friedrich, 2014 | Caucasian | ALD (105) | Healthy controls (1950) | <0.005 |
| Liangpunsakul, 2017 | 80% Caucasian | AH (189) | Heavy drinker (150) | <0.0001 |
| Atkinson, 2017 | British | ALC (898) | Heavy drinker (1188) | 2.15E-15 |
| Zhang, 2018 | Chinese | ALD (507) | Healthy control (645) | 9.05E-13 |
| Kolla, 2018 | Caucasion | ALC (169) | noncirrhotic alcohol consumer (259) | 0.00029 |
| Mancina, 2019 | Italian | ALC (90) | noncirrhotic alcohol consumer (326) | 0.008 |

Table 2: Genetic association between *PNPLA3* rs738409 (G) and alcohol-related cirrhosis.

ALD: alcoholic liver disease; ALC: alcoholic liver cirrhosis; AH: alcoholic hepatitis.

As mentioned above, the splicing SNP of *HSD17B13* (rs72613567) was also found significantly associated with a reduced risk of ALD (by 42% and 53% separately among T/TA and TA/TA genotype carriers) and alcohol-related cirrhosis (by 42% and 73% separately among T/TA and TA/TA genotype carriers). This association was confirmed in two independent cohorts (Abul-Husn et al., 2018). *HSD17B13* is abundantly expressed in liver tissue and locates on the surface of lipid droplet (Su et al., 2014). Genetic information gathered in GWAS these recent years has shed new light on the inherited aspects of ALD but more data are required to understand the pathogenesis of ALD.

All these GWAS conducted in domain of chronic liver diseases are summarized in **Table 3**.

| Chronic liver disease | Reference | Population | Phenotype | Gene | SNP |
|------------------------|-------------------------------|--|--|--------------------------------|------------|
| HBV | Kamatani et al. 2009 | Japanese | persistent HBV infection | HLA-DPA1 | rs3077 |
| | | | | HLA-DPB1 | rs9277535 |
| | Mbarek et al. 2011 | Japanese | persistent HBV infection | HLA-DQ | rs2856718 |
| | | | | HLA-DQ | rs7453920 |
| | Kim et al. 2013 | Korean | CHB | EHMT2 | rs652888 |
| | | | | TCF19 | rs1419881 |
| | Nishida et al. 2012 | Asian | HBV clearance | HLA-DPA1 | rs3077 |
| | | | | HLA-DPB1 | rs9277542 |
| | Hu et al. 2013 | Chinese | clearance of HBV infection | near HLA-C | rs3130542 |
| | | | | UBE2L3 | rs4821116 |
| Li et al. 2016 | Chinese | clearance of HBV infection | INTS10 | rs7000921 | |
| Al-Qahtani et al. 2013 | Arab | HBV related cirrhosis+HCC | FDX1 | rs2724432 | |
| Jiang et al. 2015 | Chinese | CHB | CD40 | rs1883832 | |
| | | | CFB | rs12614 | |
| | | | NOTCH4 | rs422951 | |
| | | | HLA-DOA | rs378352 | |
| | | | near HLA-C | rs2853953 | |
| HCV | Patin et al. 2012 | European | fibrosis progression | RNF7 | rs16851720 |
| | | | | MERTK | rs4374383 |
| | Urabe et al. 2013 | Japanese | fibrosis progression | TULP1 | rs9380516 |
| | | | | C6orf10 | rs910049 |
| | Ulveling et al. 2016 | European | fibrosis progression in HIV/HCV coinfection | BTNL2-HLA-DRA | rs3135363 |
| | | | | CAV3-OXTR-RAD18 | rs61183828 |
| | Tanaka et al. 2011 | Japanese | response to pegylated interferon-alpha and ribavirin therapy | IL28 | rs8099917 |
| Ge et al. 2009 | European and African American | hepatitis C treatment-induced viral clearance | IL28 | rs12979860 | |
| Suppiah et al. 2009 | Australian | response to chronic hepatitis C interferon-alpha and ribavirin therapy | IL28 | rs8099917 | |
| Rauch et al. 2010 | European | HCV persistence | IL28 | rs8099917 | |
| | | response to therapy | IL28 | rs8099917 | |
| NAFLD | Romeo et al. 2008 | Multi ethnic | nonalcoholic fatty liver disease | PNPLA3 | rs738409 |
| | Speliotes et al. 2011 | Multi ethnic | steatosis by CT scan | PNPLA3 | rs738409 |
| | | | | NCAN | rs2228603 |
| | Kawaguchi et al. 2012 | Japanese | NAFLD | PPP1R3B | rs4240624 |
| | | | | PNPLA3 | rs738409 |
| | Kitamoto et al. 2013 | Japanese | NAFLD, steatosis, fibrosis, NAS | PNPLA3 | rs738409 |
| | | | | PNPLA3 | rs738409 |
| | Feitosa et al. 2013 | Caucasian | Steatosis by CT | PPP1R3B | rs2126259 |
| | | | | ERLIN-CHUKCWF19L1 gene cluster | |
| | | | | PNPLA3 | rs738409 |
| Kozlitina et al. 2014 | Multi ethnic | Hepatic fat content | TM6SF2 | rs58542926 | |
| | | | PNPLA3 | rs738409 | |
| DiStefano et al. 2015 | Multi ethnic | Hepatic fat content | SUGP1 | rs10401969 | |
| | | | GC | rs222054 | |
| Adams et al. 2013 | Australian adolescent | NAFLD by US | LCP1 | rs7324845 | |
| | | | LPPR4 | rs12743824 | |
| | | | SLC38A8 | rs11864146 | |
| | | | HSD17B13 | rs72613567 | |
| Abul-Husn et al. 2018 | European | NAFLD-related cirrhosis | PNPLA3 | rs738409 | |
| ALD | European | ALD-related cirrhosis | TM6SF2 | rs58542926 | |
| | | | MBOAT7 | rs626283 | |
| | | | HSD17B13 | rs72613567 | |

Table 3: SNPs identified in GWAS of chronic liver diseases.

Phenotype CHB are compared between patients with persistent HBV infection and healthy population.

CHB: chronic hepatitis B; ALD: alcoholic liver disease; US: ultrasonography. NAS: NAFLD activity score.

2.2 Candidate-gene study

Chronic hepatitis B: Almost all the GWAS of CHB were conducted in Asian populations. Considering the ethnic heterogeneity, analyses of the association between HLA class II and HBV infection in western countries are required. A candidate-gene study was performed to investigate the association between *HLA-DP* variants and HBV susceptibility in 201 Caucasian HBV carriers and 235 healthy controls. It showed that *HLA-DPA1* rs3077 (C allele) was significantly associated with HBV susceptibility, whereas *HLA-DPB1* rs9277535 was not associated with HBV infection (Vermehren et al., 2012). In European–American and African–American ancestry, rs9277535 in *HLA-DPB1* had a marginal effect on HBV clearance, but a novel variant in the 3'UTR region of *HLA-DPB1* (rs9277534) was significantly associated with HBV clearance (Thomas et al., 2012). Several groups also used candidate-gene approach to investigate whether these SNPs were associated with disease progression in patients with CHB. Most of the studies failed to show a significant association between the SNPs identified in HLA class genes and the disease progression to liver cirrhosis or HCC (An et al., 2011; J. Li et al., 2011; Vermehren et al., 2012). Novel genetic variants were found by candidate-gene study, for example, the genotype distribution of rs1265163 in *OCT4* gene, a gene nearby *TCF19*, was significantly different between patients with CHB and general control population (Shin et al., 2017). As a complex disease, HBV infection and disease progression involves multiply genes. Therefore, further replicative studies in larger independent cohorts are required in different populations.

Chronic hepatitis C: Candidate-gene studies have identified SNPs in several cytokine and chemokine genes such as *IL6*, *CCL2*, *CCL8*, *TNF*, and *TGFBI*, significantly associated with progression of liver fibrosis (Falleti et al., 2010; Hellier et al., 2003; Muhlbauer et al., 2003; Powell et al., 2000). In addition, HLA class II locus were linked with inflammatory activity and disease progression in patients with chronic HCV infection (R. B. Yu et al., 2008). Several candidate-gene studies have shown that the at-risk G allele of *PNPLA3* rs738409 and T allele of *TM6SF2* rs58542926 were also associated with hepatic steatosis and fibrosis progression in patients with CHC of European and Japanese ancestry (Fan, Xiang, Li, Shi, & Guo, 2016; Milano et al., 2015; Trepo, Pradat, et al., 2011; Yasui et al., 2015; H. Zhang et al., 2015). Furthermore,

Tamaki et al. investigated several genetic variants in longitudinal studies of CHC with paired liver biopsies. They found that patients with an unfavorable *IFNL3-IFNL4* genotype or with the G allele of *PNPLA3* rs738409 were at risk of rapid fibrosis progression (Tamaki et al., 2015).

Alcoholic liver disease: Several candidate-gene studies about ALD have been performed and the candidate genes were selected based on the biological understanding of liver injury, including genes involved in alcohol metabolism; hepatic lipid turnover; modulation of endotoxin-mediated inflammation and cytokines; DNA damage and carcinogenesis; iron metabolism; immune responses; oxidative stress; tissue remodeling and fibrogenesis. However, several studies showed no association between some SNPs and ALD, and those with a possible association could not be replicated in another study or meta-analysis. The exceptions were the SNPs in promoter of *tumor necrosis factor (TNF)- α* (rs361525) and the homozygous deletions of *glutathione S-transferase mu 1 (GSTM1)*, an enzyme that are responsible for the detoxification of potentially toxic by-products of ethanol metabolism (Marcos, Gomez-Munuera, Pastor, Gonzalez-Sarmiento, & Laso, 2009; Marcos et al., 2011). There are two main reasons for these negative results: 1) the selection of these candidate genes was based on an unclear pathophysiology of ALD, making the candidate genes not the right ones; 2) the difficulty in homogenizing the cases and control regarding the complex condition of alcohol exposure and alcoholic liver disease.

Non-alcoholic fatty liver disease: Several candidate-gene studies have been performed in order to evaluate the influence of genetic determinants on the susceptibility and pathogenesis of NAFLD. *Apolipoprotein C3 (APOC3)* is a major constituent of VLDL, and inhibits the hydrolysis of TG-rich particles by lipoprotein lipase (McConathy et al., 1992). A study reported two SNPs in *APOC3* (rs2854116 and rs2854117) associated with NAFLD (Petersen et al., 2010). However, meta-analysis of prospective data finally failed to confirm this association (H. Zhang et al., 2014). In addition, SNPs in *PPAR γ* (rs1805192 and 1801282) had been examined in two meta-analyses, with discrepant association with NAFLD in East Asian population and with no association in European population (Y. H. Lee, Bae, & Song, 2014; J. Wang et al., 2013). Other candidate genes investigated could be mainly devised into two categories:

genes modulating the glucid or lipid metabolism (ex. *leptin receptor*, *Resistin*, *Phosphatidylethanolamine N-methyltransferase*, *β-adrenergic receptors*, *Apolipoprotein E*, *microsomal triglyceride transfer protein (MTTP)* and etc.) and genes involved in mechanisms of liver injury (ex. *TNF-α*, *IL-6*, *TLR4*, *IL-28B*, *superoxide dismutase 2*, *methylenetetrahydrofolate reductase (MTHFR)*, *TGF-β1* and etc.) (Macaluso, Maida, & Petta, 2015). Regarding the complex physiological function of liver, large and well-designed studies are warranted to validate the findings.

3. Genetic associations in HCC

3.1 Genome-wide association study

| SNP | Gene | Min/Maj Allele | Etiology | Case and Control | Risk Allele | p-value | Ethnicity | Study |
|------------|---------------|----------------|-----------|-------------------------------|-------------|----------|-----------|--------------------|
| rs1012068 | DEPDC5 | G/T | HCV | 212/765HCV 710/1625Healthy | G | 8.05E-08 | Japanese | D Miki (2011) |
| rs455804 | GRIK1 | A/C | HBV | 1538/1,465HBV | C | 1.86E-08 | Chinese | S Li (2012) |
| rs9272105 | HLA-DQA1/DQB1 | A/G | | | G | 1.13E-19 | | |
| rs17401966 | KIF1B | G/A | HBV | 355/360HBV | A | 5.80E-06 | Chinese | H Zhang (2010) |
| rs7574865 | STAT4 | T/G | HBV | 1161/1353HBV | G | 2.48E-10 | Chinese | D Jiang (2013) |
| rs9275319 | HLA-DQ | A/G | | | A | 2.72E-17 | | |
| rs2596542 | MICA | T/C | HCV | 721/2890Healthy | T | 4.21E-13 | Japanese | V Kumar (2011) |
| rs2551677 | DDX18 region | A/G | Cirrhosis | 386/86Cirrhosis | A | 1.41E-10 | Asian | RJ Clifford (2010) |
| rs17047200 | TLL1 | T/A | HCV | 123/333HCV | T | 2.66E-08 | Japanese | K Matsuura (2017) |
| rs10272859 | CDK14 | C/G | HBV | 355/360HBV | G | 9.46E-10 | Chinese | Y Li (2018) |

Table 4: SNPs identified in GWAS associated with the risk of HCC

Until now, 8 GWAS have been conducted in HCC (summarized in **Table 4**), mainly in Asian population with chronic hepatitis B or C. In 2010, the first GWAS focusing on HCC was performed in 355 Chinese patients with chronic hepatitis B and 360 CHB patients without HCC, then followed by replicative studies in five independent Chinese cohorts. It identified that rs17401966 located in *kinesin family member 1B* gene (*KIF1B*) has a strong correlation with HBV-related HCC (OR=0.61, combined $p=1.7 \times 10^{-18}$) (H. Zhang et al., 2010). The association was consistent with the idea that KIF1B acts as a tumor suppressor gene (Munirajan et al., 2008; Schlisio et al., 2008).

The second GWAS of HBV-related HCC was performed in a much bigger sample size (3003 Chinese CHB patients, among them 1538 with HCC). The previously reported

association between *KIF1B* rs17401966 and HBV-related HCC was confirmed but with a much weaker strength (OR=0.90 V.S. OR=0.61). This GWAS revealed two novel loci linked with the risk of HBV-related HCC, rs9272105 in *HLA-DQA1/DRB1* and rs455804 in *glutamate ionotropic receptor kainate type subunit 1 (GRIK1)*, a gene involved in the glutamate signaling (S. Li et al., 2012). It highlighted the important role of glutamate signaling in the development of HBV-related HCC. Unfortunately, no further studies were performed to validate the SNP of *GRIK1* and to investigate its functions in liver carcinogenesis.

Since *KIF1B* rs17401966 was identified in two GWAS of HBV-related HCC, numbers of candidate-gene studies were performed. Most of them failed to validate its association with HBV-related HCC, but one study showed a weak association ($p=0.044$) (Al-Qahtani et al., 2012; M. Huang et al., 2014; Sopipong, Tangkijvanich, Payungporn, Posuwan, & Poovorawan, 2013; Zhong et al., 2012). In summary, due to the inconsistent results of several investigations, additional investigations with more clinical data and new analysis models are required to study the effects of *KIF1B* polymorphism on liver carcinogenesis.

Another multi-stage GWAS of HBV-related HCC in Chinese population reported two novel genetic susceptibility loci, rs7574865 in *signal transducer and activator of transcription 4 (STAT4)* and rs9275319 in *HLA-DQ*. This study included 2514 chronic HBV carriers (1161 with HCC) in the discovery cohort and 8031 chronic HBV carriers (3737 with HCC) in five independent validation cohorts. OR for HCC was 1.22[1.15-1.29] for G allele of *STAT4* rs7574865 in meta-analysis. Strikingly, this intron variant was associated with a lower mRNA level of *STAT4* (Jiang et al., 2013). Consistently, other SNPs in *HLA-DQ* gene (rs3077, rs9277535, rs2856718, and rs7453920) have also been reported associated with CHB in a GWAS in Japanese population (Mbarek et al., 2011). Both *STAT4* and *HLA-DQ* gene are involved in immune response, indicating a biological link between their genetic variant and HBV-related HCC development. Then validation studies were performed in Asian populations, but the association between *STAT4* rs7574865 and the risk of HBV-related HCC showed conflicting results among different studies (K. Chen et al., 2013; W. Chen et al., 2015; Clark et al., 2013; L. H. Kim et al., 2015). A meta-analysis including 3467 HCC patients and 5821 HBV carriers

failed to confirm the association of *STAT4* rs7574865 with HCC development (Liao et al., 2014).

The most recent GWAS in HBV-related HCC was conducted in a collection of 205 HBV-related HCC trios, 355 chronic HBV carriers with HCC and 360 chronic HBV carriers, followed by two replication cohorts of 3796 cases and 2544 controls. It confirmed rs17401966 at *KIF1B*, rs9272105 at *HLA-DQA1/DRB1*, rs9275319 at *HLA-DQB1/DQA2*, and rs12663434 at *BACH2* as susceptibility loci of HCC, but failed to validate rs455804 at *GRIK1* and rs7574865 at *STAT4*. In addition, they identified a novel susceptibility loci within *cyclin dependent kinase 14 (CDK14)* gene, rs10272859, which was also associated with survival of HCC patients. (Y. Li et al., 2018).

Moreover, GWAS was conducted in HCV-related HCC in Japanese population, including 212 patients with HCV-related HCC and 765 CHC patients without HCC. One intronic SNP in *DEP domain containing 5 (DEPDC5)*, rs1012068, was significantly associated with susceptibility to HCC even after adjustment for age, gender and platelet count (OR = 1.96; P = 1.35×10^{-14}). Function of *DEPDC5* is poorly understood, but this study found that gene expression level of *DEPDC5* was much higher in tumor tissue than non-tumor tissue, indicating a possible role in liver carcinogenesis (W. Liu et al., 2019).

The second GWAS of HCV-related HCC was also conducted in Japanese population. In 721 cases with HCV-related HCC and 2890 HCV-negative control, 432,703 SNPs were genotyped and those with p value $< 10^{-5}$ were further genotyped in a second cohort including 673 cases and 2596 controls. A SNP (rs2596542: A) in the 5' flanking region of *MHC class I polypeptide related sequence A (MICA)* was associated with HCV-related HCC and lower level of soluble MICA protein in HCC patients. MICA was a membrane protein acting as ligand of NKG2D to active anti-tumor effects through NK cells and CD8+ T cells. Lower level of serum MICA related to A allele of rs2596542 reduces the activation of NK cells and other NKG2D-positive immune cells, which lead the progression from CHC to HCC (Kumar et al., 2011). However, detailed fine-mapping analyses around *MICA* region failed to validate *MICA* rs2596542, but revealed rs2244546 in *HCP5*, upstream of *MICA*, as associated with HCV-related HCC

in both Caucasian and Japanese populations (Lange et al., 2013). Several other studies in patients with CHC had controversial results about the associations between *MICA* rs2596542 and HCC risk, requiring further validations (Augello et al., 2018; Burza et al., 2016). After the GWAS, Kumar et al. continued to investigate the involvement of *MICA* in liver carcinogenesis related to HBV infection. They identified, by candidate-gene study, the G allele of *MICA* rs2596542 was associated with increased risk of HBV-related HCC and elevated level of serum MICA, which was associated with poor prognosis (Kumar et al., 2012). But some inconsistent data exist. In Saudi Arabian patients, allele A of *MICA* SNP was only associated with active infection of HBV but not HBV-related cirrhosis and HCC; while in Vietnamese patients with chronic hepatitis B, A allele of rs2596542 was found not only associated with HBV clearance but also HBV-related HCC in patients with cirrhosis (A. A. Al-Qahtani et al., 2013; Tong et al., 2013). The involvement of *MICA* SNP in liver carcinogenesis requires further investigations.

Finally, a recent GWAS in Japanese population identified a SNP (rs17047200) located in the intron of the *tolloid like 1* gene (*TLL1*), associated with the development of HCC after achieving SVR on IFN-based therapy. TLL1 exert several biological roles in the regulation of extracellular matrix assembly and in the transforming growth factor β (TGF- β) signaling pathway. This SNP was also found associated with mRNA of splicing of *TLL1*, which might yield short variants with high catalytic activity that accelerates hepatic fibrogenesis and carcinogenesis. (Matsuura et al., 2017).

Since all the GWAS of HCC were conducted in Asian population with chronic viral hepatitis, whether the associations of these genetic variants with HCC are still valid in other populations or in other etiologies. In addition, in the GWAS conducted in Asian population, several biases exist including unbalanced study design, inappropriate control population or absence of adjustment for the severity on chronic liver disease. Therefore, additional validations of these genetic variants are required.

3.2 Candidate-gene study

Many candidate-genes studies were performed and contributed to the discovery of susceptibility loci of HCC.

PNPLA3: The most widely studied genetic variant in the field of HCC predisposition is the lost-of-function variant of *PNPLA3*, rs738409. Since it was firstly showed associated with NAFLD in a GWAS in 2008, researchers began to investigate its possible association with progression of liver disease and development of HCC (summarized in **Table 5**). In Caucasian population, several retrospective studies and meta-analysis studies had revealed that *PNPLA3* rs738409 was associated with ALD-related HCC development in patients with liver cirrhosis (Friedrich et al., 2014; Nischalke et al., 2011; Salameh et al., 2015; A. G. Singal et al., 2014; Stickel et al., 2018a; Trepo et al., 2012; Trepo et al., 2014; S. Zhang et al., 2015). But patients with HCC developed on a non-cirrhotic liver were rarely included in these studies. Only one study reported the association between *PNPLA3* rs738409 and HCC in a case-control study including 25 patients with HCC developed on non-cirrhotic liver (F0-F2) (Donati et al., 2017). For patients with other etiologies, like CHC, there were a lot of inconsistent results in either Caucasian or non-Caucasian populations (Balasus et al., 2016; Ezzikouri et al., 2014; Guyot et al., 2013; Hai et al., 2017; Nakaoka et al., 2015; Nischalke et al., 2011; Valenti et al., 2011). Even though the association between *PNPLA3* rs738409 and NAFLD was widely validated, there is not yet a well-defined conclusion for its association with risk of NAFLD-related HCC due to limited number of cases in limited number of case-control studies (Donati et al., 2017; Y. L. Liu, Patman, et al., 2014; Seko et al., 2017).

| Author and year | Study design | Population | Etiology | Liver disease | HCC cases number | Control number | p-value |
|---------------------|---------------|------------|------------------|---------------|------------------|----------------|----------|
| Nischalke 2011 | retrospective | Caucasian | ALD | Cirrhosis | 80 | 80 | 0.013 |
| | | | HCV | Cirrhosis | 81 | 81 | ns |
| Valenti 2011 | retrospective | Caucasian | HCV | Cirrhosis | 50 | 275 | 0.002 |
| Trepo 2012 | retrospective | Caucasian | ALD | Cirrhosis | 145 | 426 | 1.83E-07 |
| Guyot 2013 | prospective | Caucasian | ALD | Cirrhosis | 66 | 213 | 0.002 |
| | | | HCV | Cirrhosis | 93 | 160 | ns |
| Hassan 2013 | retrospective | Caucasian | Mixed | Mixed | 257 | 494H | <0.05 |
| Friedrich 2014 | retrospective | Caucasian | ALD | Cirrhosis | 29 | 76 | 0.008 |
| Liu 2014 | retrospective | Caucasian | NAFLD | Mixed | 100 | 275 | 0.0001 |
| Ezzikouri 2014 | retrospective | African | HCV | Cirrhosis | 101 | 129 | 0.004 |
| Khlaiphuengsin 2015 | retrospective | Asian | Mixed | Mixed | 388 | 200 | <0.05 |
| Nakaoka 2015 | retrospective | Japanese | HCV | Cirrhosis | 48 | 183 | 0.02 |
| Falleti 2016 | retrospective | Caucasian | ALD | Cirrhosis | 75 | 151 | 0.03 |
| | | | HCV/HBV patients | Cirrhosis | 75 | 210 | ns |
| Ueyama 2016 | retrospective | Japanese | T2D | Mixed | 59 | 198 | 1.05E-05 |
| Balalus 2016 | retrospective | Caucasian | HCV | Cirrhosis | 170 | 304H | 0.006 |
| Casper 2016 | retrospective | Caucasian | Mixed | Mixed | 88 | 348H | 3.30E-04 |
| Seko 2017 | retrospective | Japanese | NAFLD | Mixed | 10 | 238 | 0.02 |
| Hai 2017 | retrospective | Japanese | HCV | Mixed | 142 | 575 | ns |
| Stickel 2018 | retrospective | Caucasian | ALD | Cirrhosis | 751 | 1165 | 5.13E-05 |

Table 5: Genetic association between *PNPLA3* rs738409 (G) and HCC risk according to the type of underlying liver disease

H: healthy population; *ns*: non-significant; *T2D*: type 2 diabetes.

TM6SF2: Loss-of-function genetic variant of *TM6SF2* was associated with increased hepatic triglyceride content and risk of NAFLD in a GWAS. Like *PNPLA3*, *TM6SF2* is also involved in lipid metabolism. The first study compared the genotype distribution of *TM6SF2* rs58542926 in 99 patients with NAFLD-related HCC and 1074 patients with NAFLD but without HCC. The significant association with NAFLD-related HCC was found in univariate analysis, but not confirmed in multivariate analysis incorporating known risk factors of HCC such as age, gender, BMI, T2D and presence of cirrhosis (Y. L. Liu, Reeves, et al., 2014). In contrast, a second study confirmed an independent association between *TM6SF2* rs58542926 and NAFLD-related HCC after adjustment for age, sex, obesity, T2D and presence of advanced fibrosis (Donati et al., 2017). Another study investigated the association between *TM6SF2* rs58542926 and HCC in patients with liver cirrhosis. The authors found that *TM6SF2* rs58542926 was associated with risk of HCC in patients with ALD but not with chronic viral hepatitis (Falleti, Cussigh, Cmet, Fabris, & Toniutto, 2016).

MBOAT7: An intronic SNP in *MBOAT7* (rs626283 C>G) was associated with increased risk of ALD-related cirrhosis in a GWAS (Buch et al., 2015). Moreover, it

has been demonstrated that the *MBOAT7* rs641738 T allele was associated with reduced *MBOAT7* expression and may predispose to HCC in patients without cirrhosis.(Donati et al., 2017).

Moreover, several studies based on a candidate-gene strategy about HCC susceptibility were conducted and based on the knowledge about liver tumorigenesis. The SNPs analyzed belong to genes involved in modulation of oxidative stress and detoxifying systems, iron metabolism, inflammation and immune responses, DNA damage repair or in cell cycle regulation. Some of them have revealed genetic variants potentially associated with risk of HCC. They were summarized in **Table 6**. Unfortunately, due to the heterogeneous population, the case/control design, the monocentric aspect, many studies suffer from major methodological biases which induce false-positive association. Therefore, further validations in large cohort of well-defined and prospectively followed-up patients, and investigations of potential functions and mechanisms are needed. Other well-known inherited traits influenced by multigenic variants, such as T2D and obesity, are also strong risk factors of HCC. Considering this fact, these trait-associated genetic variants might be shared in the development of HCC. This issue remained to be investigated.

| Biological pathway | Main study(ies) | SNP | SNP | Aetiology | Type of study | Cases/controls | Odds ratio or hazard ratio | Ethnicity |
|-------------------------------------|-------------------------|-------------------------|------------|-----------------------|------------------------|----------------|----------------------------|---------------------|
| Oxidative stress/iron metabolism | Wang et al., 2010 | GSTM1 | Deletion | HBV/HCV | Meta-analysis (n = 24) | 3349/5609 | OR = 1.26 | Mixed |
| | | GSTT1 | Deletion | HBV/HCV | Meta-analysis (n = 19) | 2884/4898 | OR = 1.28 | Mixed |
| | Nahon et al., 2009 | SOD2 A16V | rs4880 | Alcohol | Prospective | 190 | HR = 1.70 | Caucasian |
| | | MPO G-463A | rs2333227 | Alcohol | Prospective | 190 | HR = 3.80 | Caucasian |
| | Nahon et al., 2011 | | | HCV | Prospective | 205 | HR = 2.80 | Caucasian |
| | | CAT T-262C | rs1001179 | HCV | Prospective | 205 | HR = 1.74 | Caucasian |
| | Nahon et al., 2008 | HFE C282Y | rs1800562 | Alcohol | Prospective | 165 | HR = 2.70 | Caucasian |
| | Zhu et al., 2013 | GRP78 | rs391957 | Mixed | retrospective | 1007/810 | OR = 1.24 | Asian |
| | Sakurada et al., 2018 | MUTYH | rs3219487 | HCV | retrospective | 20/20 | OR = 9.27 | Asian |
| | Chen et al., 2013 | PPP2R1A | rs11453459 | Mixed | retrospective | 251/252 | OR = 0.52 | Asian |
| Wu et al., 2014 | ATF6 | rs2070150 | HBV | retrospective | 816/1082 | OR = 0.66 | Asian | |
| Jin et al., 2010 | HFE C282Y | rs1800562 | Alcohol | Meta-analysis (n = 9) | 1102/3766 | OR = 4.06 | Caucasian | |
| Detoxifying systems | Wang et al., 2004 | UGT1A7 R131K | rs17868323 | HCV | Case-control | 122/158 | OR = 2.33 | Asian |
| | Tseng et al., 2005 | N129K | rs1692021 | HBV/HCV | Case-control | 217/291 | OR = 3.06 | Asian |
| | Imaizumi et al., 2009 | CYP1A2 G-3860A | rs2069514 | HBV/HCV | Case-control | 205/652 | OR = 4.08 | Asian |
| Inflammation | Wei et al., 2011 | TNF G-308A | rs1800629 | HBV/HCV | Meta-analysis (n = 14) | 1835/2207 | OR = 1.80 | Mixed |
| | Yang et al., 2011 | | | HBV/HCV | Meta-analysis (n = 20) | 2763/4152 | OR = 1.74 | Mixed |
| | Wei et al., 2011 | TNF α G-238A | rs361525 | HBV/HCV | Meta-analysis (n = 8) | 938/1370 | OR = 1.62 | Mixed |
| | | TNF α C-863A | rs1800630 | HBV/HCV | Meta-analysis (n = 5) | 627/1004 | OR = 1.71 | Mixed |
| | Wang et al., 2003 | IL18 C-511T | rs16944 | HCV/Alcohol | Case-control | 125/149 | OR = 1.51 | Asian |
| | | IL1 β C-31T | rs143627 | HCV/Alcohol | Case-control | 125/149 | OR = 3.73 | Asian |
| | Wei et al., 2011 | IL10 A-592C | rs1800872 | HBV/HCV | Meta-analysis (n = 7) | 1012/2308 | OR = 1.29 | Mixed |
| | Migita et al., 2005 | TGF β 1 | rs13447341 | HCV | Case-control | 48/138 | OR = 2.77 | Asian |
| | Charni et al., 2011 | RANTES G-403A | rs2107538 | Alcohol | Prospective | n = 253 | HR = 2.70 | Caucasian |
| | He et al., 2009 | NFKBIA G881A | rs3138053 | HBV | Case-control | 202/886 | OR = 3.53 | Asian |
| | Han et al., 2014 | miR-218 | rs11134527 | HBV | retrospective | 1021/990 | OR = 1.21 | Asian |
| | Wang et al., 2017 | CCL4 | rs10491121 | Mixed | retrospective | 346/1200 | OR = 0.67 | Asian |
| | Asahina et al., 2014 | IL-28 | rs8099917 | HCV | prospective | 792 | NA | Asian |
| | Hu et al., 2016 | NTCP | rs2296651 | HBV without cirrhosis | retrospective | 370/2789 | OR = 0.65 | Asian |
| | Kim et al., 2013 | interferon gamma | rs2430561 | HBV | retrospective | 170/171 | NA | Asian |
| | Fatemipour et al., 2017 | STAT3 | rs1053004 | HBV | retrospective | 33/50 | NA | Iranian |
| | Lio et al., 2019 | TLL1 | rs17047200 | HCV | prospective | 1010 | HR = 2.8 | Asian |
| Zhang et al., 2014 | NFKBIA | rs2233406 | HBV | retrospective | 1495/1481 | OR = 1.33 | Asian | |
| DNA synthesis and repair mechanisms | Yuan et al., 2007 | MTHFR C-677T | rs1801133 | HBV/HCV/Alcohol | Case-control | 365/467 | OR = 2.17 | Mixed |
| | | MTHFR A1298C | rs1801131 | HBV/HCV/Alcohol | Case-control | 365/467 | OR = 2.17 | Mixed |
| | | TYMS 3' UTR1949del6 | rs16430 | HBV/HCV/Alcohol | Case-control | 365/468 | OR = 2.17 | Mixed |
| | Long et al., 2010 | XPC L-939G | rs2228001 | HBV/HCV | Case-control | 1156/1402 | OR = 1.81 | Asian |
| | Jung et al., 2014 | TEP1 | rs1713449 | HBV | retrospective | 702/351 | OR = 1.42 | Asian |
| | Li et al., 2016 | TERT | rs2736098 | NS | meta n=3 | 1901/1954 | OR = 1.38 | Asian |
| | Zhu et al., 2017 | RYBP | rs12956 | HBV | retrospective | 1100/1100 | OR = 0.72 | Asian |
| Long et al., 2008 | XRCC3 C-1806TT | rs861539 | HBV/HCV | Case-control | 1100/1100 | OR = 1.53 | Asian | |
| Cell cycle regulation | Yoon et al., 2008 | | | HBV | Case-control | 491/862 | OR = 7.19 | Asian |
| | Dharef et al., 2006 | MDM2 G-309T | rs2279744 | HBV | Case-control | 287/296 | OR = 4.89 | Asian |
| | | | | HCV | Case-control | 187/203 | OR = 2.27 | Asian |
| | Jin et al., 2011 | | | HBV/HCV | Meta-analysis (n = 5) | 738/1062 | OR = 1.57 | Mixed |
| | Liu et al., 2013 | p21 | rs1801270 | Mixed | retrospective | 476/526 | OR = 1.50 | Asian |
| | p27 | rs34330 | Mixed | retrospective | 476/526 | OR = 1.42 | Asian | |
| Yoon et al., 2008 | P53 R72P | rs1042522 | HBV | Case-control | 287/296 | OR = 3.03 | Asian | |
| Hormone metabolism | Zhai et al., 2006 | ESR1 T29C | rs2077647 | HBV | Case-control | 248/239 | OR = 2.31 | Asian |
| | | ESR1 (TA)n repeat | rs3138774 | | | | OR = 2.66 | |
| | | ESR1 PvuII CG | rs2234693 | | | | OR = 2.19 | |
| Growth factors | Abu Dayyeh et al., 2011 | EGF A61G | rs4444903 | HCV | Case-control | 816 | HR = 2.10 | Caucasian |
| | Tanabe et al., 2008 | | | HBV/HCV/Alcohol | Case-control | 103/225 | OR = 3.49 | Afro-American Mixed |
| Micro-RNAs | Gao et al., 2009 | IL1 α TTCA Indel | rs3783553 | HBV/HCV/Alcohol | Case-control | 1477/1673 | OR = 1.61 | Asian |
| | Tian et al., 2017 | miR-196a-2 | rs11614913 | HBV/HCV | meta n=9 | 3217/4163 | OR = 0.85 | Asian |
| | Xu et al., 2008 | miR-146a GC | rs2910164 | HBV/HCV | Case-control | 479/504 | OR = 2.01 | Asian |
| Immune response | Li et al., 2009 | CD24 C170T | rs8734 | HBV | Prospective | 383 | NA | Asian |
| | Sheng et al., 2011 | | | HBV | Case-control | 235/268 | OR = 2.96 | Asian |
| Cell death | Alsawalay et al., 2018 | DR4 | rs20575 | HCV | retrospective | 80/80 | OR = 2.01 | Egyptian |
| | Liu et al., 2017 | CASP8 | rs3834129 | HBV | retrospective | 600/600 | OR = 0.71 | Asian |
| | Khalifa et al., 2016 | FasL | rs763110 | HBV | retrospective | 100/100 | OR = 1.97 | Egyptian |
| | Jin et al., 2015 | TNF- α | rs1799724 | HBV with cirrhosis | prospective | 430 | HR = 6.29 | Asian |
| | Wang et al., 2016 | GLB1 | rs4678680 | HBV | retrospective | 1186/508 | OR = 1.52 | Asian |

Table 6: SNPs associated with HCC in candidate-gene studies or meta-analysis.

OR and HR take into account the highest reported value according to allelic or genotypic expression of results and, when available, after multivariate analysis; NA: not available

Results

Context and aim of the project

HCC is among the most lethal and prevalent cancer worldwide. Patients with HCC diagnosed at early stage could achieve curative treatment, while in current practice nearly half of HCC are diagnosed at an intermediate or advanced tumor stage with a very poor prognosis. Most HCC develops on chronic liver diseases (CLD), usually liver cirrhosis. Due to individual genetic heterogeneity and different environmental risk exposures, patients underlying CLD have different risk to develop HCC. Therefore, hope remains in the identification of novel and promising genetic markers for better definition and management of surveillance and chemoprevention in these populations.

One aim of the laboratory is to identify the genetic markers that could predispose the risk of HCC. We use two different methods to study genetic markers: GWAS and candidate-gene study. My project focused on candidate-gene study based on a selection of a panel of SNPs. Most of them were closely associated with the risk of different CLDs or liver cirrhosis. The question was whether they contribute to risk of HCC progression with the underlying CLD.

The main strategy of our study was to compare patients with or without HCC taken into account the various etiologies and different severity of their underlying CLDs. Several of them were previously associated with risk of HBV- or HCV-related HCC of genome-wide significance in Asians, requiring a validation in Europeans due to difference in ethnicity and some bias of study design. Thanks to all our collaborators in hospitals across France and Belgium, we collected a large number of patients with CLD with or without HCC. In addition, two prospective cohorts of patients followed up for HCC occurrence were analyzed to study the predictive effect of SNPs.

Brief summarization of the current results:**Article 1: “*PNPLA3* and *TM6SF2* variants as risk factors of hepatocellular carcinoma across various etiologies and severity of underlying liver diseases”**

(International Journal of Cancer, 2018)

Jie Yang, Eric Trépo, Pierre Nahon, Qian Cao, Christophe Moreno, Eric Letouzé, Sandrine Imbeaud, Thierry Gustot, Jacques Deviere, Stéphanie Debette, Philippe Amouye, Paulette Bioulac-Sage, Julien Calderaro, Nathalie Ganne-Carrié1, Alexis Laurent, Jean Frédéric Blanc, Erwan Guyot, Angela Sutton, Marianne Ziol*, Jessica Zucman-Rossi*, Jean-Charles Nault*

PNPLA3 rs738409 and *TM6SF2* rs58542926 are two SNPs closely associated with the risk of NAFLD- and ALD-related cirrhosis. Their predisposing effect on the risk of HCC needs to be well-validated in a prospective manner. Moreover, several SNPs have been identified associated with risk of HCC in GWAS performed in Asia. However, few of them were cross-validated, requiring further investigation in European population.

The goal of our study was to test the association between SNPs previously identified in GWAS and HCC development taking into account the heterogeneity of etiologies and the severity of underlying CLDs. It allowed a better assessment of the role of candidate genetic variants in liver carcinogenesis. We firstly genotyped these 9 SNPs in a retrospective cohort including 1020 patients with HCC and 2021 patients with only CLD. The genotype distributions of these SNPs were compared between cases and controls taken into account the etiology and severity of underlying liver diseases. Among the 7 SNPs from GWAS of HCC, only *STAT4* rs7574865 was re-validated as genetic risk factor of HBV-related HCC. *PNPLA3* and *TM6SF2* variants were both independently associated with HCC in patient with ALD, but not with other etiologies. Due to limited number of cases in our study, the association between *PNPLA3/ TM6SF2* variants and NAFLD-related HCC needs further investigations. Strikingly, we observed an association between *PNPLA3* rs738409 and HCC developed on non-fibrotic liver, suggesting a direct role of

PNPLA3 in tumorigenesis. In addition, *PNPLA3* and *TM6SF2* variants were confirmed as associated with HCC incidence in a dose-dependent manner in a prospective cohort of patients with ALD-related cirrhosis.

In conclusion, *PNPLA3* rs738409 and *TM6SF2* rs58542926 are risk factors of HCC in patients with ALD.

PNPLA3 and TM6SF2 variants as risk factors of hepatocellular carcinoma across various etiologies and severity of underlying liver diseases

Jie Yang¹, Eric Trépo^{1,2}, Pierre Nahon^{1,3,4}, Qian Cao⁵, Christophe Moreno², Eric Letouzé¹, Sandrine Imbeaud¹, Thierry Gustot², Jacques Deviere², Stéphanie Debette^{6,7}, Philippe Amouyel⁸, Paulette Bioulac-Sage^{9,10}, Julien Calderaro^{11,12}, Nathalie Ganne-Carrié^{1,3,4}, Alexis Laurent^{13,14}, Jean Frédéric Blanc¹⁵, Erwan Guyot^{16,17}, Angela Sutton^{16,17}, Marianne Ziol^{1,4,18}, Jessica Zucman-Rossi^{1,19} and Jean-Charles Nault^{1,3,4}

¹Inserm UMR-1162, Génomique fonctionnelle des Tumeurs solides, Université Paris Descartes, Université Paris Diderot, Université Paris 13, Labex Immuno-Oncology, Paris, France

²Department of Gastroenterology, Hepatopancreatology and Digestive Oncology, Cliniques Universitaires de Bruxelles Hôpital Erasme, Université Libre de Bruxelles, Bruxelles, Belgium

³Liver unit, Hôpital Jean Verdier, Hôpitaux Universitaires Paris-Seine-Saint-Denis, Assistance-Publique Hôpitaux de Paris, Bondy, France

⁴Unité de Formation et de Recherche Santé Médecine et Biologie Humaine, Université Paris 13, Communauté d'Universités et Etablissements Sorbonne Paris Cité, Paris, France

⁵Department of Infectious Diseases, Zhongnan Hospital of Wuhan University, Wuhan, China

⁶Univ. Bordeaux, Inserm, Bordeaux Population Health Research Center, Bordeaux, France

⁷CHU de Bordeaux, Department of Neurology, Bordeaux, France

⁸University of Lille, Institut National de la Santé et de la Recherche Médicale, CHU Lille, Institut Pasteur de Lille, Lille, France

⁹Univ. Bordeaux, UMR1053 Bordeaux Research in Translational Oncology, Bordeaux, France

¹⁰Service de Pathologie, Hôpital Pellegrin, CHU de Bordeaux, Bordeaux, France

¹¹Service d'anatomopathologie, Hôpital Henri Mondor, Créteil

¹²Université Paris Est Créteil, Inserm U955, Team 18, Institut Mondor de Recherche Biomédicale, France

¹³Service de chirurgie digestive, Hôpital Henri Mondor, Créteil

¹⁴Université Paris Est Créteil, Institut Mondor de Recherche Biomédicale, France

¹⁵Service Hépato-Gastroentérologie et oncologie digestive, Centre Medico-Chirurgical Magellan, Hôpital Haut-Lévêque, CHU de Bordeaux, Bordeaux, France

¹⁶Laboratoire de biochimie, Hôpital Jean Verdier, Hôpitaux Universitaires Paris-Seine-Saint-Denis, Assistance-Publique Hôpitaux de Paris, Bondy, France

Key words: single nucleotide polymorphisms, STAT4, cirrhosis, predisposition, liver cancer

Abbreviations: ALD: alcoholic liver disease; ARID1A: AT-rich interaction domain 1A; ARID2: AT-rich interaction domain 2; BMI: body mass index; CI: confidence interval.; CLD: chronic liver disease; CTNNB1: catenin beta 1; DDX18: DEAD-Box Helicase 18; DEPDC5: DEP domain-containing 5; DLC1: deleted in liver cancer 1; GRIK1: glutamate ionotropic receptor kainate type subunit 1; GWAS: genome-wide association study; HBV: hepatitis B virus; HCC: hepatocellular carcinoma; HCV: hepatitis C virus; HWE: Hardy-Weinberg equilibrium; KEAP1: kelch-like ECH-associated protein 1; KIF1B: kinesin family member 1B; MAF: minor allele frequency; MICA: MHC class I polypeptide-related sequence A; NAFLD: nonalcoholic fatty liver disease; NFE2L2: nuclear factor (erythroid-derived 2)-like 2; OR: odds ratio; PNPLA3: patatin-like phospholipase domain-containing 3; RPS6KA3: ribosomal protein S6 kinase A3; RT-PCR: reverse transcription polymerase chain reaction; SNP: single nucleotide polymorphisms; STAT4: signal transducer and activator of transcription 4; TERT: telomerase reverse transcriptase; TM6SF2: transmembrane 6 superfamily 2; TP53: tumor protein p53

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: Christophe Moreno declared personnel fees from Bayer. Other authors have no conflict of interest to declare.

Authors Marianne Ziol, Jessica Zucman-Rossi, and Jean-Charles Nault shared the co-authorship.

Grant sponsor: HECAM project (BPI); **Grant sponsor:** International Cancer Genome Consortium ; **Grant sponsor:** MUTHEC projects ;

Grant sponsor: NoFLIC inca; **Grant sponsor:** Fondation Mérieux; **Grant sponsor:** SIRIC CARPEM; **Grant sponsor:** Coup d'Élan de la Fondation Bettencourt-Shueller; **Grant sponsor:** Labex OncoImmunology; **Grant sponsor:** Ligue Nationale contre le Cancer; **Grant sponsor:** INSERM; **Grant sponsor:** Institut National du Cancer (INCa)

DOI: 10.1002/ijc.31910

History: Received 16 Jul 2018; Accepted 10 Sep 2018; Online 5 Oct 2018

Correspondence to: Jean-Charles Nault, Service d'hépatologie, hôpital Jean Verdier Avenue du 14 juillet 93140 Bondy, Tel.: +33 6 10 67

94 61, Fax: +33 1 53 72 51 92, E-mail: naultjc@gmail.com; or Jessica Zucman-Rossi; MD, PhD INSERM U 1162, Génomique fonctionnelle des tumeurs solides 27 Rue Juliette Dodu 75010 Paris, France, Tel.: +33 1 53 72 51 66, Fax: +33 1 53 72 51 92, E-mail: jessica.zucman-rossi@inserm.fr

¹⁷INSERM U1148 LVTS, UFR SMBH, Université Paris 13, PRES Paris Sorbonne Cité, Bobigny, France

¹⁸Centre de Ressources Biologiques (BB-0033-00027) Hôpitaux Universitaires Paris-Seine-Saint-Denis, Assistance-Publique Hôpitaux de Paris, Bondy, France

¹⁹Hôpital Européen Georges Pompidou, Paris, France

Few single nucleotide polymorphisms (SNPs) have been reproducibly associated with hepatocellular carcinoma (HCC). Our aim was to test the association between nine SNPs and HCC occurrence. SNPs in genes linked to HCC (*DEPDC5*, *GRIK1*, *KIF1B*, *STAT4*, *MICA*, *DLC1*, *DDX18*) or to liver damage (*PNPLA3*-rs738409, *TM6SF2*-rs58542926) in GWAS were genotyped in discovery cohorts including 1,020 HCC, 2,021 controls with chronic liver disease and 2,484 healthy individuals and replication was performed in prospective cohorts of cirrhotic patients with alcoholic liver disease (ALD, $n = 249$) and hepatitis C ($n = 268$). In the discovery cohort, *PNPLA3* and *TM6SF2* SNPs were associated with HCC (OR = 1.67 [CI95%:1.16–2.40], $p = 0.005$; OR = 1.45 [CI95%:1.08–1.94], $p = 0.01$) after adjustment for fibrosis, age, gender and etiology. In contrast, *STAT4*-rs7574865 was associated with HCC only in HBV infected patients ($p = 0.03$) and the other tested SNP were not linked with HCC risk. *PNPLA3* and *TM6SF2* variants were independently associated with HCC in patients with ALD (OR = 3.91 [CI95%:2.52–6.06], $p = 1.14 \times 10^{-9}$; OR = 1.79 [CI95%:1.25–2.56], $p = 0.001$) but not with other etiologies. *PNPLA3* SNP was also significantly associated with HCC developed on a nonfibrotic liver (OR = 2.19 [CI95%:1.22–3.92], $p = 0.007$). The association of *PNPLA3* and *TM6SF2* with HCC risk was confirmed in the prospective cohort with ALD. A genetic score including *PNPLA3* and *TM6SF2* minor alleles showed a progressive significant increased risk of HCC in ALD patients. In conclusion, *PNPLA3*-rs738409 and *TM6SF2*-rs58542926 are inherited risk variants of HCC development in patients with ALD in a dose dependent manner. The link between *PNPLA3* and HCC on nonfibrotic liver suggests a direct role in liver carcinogenesis.

What's new?

Multiple single nucleotide polymorphisms (SNPs) have been associated with hepatocellular carcinoma (HCC) through genome-wide association studies (GWAS). Few of these variants, however, have been cross-validated, raising questions about their relation to HCC. Here, nine SNPs previously linked to HCC or liver damage were selected for investigation in European patients. Two of the variants, *PNPLA3* rs738409 and *TM6SF2* rs58542926, were found to be associated with HCC risk specifically in patients with alcoholic liver disease. *PNPLA3* rs738409 was also associated with HCC in patients with nonfibrotic livers, where it potentially influences liver carcinogenesis *via* modification of protein function without altering gene expression.

Introduction

Hepatocellular carcinoma (HCC) ranks as the fifth most common cancer worldwide with 9.1% of global cancer mortality. In 90% of the cases, HCC arises on chronic liver injury, mainly on cirrhosis, due to various etiologies that have heterogeneous distribution worldwide. In Eastern countries, chronic HBV infection is the main cause of chronic liver disease in contrast with Western countries where chronic hepatitis C infection, high alcohol consumption and nonalcoholic fatty liver disease (NAFLD) due to metabolic syndrome are the most prevalent etiologies.¹ However, only a subset of these patients develops HCC, suggesting the role of potential genetic modifiers in the occurrence of HCC in chronic liver disease. The association between a family history of HCC and a higher risk of HCC in HBV infected patients in Asia but also in Western patients with chronic HCV infection supports the role of genetic predisposition in liver carcinogenesis.^{2,3}

Several genome-wide association studies (GWAS) using a case-control design were performed in the field of HCC and have identified single nucleotide polymorphisms (SNP)

associated with HCC risk: *DDX18* rs2551677, *KIF1B* rs17401966, *DEPDC5* rs1012068, *MICA* rs2596542, *GRIK1* rs455804 and *STAT4* rs7574865.^{4–10} All of these studies were conducted in Asian populations with chronic viral infection (HBV or HCV) but none of these SNPs was cross-validated in another GWAS. Moreover, several of these GWAS used only healthy individuals as controls and most of them did not take into account the fibrosis stage reflecting the severity of liver diseases. These major drawbacks introduce bias leading to inconclusive results whether these SNPs were related to viral infection, fibrosis severity and/or HCC development.

Importantly, additional GWAS have identified a nonsynonymous SNP rs738409 (C > G; p. I148M) in patatin-like phospholipase domain-containing 3 (*PNPLA3*) and a SNP rs58542926 (C > T; p. E167K) of transmembrane 6 superfamily 2 (*TM6SF2*), two genes involved in lipid metabolism, as associated with development of NAFLD.^{11–14} *PNPLA3* rs738409 has also been associated with HCC in case-control study and prospective cohorts in alcohol-related cirrhosis whereas the role in HCC development on HCV related cirrhosis or on

NAFLD remains controversial in the literature.^{15–20} *TM6SF2* rs58542926 was associated with HCC occurrence in patients with alcoholic cirrhosis but few data were available in other etiologies.^{13,21–23} Moreover, whether these aforementioned variants predispose to HCC developed on a nonfibrotic liver is currently unknown. Individuals who develop HCC on a nonfibrotic liver define a rare but important subgroup of patients without liver disease, which is helpful to identify inherited variants related to HCC development outside the background of chronic liver injury.²⁴

Using both case–control and prospective design with a large number of patients, we aimed to test the association between genetic variants previously identified in GWAS and HCC development taking into account the heterogeneity of etiologies and the severity of underlying liver disease allowing a better assessment of the role of candidate genetic variants in liver carcinogenesis.

Materials and Methods

Selection of the population

We included 1,020 European HCC patients and 3 control populations without HCC in the retrospective case–control study: (1) 2021 control patients with chronic liver disease (2) 503 Europeans from 1000Genomes^{25,26} and (3) 1981 individual from the general French population (3city population).^{27,28} The study was approved by the Ethics Committee (IRB HECAM IRB00003888), and informed consent was obtained for all analyzed individuals according to the French laws.

HCC patients were recruited in France, all were diagnosed either using histological analysis or noninvasive criteria in cirrhotic patients.²⁹ Liver fibrosis was assessed histologically, and classified, according to the METAVIR staging,³⁰ as nonfibrotic liver (F0–F1), fibrotic liver (F2–F3) and cirrhosis (F4). Chronic hepatitis C infection was defined by positive anti HCV serology with positive RNA at PCR at the diagnosis or at any time of the history of the patient. Chronic hepatitis B infection was defined by a positive AgHBs. Chronic alcohol intake was defined as the main cause of liver disease in patients with persistent high alcohol intake (>80 g per day in males and > 60 g per day in females for at least 10 years) and without chronic HBV and HCV infection. NAFLD was defined as the main cause of liver disease in patients with metabolic syndrome and without high alcohol intake, chronic HBV infection and chronic HCV infection.

In the control groups, we included patients with chronic liver disease of various etiology and without HCC from France ($n = 991$) and Belgium ($n = 1,030$). Etiologies of the patients were defined as HCC patients and included chronic hepatitis C infection, chronic hepatitis B infection, alcohol intake, and NAFLD. Considering that 212 HCC patients were without any evident fibrosis, we used the 503 general Europeans from 1000 Genome Project with accessible genotyping data as control. We also included in our analysis a second control of

1981 European individuals from the general French population randomly selected from the “3city cohort” designed for a population-based longitudinal study of the relation between vascular diseases and dementia in persons aged 65 years and older. These individuals were recruited in three French cities (Bordeaux, Dijon and Montpellier).

We also analyzed two prospective cohorts of patients with alcohol- (249 patients) or HCV- (268 patients) related cirrhosis. Results of *PNPLA3* genotyping and the association with HCC incidence in these 2 cohorts have been already published.¹⁶ The present analyses are focused on the longitudinal impact of *TM6SF2* rs58542926 and its potential association with *PNPLA3* on HCC occurrence in these cohorts. All patients were Caucasian and had a histologically proven cirrhosis diagnosed by liver biopsy between January 1999 and June 2010. Clinical features were recorded at baseline: age, gender, liver blood tests, platelets, body mass index or diabetes type 2. HCV-infected patients were part of a “historical” cohort of patients followed-up before direct antivirals became commercially available in France. All patients were included in HCC surveillance programs which comprised a semi-annual ultrasound examination.

Genotyping of the single nucleotide polymorphisms

We extracted DNA samples either from frozen blood samples or from frozen nontumor liver of patients with Maxwell tissue DNA purification kit (Promega). After quantification and quality assessment, DNA samples were pre-amplified to normalize the concentration of each allele. A randomly selected sample was included in every plate to ensure consistent genotype. According to the manufacturer’s protocol, we used the Fluidigm technology to discriminate the genotype of the after SNP: *DEPDC5* rs1012068, *GRIK1* rs455804, *KIF1B* rs17401966, *STAT4* rs7574865, *MICA* rs2596542, *DLC1* rs2275959 and *DDX18* rs2551677. The Taqman probes used were described in Supporting Information Table 1. Genotyping analysis was performed using Fluidigm SNP genotyping analysis software (4.1.3). To ensure quality control, genotyping cluster plots have been inspected visually and samples which failed >5% of assays have been removed from analysis. Besides, we also confirmed the genotype of 30 patients with Exome sequencing data already available. For the 503 Europeans and 1981 general French population, genotypes of these 9 SNP were extracted from the database of 1000Genomes project and 3City study.

Molecular analysis of HCC

To analyze the correlation between SNP and histological and molecular tumor features, we only included 563 patients with a unique HCC treated by liver resection. We search for somatic mutations in 9 driver genes as previously described³¹: *TERT* promoter, *CTNNB1*, *TP53*, *ARID1A*, *ARID2*, *AXIN1*, *RPS6KA3*, *NFE2L2* and *KEAP1*.

Quantitative RT-PCR was performed to assess mRNA expression of 16 genes in order to predict the G1-G6 molecular classification in HCC as previously described.³² Moreover, we assess expression of *PNPLA3* and *TM6SF2* by quantitative RT-PCR in a panel of 340 normal liver using the after Taqman probes: Hs00228747_m1 and Hs00403495_m1. We performed reverse transcription of 500 ng of total RNA using the High-Capacity Transcription Kit (Life Technologies) and assess gene expression using TaqMan predesigned assays (Life Technologies) on Fluidigm 96 dynamic arrays with the BioMark Real-Time PCR system. Expression level (Ct values) was assessed using the Fluidigm Real-Time PCR Analysis software (4.1.3) and gene expression data were expressed using the 2DDCt method relative to ribosomal 18S (R18S) and the expression level of the corresponding gene in normal liver samples.

Statistical analysis

Comparison between categorical variable was performed using the Chi-square test and between continuous variable using the Kruskal-Wallis test. In the case-control analyses, we adjusted our results using few selected well-known risk factors related to HCC risk in the literature and available in our database. After this strategy, we introduced only a limited number of features (age, Gender, degree of fibrosis in the nontumor liver, and etiology) in the multivariate analyses. Adjustment for potential known confounder was performed using a multivariate logistic regression analysis. In the analysis, we compared the homozygote genotype of the minor allele vs. homozygote genotype of the major allele and the heterozygote genotype vs. the homozygote genotype of the major allele except for *STAT4* and *TM6SF2* genotype where we combined the homozygote genotype of minor allele with the heterozygote genotype for the analysis because the number of patients with the homozygote genotype of minor allele was low.

Patients from the 2 prospective cohorts were screened every 6 months for HCC development using liver ultrasonography. The date of inclusion was the date of the first liver biopsy that showed cirrhosis. HCC were diagnosed either using histological analysis or noninvasive criteria using contrast enhanced computed tomodensitometry, and/or magnetic-resonance imaging in cirrhotic patients.²⁹ The main end point was the occurrence of HCC during follow-up. Follow-up ended at the date of death or liver transplantation (considered as death), or at the last recorded visit (or information taken) within the previous 6 months before June 2012. Survival curves were performed using the Kaplan Meier method with the log-rank test. Adjustment for potential known confounders (gender, age, BMI and diabetes) was performed using a multivariate Cox regression model. SNPs were in Hardy-Weinberg equilibrium except for *MICA* in patients with chronic liver disease ($p = 0.01$) and for *PNPLA3* in the prospective cohort of HCV related cirrhosis ($p < 0.001$). A two-tailed p value < 0.05 was considered as significant. Statistical analysis was performed using the GraphPad Prism software and the R software (<https://www.r-project.org/>).

Results

Clinical features of the case-control populations

A total of 5,525 subjects were analyzed in the case-control part of our study, including 1,020 patients with HCC, 2021 patients with chronic liver diseases (CLD) but without HCC and 2,484 healthy subjects from the European population ($n = 503$ from the 1000genomes project) and from the French population ($n = 1981$ from the 3City project) (See Supporting Information Fig. 1 for the flow chart). The characteristics of the HCC patients and control CLD population are described in Table 1. Patients with HCC ($n = 1,020$) have a median age of 64 years old, with 83% of male; the main etiologies of underlying liver diseases were alcohol intake (35%), HCV (25%), NAFLD (12%), HBV (9%) and 51% of the patients had

Table 1. Demographic characteristics of the study subjects

| Variables | | CLD Patients ($n = 2021$) | HCC Patients ($n = 1,020$) |
|-------------------------------|-----------------------------|-----------------------------|------------------------------|
| Age (years) ¹ | | 55 ± 11 | 64 ± 12 |
| Gender ² | Male | 1,383 (68%) | 849 (83%) |
| Etiology ² | Alcoholic liver disease | 1,222 (61%) | 360 (35%) |
| | Hepatitis B | 87 (4%) | 90 (9%) |
| | Hepatitis C | 480 (24%) | 256 (25%) |
| | NAFLD | 231 (11%) | 121 (12%) |
| | Without or other etiologies | - | 193 (19%) |
| Nontumoral liver ² | F0-F1 | 212 (11%) | 261 (26%) |
| Fibrosis according to METAVIR | F2-F3 | 193 (10%) | 235 (23%) |
| | F4 | 1,555 (79%) | 514 (51%) |

CLD: chronic liver disease; HCC: hepatocellular carcinoma, NAFLD: nonalcoholic fatty liver disease.

¹Mean ± SD.

²Number (percentage).

Table 2. List of SNPs that are associated with risk of HCC in GWAS and validation analysis in patients with or without HCC in our study

| SNP | Original GWAS | | | | Current study | | | | | | | |
|------------------|---------------|------------------|-------------|----------|---------------|------------------------------|----------------|--------------------|------------------|------|-------------------------|-------|
| | Etiology | Case and Control | Risk Allele | p Value | Ethnicity | Study | Genotype | Case and Control | OR (95%CI) | p | Multivariate OR (95%CI) | p |
| | | | | | | | | | | | | |
| DEPDC5 rs1012068 | HCV | 212/765 HCV | G | 8.05E-08 | Japanese | Miki et al. ⁵ | GT vs. TT | 256/480 HCV | 1.02 [0.74–1.40] | 0.94 | 1.47 [0.78–2.78] | 0.24 |
| GRIK1 rs455804 | HBV | 1538/1465 HBV | C | 1.65E-04 | Chinese | Li et al. ⁷ | GG vs. TT | 90/87 HBV | 0.85 [0.49–1.49] | 0.57 | 1.45 [0.78–2.72] | 0.24 |
| KIF1B rs17401966 | HBV | 355/360 HBV | A | 5.8E-06 | Chinese | Zhang et al. ⁹ | CA vs. AA | 90/87 HBV | 1.16 [0.34–3.98] | 1.00 | 1.46 [0.39–5.44] | 0.58 |
| STAT4 rs7574865 | HBV | 1161/1353 HBV | G | 1.53E-04 | Chinese | Jiang et al. ⁸ | CC vs. AA | 90/87 HBV | 1.82 [0.54–6.16] | 0.33 | 1.78 [0.49–6.48] | 0.38 |
| MICA rs2596542 | HCV | 721/2890 Healthy | T | 4.50E-06 | Japanese | Kumar et al. ⁶ | AG vs. GG | 256/480 HCV | 1.57 [0.38–6.44] | 0.72 | 1.13 [0.22–5.85] | 0.88 |
| DLC1 rs2275959 | HBV | 95/97 HBV | A | 6.35E-06 | Chinese | Li et al. ⁷ | AA vs. GG | 90/87 HBV | 1.18 [0.30–4.64] | 0.81 | 1.05 [0.21–5.29] | 0.95 |
| DDX18 rs2551677 | Cirrhosis | 386/86 Cirrhosis | A | 1.41E-10 | Asian | Clifford et al. ⁴ | GG vs. GT + TT | 514/1555 Cirrhosis | 2.07 [1.07–4.02] | 0.03 | 2.06 [1.00–4.24] | 0.049 |
| | | | | | | | TC vs. CC | | 1.10 [0.78–1.56] | 0.60 | 1.22 [0.83–1.79] | 0.30 |
| | | | | | | | TT vs. CC | | 1.05 [0.69–1.59] | 0.82 | 1.07 [0.70–1.70] | 0.78 |
| | | | | | | | AG vs. GG | | 0.63 [0.26–1.49] | 0.39 | 0.60 [0.23–1.55] | 0.29 |
| | | | | | | | AA vs. GG | | 0.65 [0.24–1.74] | 0.39 | 0.69 [0.24–1.99] | 0.50 |
| | | | | | | | AG vs. GG | | 0.98 [0.78–1.23] | 0.91 | 0.96 [0.70–1.33] | 0.80 |
| | | | | | | | AA vs. GG | | 1.00 [0.58–1.73] | 0.99 | 1.23 [0.58–2.59] | 0.59 |

Abbreviations: CLD, chronic liver disease; HCC, hepatocellular carcinoma; SNP, single nucleotide polymorphism. GG and GT genotypes of STAT4 were combined due to the low numbers of patients with GG genotype. Multivariate analysis was adjusted for age, gender and fibrosis. Significant p values are shown in bold type.

cirrhosis (F4 according to the METAVIR classification); 23% significant fibrosis (F2-F3) and 26% a nonfibrotic liver (F0-F1). In the control population with CLD and without HCC ($n = 2021$), the median age was 55 years old with 68% of male; the main etiologies of CLD were alcohol intake (61%), HCV (24%), NAFLD (11%), HBV (4%) and 79% of the patients have cirrhosis (F4 according to the METAVIR classification); 10% significant fibrosis (F2-F3) and 11% a non-fibrotic liver (F0-F1).

Genetic variants reported in previous GWAS as risk factor of HCC

We genotyped the seven SNP (*DEPDC5* rs1012068, *GRIK1* rs455804, *KIF1B* rs17401966, *STAT4* rs7574865, *MICA* rs2596542, *DLC1* rs2275959 and *DDX18* rs2551677) previously identified in GWAS as associated with HCC risk in non-Caucasian ancestry (Table 2). We firstly assessed the distribution of their genotype in HCC patients vs. CLD patients without HCC and did not identify any association with HCC risk even after adjustment for age, gender, etiology and severity of chronic liver disease (Supporting Information Table 3). Next, we performed subgroups analysis within the same etiology reported in the original GWAS and observed that only *STAT4* rs7574865 was linked to HCC risk (Table 2). In patients with chronic HBV infection, we found an increase frequency of the major allele (GG) of rs7574865 *STAT4* in HCC patients (78%) vs. CLD patients without HCC (63%, $p = 0.03$) confirming in European ancestry the association between *STAT4* rs7574865 and HCC risk in HBV patients as described in the initial GWAS (Table 2). The association was still significant ($p = 0.049$) after adjustment for age, gender and severity of liver disease.

PNPLA3 rs738409 and TM6SF2 rs58542926 and HCC risk according to the severity and etiologies of underlying liver disease

Firstly, we investigated the link between *PNPLA3* rs738409 (minor allele frequency (MAF) = 0.35 in CLD patients without HCC, GG vs. GC vs. CC) and *TM6SF2* rs58542926 (MAF = 0.08 in CLD patients without HCC, TT + CT vs. CC) and HCC risk comparing HCC patients vs. CLD patients. HCC patients have different genotype distribution of *PNPLA3* rs738409 (15% of GG) compared to CLD patients (12% of GG, $p = 0.009$) and, after adjustment for age, gender, etiologies and fibrosis stage, *PNPLA3* rs738409 was still associated with HCC risk (OR = 1.67 [CI95%: 1.16–2.40], $p = 0.005$) (Table 3). Similarly, we observed a higher frequency of *TM6SF2* TT + CT genotype in HCC patients (21%) compared to CLD patients (15%, $p = 9.0E-05$) and this difference remains significant after adjustment (OR = 1.45 [CI95%: 1.08–1.94], $p = 0.01$) (Table 3).

Next, we tested the association of *PNPLA3* rs738409 and *TM6SF2* rs58542926 in subgroup analysis taking into account fibrosis stage of the nontumor liver (Table 3). In patients with

Table 3. Genotype of PNPLA3 rs738409 and TM6SF2 rs58542926 in HCC cases/ control study according to severity of liver disease

| All patients | | | | | | | | |
|--|----------|--------------------------|-------------------------------|------------------------|-------------------------|-------------------------|----------------------------|-------------------------|
| Gene (SNP) | Genotype | CLD Patients (n = 2021) | HCC Patients (n = 1,020) | p Value | OR (95%CI) | p-adj# | OR (95%CI)-adj# | |
| PNPLA3 rs738409 | CC | 925 (46%) | 435 (43%) | - | 1 | - | 1 | |
| | CG | 856 (46%) | 433 (42%) | 0.38 | 1.08 [0.91–1.27] | 0.15 | 1.20 [0.93–1.53] | |
| | GG | 233 (12%) | 150 (15%) | 0.009 | 1.37 [1.08–1.73] | 0.005 | 1.67 [1.16–2.40] | |
| TM6SF2 rs58542926 | CC | 1,715 (85%) | 810 (79%) | - | 1 | - | 1 | |
| | CT + TT | 300 (15%) | 210 (21%) | 9.0E-05 | 1.48 [1.22–1.80] | 0.01 | 1.45 [1.08–1.94] | |
| Patients with cirrhosis (F4) | | | | | | | | |
| Gene (SNP) | Genotype | CLD Patients (n = 1,555) | HCC Patients (n = 514) | p Value | OR (95%CI) | p-adj* | OR (95%CI)-adj* | |
| PNPLA3 rs738409 | CC | 687 (44%) | 193 (38%) | - | 1 | - | 1 | |
| | CG | 661 (43%) | 226 (44%) | 0.08 | 1.22 [0.98–1.52] | 0.03 | 1.41 [1.03–1.92] | |
| | GG | 200 (13%) | 93 (18%) | 0.0007 | 1.66 [1.23–2.22] | 0.0005 | 2.13 [1.39–3.26] | |
| TM6SF2 rs58542926 | CC | 1,310 (84%) | 406 (79%) | - | 1 | - | 1 | |
| | CT + TT | 242 (16%) | 108 (21%) | 0.006 | 1.44 [1.12–1.85] | 0.04 | 1.45 [1.01–2.08] | |
| Patients with significant fibrosis (F2-F3) | | | | | | | | |
| Gene (SNP) | Genotype | CLD Patients (n = 193) | HCC Patients (n = 235) | p Value | OR (95%CI) | p-adj* | OR (95%CI)-adj* | |
| PNPLA3 rs738409 | CC | 97 (50%) | 108 (46%) | - | 1 | - | 1 | |
| | CG | 83 (43%) | 96 (41%) | 0.85 | 1.04 [0.70–1.55] | 0.87 | 0.95 [0.49–1.84] | |
| | GG | 13 (7%) | 31 (13%) | 0.03 | 2.14 [1.06–4.33] | 0.67 | 1.26 [0.43–3.75] | |
| TM6SF2 rs58542926 | CC | 167 (87%) | 178 (76%) | - | 1 | - | 1 | |
| | CT + TT | 25 (13%) | 57 (24%) | 0.005 | 2.14 [1.28–3.58] | 0.23 | 1.67 [0.72–3.89] | |
| Patients without fibrosis (F0-F1) | | | | | | | | |
| Gene (SNP) | Genotype | 1000Genome EUR (n = 503) | Healthy population (n = 1981) | HCC Patients (n = 261) | 1000G vs. HCC | | Healthy population vs. HCC | |
| | | | | | p Value | OR (95%CI) | p Value | OR (95%CI) |
| PNPLA3 rs738409 | CC | 303 (61%) | 1,134 (58%) | 128 (49%) | - | 1 | - | 1 |
| | CG | 173 (34%) | 685 (35%) | 108 (41%) | 0.02 | 1.48 [1.08–2.03] | 0.02 | 1.40 [1.06–1.84] |
| | GG | 27 (5%) | 130 (7%) | 25 (10%) | 0.007 | 2.19 [1.22–3.92] | 0.02 | 1.70 [1.07–2.71] |
| TM6SF2 rs58542926 | CC | 438 (87%) | 1,722 (87%) | 219 (84%) | - | 1 | - | 1 |
| | CT + TT | 65 (13%) | 254 (13%) | 42 (16%) | 0.28 | 1.29 [0.85–1.97] | 0.18 | 1.30 [0.91–1.86] |

Abbreviations: CLD, chronic liver disease; HCC, hepatocellular carcinoma; SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval. Analysis was performed using Chi-square test then # adjusted by gender, age, severity of liver disease and etiology and * adjusted by gender, age and etiology using logistic regression analysis. Patients with HCC but without fibrosis on nontumor liver were compared to healthy population from 3C study and 1000 genome. Significant p values and relative OR values are shown in bold type.

cirrhosis (F4, 514 cirrhotic patients with HCC vs. 1,555 cirrhotic patients without HCC), the at-risk GG genotype of *PNPLA3* and CT + TT genotype of *TM6SF2* were associated with HCC risk even after adjustment for age, gender and etiologies ($p = 0.0005$, OR = 2.13 [CI95%: 1.39–3.26] for *PNPLA3* rs738409 and $p = 0.04$, OR = 1.45 [CI95%: 1.01–2.08] for *TM6SF2* rs58542926, Table 4). In patients with significant fibrosis (F2-F3), neither *TM6SF2* rs58542926 nor *PNPLA3* rs738409 were associated with HCC risk after adjustment (Table 3). Then, we compared genotype distribution in HCC patients without fibrosis (F0-F1, $n = 261$), in the 503 healthy Europeans from 1000Genomes and in 1981 healthy

individuals from the French population (3City project). No significant association was found between *TM6SF2* rs58542926 and HCC development on nonfibrotic liver (Table 3). In contrast, patients with HCC developed on nonfibrotic liver (F0-F1) harbored a different distribution of *PNPLA3* genotype with a significant enrichment of at-risk genotypes (10% of GG and 41% of GC) compared to the 1000 genomes population (5% of GG, $p = 0.007$; 34% of GC, $p = 0.02$) and compared to the 3City population (7% of GG; $p = 0.02$; 35% of GC, $p = 0.02$) (Table 3).

We also analyzed *PNPLA3* rs738409 and *TM6SF2* rs58542926 among the different etiologies of underlying liver

Table 4. Multivariate analysis of genotype of PNPLA3 rs738409 and TM6SF2 rs58542926 in all patients with CLD and in patients with cirrhotic CLD in relation to risk of HCC

| Genotype | All patients | | | | | | | | | | | |
|------------------------------|--------------------|----------------------------|-----------------|----------|---------------------|------------|---------------|----------------------------|-------------|---------------|---------------------|---------|
| | Alcoholic etiology | | | NAFLD | | | HCV infection | | | HBV infection | | |
| | HCC/ CLD | n = 360/1222 | | HCC/ CLD | n = 121/231 | | HCC/ CLD | n = 256/480 | | HCC/ CLD | n = 90/87 | |
| | | OR (95%CI) | p Value | | OR (95%CI) | p Value | | OR (95%CI) | p Value | | OR (95%CI) | p Value |
| PNPLA3 rs738409 | | | | | | | | | | | | |
| CC | 108/530 | 1 | - | 43/85 | 1 | - | 144/257 | 1 | - | 47/53 | - | - |
| CG | 172/528 | 2.11 [1.49–2.98] | 2.31E-05 | 59/109 | 1.49 [0.75–2.98] | 0.25 | 88/192 | 0.79 [0.55–1.14] | 0.21 | 36/26 | 1.18 [0.59–2.37] | 0.64 |
| GG | 80/157 | 3.91 [2.52–6.06] | 1.14E-09 | 19/37 | 2.15 [0.82–5.61] | 0.12 | 22/31 | 1.69 [0.88–3.26] | 0.12 | 7/8 | 0.93 [0.26–3.27] | 0.90 |
| TM6SF2 rs58542926 | | | | | | | | | | | | |
| CC | 271/1026 | 1 | - | 91/184 | 1 | - | 213/425 | 1 | - | 79/79 | - | - |
| CT + TT | 89/193 | 1.79 [1.25–2.56] | 0.001 | 30/45 | 1.44 [0.67–3.07] | 0.35 | 43/55 | 1.78 [1.09–2.90] | 0.02 | 11/7 | 1.44 [0.47–4.37] | 0.52 |
| Patients with cirrhosis (F4) | | | | | | | | | | | | |
| Genotype | Alcoholic etiology | | | NAFLD | | | HCV infection | | | HBV infection | | |
| | HCC/ CLD | n = 294/1168 | | HCC/ CLD | n = 66/149 | | HCC/CLD | n = 242/361 | | HCC/ CLD | n = 78/69 | |
| OR (95%CI) | | p value | OR (95%CI) | | p value | OR (95%CI) | | p value | OR (95%CI) | | p value | |
| PNPLA3 rs738409 | | | | | | | | | | | | |
| CC | 52/487 | 1 | - | 5/33 | 1 | - | 91/137 | 1 | - | 30/30 | - | - |
| CG | 110/491 | 2.36 [1.61–3.46] | 1.2E-05 | 15/49 | 2.09 [0.58–7.51] | 0.26 | 66/103 | 0.91 [0.60–1.39] | 0.67 | 24/17 | 1.18 [0.51–2.72] | 0.70 |
| GG | 56/151 | 4.05 [2.54–6.45] | 4.1E-09 | 8/24 | 3.93 [0.93–16.6] | 0.06 | 19/18 | 2.01 [0.96–4.19] | 0.06 | 6/7 | 0.85 [0.24–3.05] | 0.80 |
| TM6SF2 rs58542926 | | | | | | | | | | | | |
| CC | 160/945 | 1 | - | 17/79 | 1 | - | 150/227 | 1 | - | 54/49 | - | - |
| CT + TT | 58/180 | 1.84 [1.26–2.68] | 0.0015 | 11/26 | 1.81 [0.64–5.14] | 0.27 | 28/31 | 1.47 [0.83–2.60] | 0.19 | 6/5 | 1.22 [0.33–4.53] | 0.76 |

Abbreviations: CLD, chronic liver disease; HCC, hepatocellular carcinoma; SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval. The associations are adjusted by age, gender and severity of liver disease by logistic approach. Significant p values and relative OR values are shown in bold type.

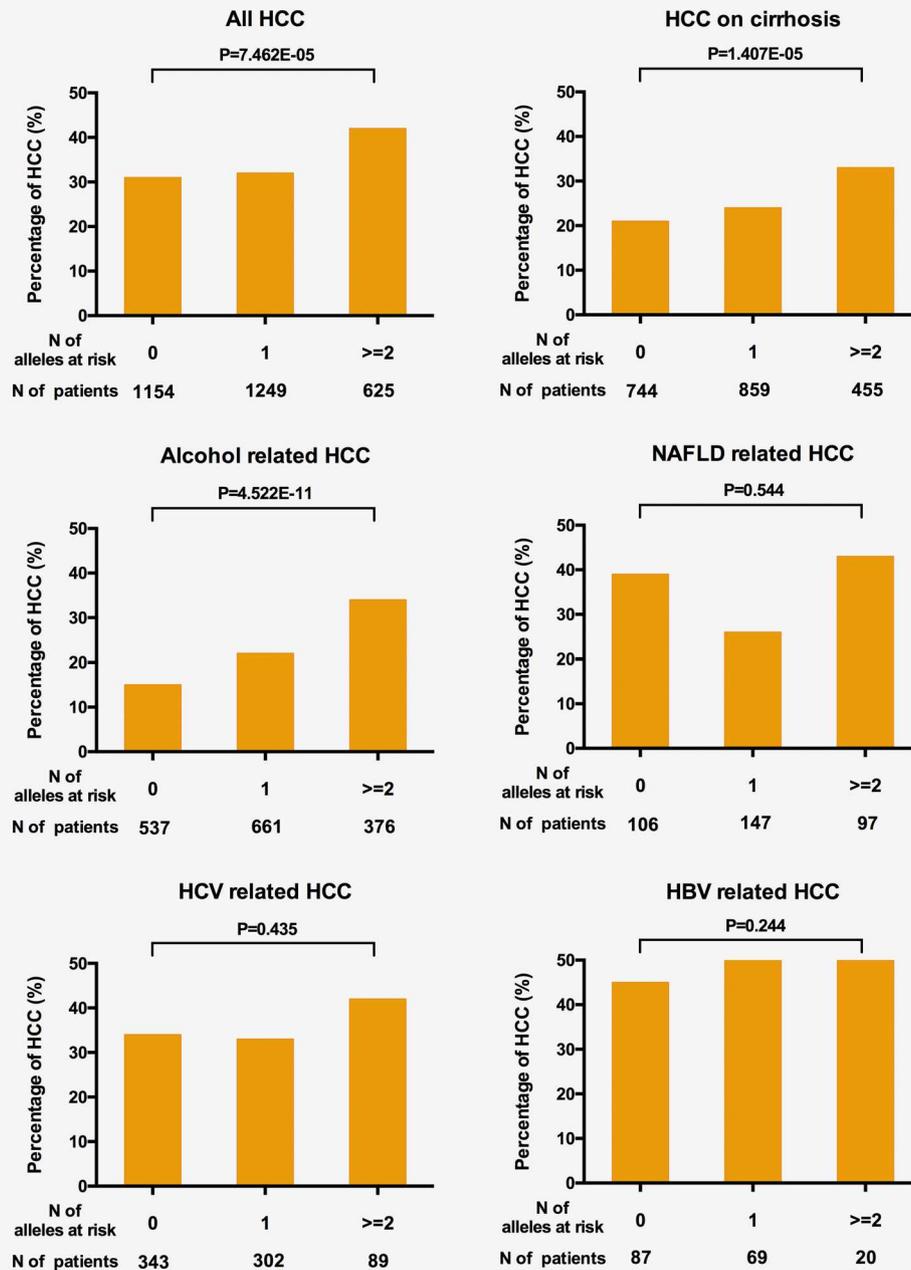


Figure 1. Risk of HCC according to the combination of at-risk alleles of *PNPLA3* rs738409 and *TM6SF2* rs58542926. We analyzed the percentages of patients with HCC according to the numbers of allele at-risk of *PNPLA3* rs738409 and *TM6SF2* rs58542926 (G for *PNPLA3* rs738409 and T for *TM6SF2* rs58542926). Analysis was performed in all HCC patients, thus in subgroups analysis according to the presence of cirrhosis or to the different etiologies. The x-axis represented three groups of patients with 0, 1 and more than 2 alleles at risk. The y-axis represented the percentage of patients with HCC in each group. Differences between distribution were assessed using the Chi-square test for trend.

disease (Table 4). In patients with high alcohol intake (360 HCC patients vs. 1,222 CLD patients without HCC), GG and GC genotype of *PNPLA3* was associated with HCC risk after adjustment for age, gender and degree of fibrosis (OR = 3.91 [CI95%: 2.52–6.06], $p = 1.14E-09$ and OR = 2.11 [CI95%: 1.49–2.98], $p = 2.31E-05$ respectively). Similarly,

TT + CT genotype of *TM6SF2* was also significantly associated with HCC occurrence in patients with alcoholic liver disease (OR = 1.79 [CI95%: 1.25–2.56], $p = 0.001$) after adjustment. No significant association was observed between *PNPLA3* rs738409 or *TM6SF2* rs58542926 and HCC risk in patients with NAFLD or chronic HBV infection whereas

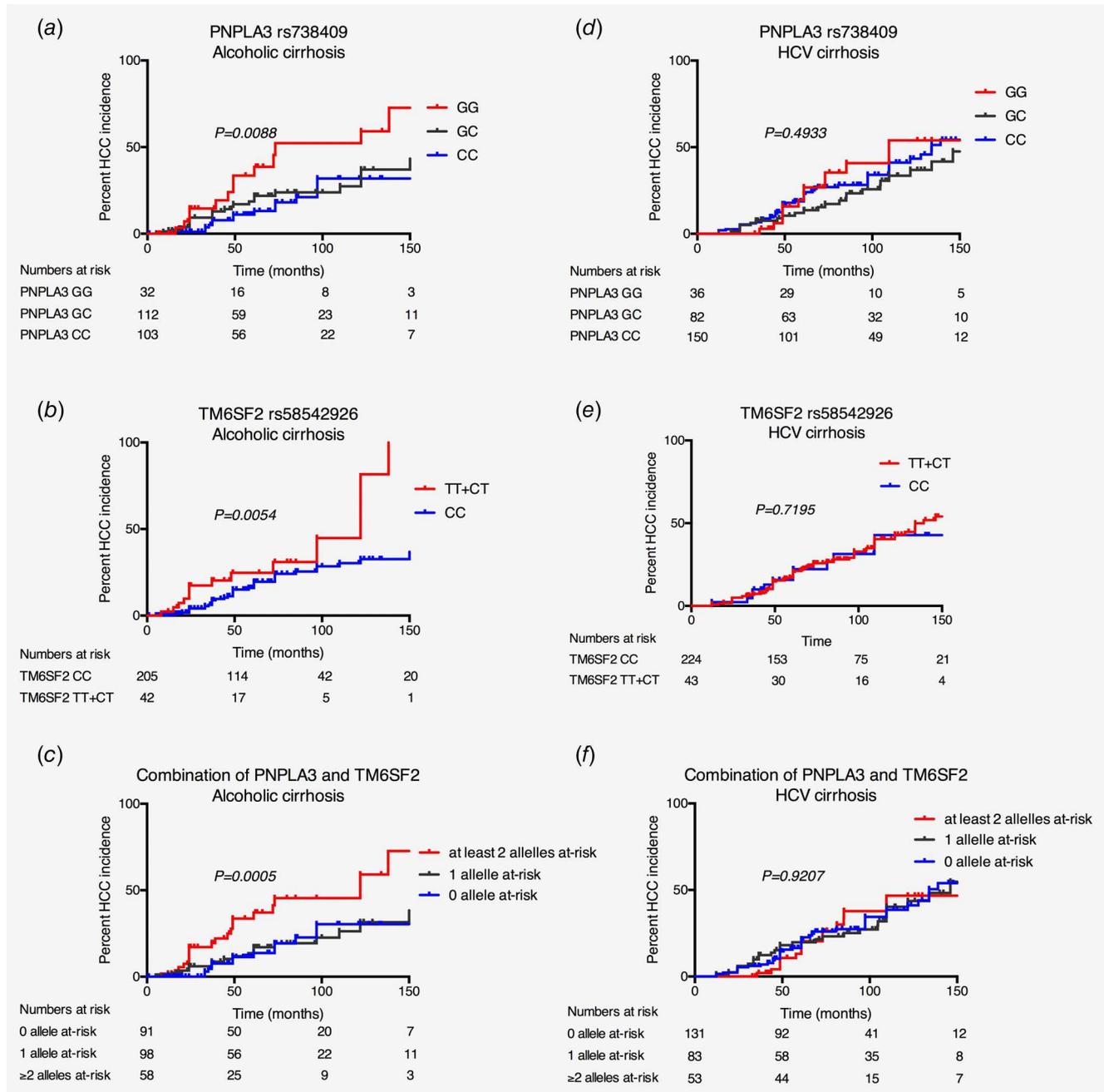


Figure 2. Incidence of HCC according to *PNPLA3* rs738409 and *TM6SF2* rs58542926 in prospective cohorts of HCV-related cirrhosis and alcohol-related cirrhosis. The incidence of HCC were represented using the Kaplan–Meier method with the log-rank test in prospective cohorts of alcohol-related cirrhosis and HCV-related cirrhosis: according to *PNPLA3* rs738409 in alcohol-related cirrhosis (a), according to *TM6SF2* rs58542926 in alcohol-related cirrhosis (b), according to the combination of at-risk alleles of *PNPLA3* rs738409 and *TM6SF2* rs58542926 in alcohol-related cirrhosis (c), according to *PNPLA3* rs738409 in HCV-related cirrhosis (d), according to *TM6SF2* rs58542926 in HCV-related cirrhosis (e), according to the combination of at-risk alleles of *PNPLA3* rs738409 and *TM6SF2* rs58542926 in HCV-related cirrhosis (f), The numbers at risk were represented under the x-axis.

TT + CT genotype of *TM6SF2* was associated with HCC development in HCV patients (OR = 1.78 [CI95%: 1.09–2.90], $p = 0.02$). In the subgroup of cirrhotic patients, we confirmed the association between *PNPLA3* rs738409 and *TM6SF2* rs58542926 and HCC risks in alcohol related cirrhosis but not in cirrhosis due to NAFLD, chronic HCV infection or chronic HBV infection (Table 4).

Finally, we combined the at-risk allele of *PNPLA3* rs738409 and *TM6SF2* rs58542926 and divided the population in carrier with 0, 1 and ≥ 2 alleles at-risk (Fig. 1). In patients with alcohol-related HCC, we observed a progressive increase of HCC risk with the number of at-risk allele (1 allele at-risk vs. no allele at-risk OR = 1.88 [CI95%: 1.29–2.75], $p = 0.001$ and ≥ 2 alleles at-risk vs. no allele at risk OR = 4.30 [CI95%: 2.85–6.48], $p = 3.2E-12$) (Fig. 1).

***PNPLA3* rs738409 and *TM6SF2* rs58542926 and HCC incidence in prospective cohorts of HCV- and alcohol-related cirrhosis**

We tested the longitudinal impact of *PNPLA3* rs738409 and *TM6SF2* rs58542926 on HCC incidence in one cohort of patients with alcoholic cirrhosis ($n = 249$) and another cohort of patients with HCV-related cirrhosis ($n = 268$), both prospectively screened every 6 months for HCC occurrence (see Supporting Information Table 3). The median age of patients with alcoholic cirrhosis was 57 years, with 79.5% of male and a mean Child Pugh score of 6.8. The median age of the patients with HCV-related cirrhosis was 56 years with 55% of male and a mean Child Pugh score of 5.3 (see Supporting Information Table 3 for the extensive description of the two cohorts at baseline). The incidence of HCC was 21% ($n = 53$) after a median follow-up of 59 months in patients with alcoholic cirrhosis and 32% ($n = 87$) after a median 85 months of follow-up in patients with HCV-related cirrhosis. As previously described,¹⁶ GG genotype of *PNPLA3* was associated with an increased HCC incidence compared to GC or CC genotype (log-rank test $p = 0.009$, Fig. 2). Patients with *TM6SF2* TT + CT genotypes had a higher incidence of HCC development compared to patients with CC genotype (log-rank test $p = 0.005$, Fig. 2). In multivariate analysis, combination of at-risk alleles of *PNPLA3* and *TM6SF2* SNPs remained an independent predictor of HCC development (HR = 4.38 [CI95% = 2.00–9.60], $p = 0.0002$), along with older age (HR = 1.09 [CI95%: 1.05–1.12], $p = 2.84E-06$), male gender (HR = 9.68 [CI95%: 1.32–70.84], $p = 0.03$) and a higher BMI (HR = 1.12 [CI95%: 1.06–1.19], $p = 0.0001$), (Supporting Information Table 4), in patients with alcohol-related cirrhosis. In contrast, neither *PNPLA3* rs738409 (log-rank test $p = 0.49$) nor *TM6SF2* rs58542926 (log-rank test $p = 0.72$) nor the combination of at-risk alleles (log-rank test $p = 0.92$) were not associated with HCC occurrence in patients with HCV-related cirrhosis (Fig. 2).

Association between SNP and pathological and molecular features in HCC patients

Finally, we search for links between *PNPLA3* and *TM6SF2* polymorphisms and tumor pathological features and somatic mutations in 9 driver genes in 563 patients with a unique HCC treated by liver resection. We did not identify any significant association between *PNPLA3* rs738409, *TM6SF2* rs58542926 and histological features of the tumor (Supporting Information Table 5). In contrast, patients with GG genotype of *PNPLA3* rs738409 were more likely to have a tumor with somatic *ARID1A* mutations compared to carriers of GC or CC genotype (26% vs. 12% or 11% respectively; $p = 0.05$). Patients with TT + CT genotype of *TM6SF2* rs58542926 had a higher rate of HCC with *TERT* promoter mutations (70% vs. 55%; $p = 0.01$) with less somatic mutations in *ARID2* (0% vs. 11% for *ARID2*, $p = 0.02$) compared to patients with CC genotype. We did not observe any significant association between the genotype of *PNPLA3* rs738409 and *TM6SF2* rs58542926 and *PNPLA3* and *TM6SF2* mRNA assessed by quantitative RT-PCR in 340 F0-F1 nontumor tissues (Supporting Information Fig. 2).

Discussion

Based on the analysis of a large number of 5,525 individuals, we refined our understanding of the role of *PNPLA3* and *TM6SF2* in genetic predisposition, with a strong role of these variants in HCC developed on alcoholic liver disease, but also with a role of *PNPLA3* in HCC development on nonfibrotic liver.

We also tested 7 SNP previously described as associated with HCC risk in GWAS. In the literature, SNP identified in GWAS failed to be validated in other GWAS and in case-control studies available in the literature.^{9,10,33–37} Similarly, we failed to confirm an association between these SNP and HCC development except for rs7574865 *STAT4* that were linked with HCC occurrence in HBV related HCC in our study, confirming the results of the original GWAS. In this analysis, we did not perform correction for multiple testing because we considered this part of our study as a replication of previous findings.^{9,10,33–37} The different results between our study and GWAS studies could be explained by the difference in term of control group (chronic liver disease population vs. healthy population) and the lack of adjustment on fibrosis stage in GWAS. This discrepancy could also be explained by the difference in ethnicity with Asian population included in GWAS studies and European ancestries in our study.

We observed, in both the case-control study and the prospective cohort, an association between *PNPLA3* rs738409 and *TM6SF2* rs58542926 and HCC development in patients with alcoholic liver disease. In line with a recently published case control study,²³ the present results confirmed the association between *PNPLA3* rs738409 and *TM6SF2* rs58542926 and HCC risk in alcoholic liver disease previously described in the literature.^{15,22} However, prospective well-defined cohorts are

of the gold standard to assess the association between genetic traits and the development of liver cancer in order to avoid bias link with case/control study. In this line, we confirmed for the first time in a prospective manner the link between *TM6SF2* rs58542926 and HCC occurrence in cohort of alcohol related cirrhosis.

In contrast, the potential association between *PNPLA3* rs738409 and HCC risk in patients with HCV-related cirrhosis remains controversial.^{38–41} The association between this variant and HCC risk in HCV-related cirrhosis was either marginal in our case–control study or negative in our prospective cohort. We observed the same results for *TM6SF2* rs58542926 that was not associated with HCC development in case–control or prospective cohort of HCV-related cirrhosis. Moreover, we also did not identify any significant association with HCC development in HBV-related liver disease or with NAFLD.¹⁷ We have to acknowledge that this absence of association could be potentially explained by a lack of power due to the limited numbers of patients with these etiologies analyzed in our study; Additional studies including a larger number of patients are required to fully investigate the role of *PNPLA3* rs738409 and *TM6SF2* rs58542926 in liver carcinogenesis due to NAFLD. In Asia, considering the increased incidence of HCC due to alcohol and metabolic syndrome, further analyses of *PNPLA3* and *TM6SF2* variants remain to be performed in various well-defined populations.

We also observed a higher rate of the at-risk G allele of *PNPLA3* rs738409 in patients with HCC developed on nonfibrotic liver. This association was firstly suggested in a case/control study including 25 HCC developed on noncirrhotic liver (F0–F2) in patients with HCV infection or alcoholic liver disease.⁴² Now, we confirmed this association in a larger series of 212 patients with HCC developed on nonfibrotic liver (F0–F1) and mostly without any etiologies of liver diseases and using two different control populations composed of a total of 2,484 healthy subjects.

This result showed a role of *PNPLA3* rs738409 in development of HCC on nonfibrotic liver suggesting a direct carcinogenic role of this SNP. We did not find any difference in *PNPLA3* mRNA level according to the genotype, in line with previous studies suggesting that this polymorphism modify the function of the protein but not the gene expression.^{43,44}

The rs738409 [G] variant in *PNPLA3* induces a loss of function of hydrolase activity of the protein with an increase of triglyceride content in hepatocyte, the shift to anaerobic metabolism and mitochondrial dysfunction as well modification of the microenvironment by hepatic stellate cells that could promote liver carcinogenesis.^{45–48} However, additional works are warranted to better understand the potential direct carcinogenic role of *PNPLA3*.

Dissection of HCC genetic predisposition will help designing chemo-preventive treatment based on a better comprehension of the biological pathway involved in tumor initiation. Although risk effect sizes for *PNPLA3* and *TM6SF2* reported in the current study and previous ones are of remarkable magnitude for common variants, it is unlikely that these SNPs, alone, will be valuable to daily clinical practice.⁴⁹ Indeed, as in other complex diseases, the variance explained by few common variants is too limited.⁵⁰ However, it can be anticipated that incorporating genetic data into clinical scoring systems, including these variants and others, as well as clinical factors, will allow the refinement of specific HCC risk subgroups.

In conclusion, *TM6SF2* rs58542926 and *PNPLA3* rs738409 are associated with HCC occurrence in alcoholic liver disease with an additive effect of these two polymorphisms in HCC risks. Moreover, the role of *PNPLA3* rs738409 in HCC development in nonfibrotic liver suggests a direct carcinogenic role of this polymorphism.

Acknowledgements

We thank the “Réseau National des Centres de Ressources Biologiques (CRB) Foie, the French Liver Biobanks network (INCa, BB-0033-00085, Hepatobio bank), CRB CHU Bordeaux (BB-0033-00036), the CRB Maladies du foie des Hôpitaux Universitaires Paris-Seine-Saint-Denis (BB-0033-00027), for the contributing to the sample collection. We are grateful to Nathalie Barget and Lucie Del Pozo for their valuable help in the preparation of samples. This work was supported by Institut National du Cancer (INCa) with the International Cancer Genome Consortium (ICGC LICA-FR project) and NoFLIC projects (PAIR HCC, INCa and ARC), INSERM with the « Cancer et Environnement » (plan Cancer), MUTHEC projects (INCa) and the HECAM project (BPI). The group is supported by the Ligue Nationale contre le Cancer (Equipe Labellisée), Labex OncoImmunology (investissement d’avenir), grant IREB, Coup d’Elan de la Fondation Bettencourt-Shueller, the SIRIC CARPEM and Fondation Mérieux.

References

- Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet* 2018;391:1301–14.
- Turati F, Edefonti V, Talamini R, et al. Family history of liver cancer and hepatocellular carcinoma. *Hepatology (Baltimore, Md.)* 2012;55:1416–25.
- Yu MW, Chang HC, Liaw YF, et al. Familial risk of hepatocellular carcinoma among chronic hepatitis B carriers and their relatives. *J Natl Cancer Inst* 2000;92:1159–64.
- Clifford RJ, Zhang J, Meerzaman DM, et al. Genetic variations at loci involved in the immune response are risk factors for hepatocellular carcinoma. *Hepatology (Baltimore, Md.)* 2010;52:2034–43.
- Miki D, Ochi H, Hayes CN, et al. Variation in the DEPDC5 locus is associated with progression to hepatocellular carcinoma in chronic hepatitis C virus carriers. *Nat Genet* 2011;43:797–800.
- Kumar V, Kato N, Urabe Y, et al. Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma. *Nat Genet* 2011;43:455–8.
- Li S, Qian J, Yang Y, et al. GWAS identifies novel susceptibility loci on 6p21.32 and 21q21.3 for hepatocellular carcinoma in chronic hepatitis B virus carriers. *PLoS Genet* 2012;8:e1002791.
- Jiang DK, Sun J, Cao G, et al. Genetic variants in STAT4 and HLA-DQ genes confer risk of hepatitis B virus-related hepatocellular carcinoma. *Nat Genet* 2013;45:72–5.
- Zhang H, Zhai Y, Hu Z, et al. Genome-wide association study identifies 1p36.22 as a new susceptibility locus for hepatocellular carcinoma in

- chronic hepatitis B virus carriers. *Nat Genet* 2010; 42:755–8.
10. Chan KY, Wong CM, Kwan JS, et al. Genome-wide association study of hepatocellular carcinoma in southern Chinese patients with chronic hepatitis B virus infection. *PLoS One* 2011;6:e28798.
 11. Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2008;40: 1461–5.
 12. Kozlitina J, Smagris E, Stender S, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2014;46:352–6.
 13. Liu YL, Reeves HL, Burt AD, et al. TM6SF2 rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. *Nat Commun* 2014;5:4309.
 14. Chen LZ, Xia HH, Xin YN, et al. TM6SF2 E167K variant, a novel genetic susceptibility variant, contributing to nonalcoholic fatty liver disease. *J Clin Transl Hepatol* 2015;3:265–70.
 15. Trepo E, Guyot E, Ganne-Carrie N, et al. PNPLA3 (rs738409 C>G) is a common risk variant associated with hepatocellular carcinoma in alcoholic cirrhosis. *Hepatology (Baltimore, Md.)* 2012;55:1307–8.
 16. Guyot E, Sutton A, Rufat P, et al. PNPLA3 rs738409, hepatocellular carcinoma occurrence and risk model prediction in patients with cirrhosis. *J Hepatol* 2013;58:312–8.
 17. Liu YL, Patman GL, Leathart JB, et al. Carriage of the PNPLA3 rs738409 C >G polymorphism confers an increased risk of non-alcoholic fatty liver disease associated hepatocellular carcinoma. *J Hepatol* 2014;61:75–81.
 18. Nakaoka K, Hashimoto S, Kawabe N, et al. PNPLA3 I148M associations with liver carcinogenesis in Japanese chronic hepatitis C patients. *SpringerPlus* 2015;4:83.
 19. Seko Y, Sumida Y, Tanaka S, et al. Development of hepatocellular carcinoma in Japanese patients with biopsy-proven non-alcoholic fatty liver disease: association between PNPLA3 genotype and hepatocarcinogenesis/fibrosis progression. *Hepatol Res* 2017;47:1083–92.
 20. Balasus D, Way M, Fusilli C, et al. The association of variants in PNPLA3 and GRP78 and the risk of developing hepatocellular carcinoma in an Italian population. *Oncotarget* 2016;7:86791–802.
 21. Buch S, Stickel F, Trepo E, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. *Nat Genet* 2015;47:1443–8.
 22. Falletti E, Cussigh A, Cmet S, et al. PNPLA3 rs738409 and TM6SF2 rs58542926 variants increase the risk of hepatocellular carcinoma in alcoholic cirrhosis. *Digest Liver Dis* 2016;48:69–75.
 23. Stickel F, Buch S, Nischalke HD, et al. Genetic variants in PNPLA3 and TM6SF2 predispose to the development of hepatocellular carcinoma in individuals with alcohol-related cirrhosis. *Am J Gastroenterol* 2018.
 24. Trevisani F, Frigerio M, Santi V, et al. Hepatocellular carcinoma in non-cirrhotic liver: a reappraisal. *Digest Liver Dis* 2010;42: 341–7.
 25. Genomes Project C, Abecasis GR, Altshuler D, et al. A map of human genome variation from population-scale sequencing. *Nature* 2010;467: 1061–73.
 26. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015;526:68–74.
 27. Alperovitch A, Amouyel P, Dartigues JF, et al. Epidemiological studies on aging in France: from the PAQUID study to the Three-City study. *C R Biol* 2002;325:665–72.
 28. Group CS. Vascular factors and risk of dementia: design of the Three-City study and baseline characteristics of the study population. *Neuroepidemiology* 2003;22:316–25.
 29. European Association For The Study Of The L, European Organisation For R, Treatment Of C. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2012;56:908–43.
 30. Bedossa P, Poinard T. An algorithm for the grading of activity in chronic hepatitis C. the METAVIR cooperative study group. *Hepatology (Baltimore, Md.)* 1996;24:289–93.
 31. Schulze K, Imbeaud S, Letouze E, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet* 2015;47:505–11.
 32. Boyault S, Rickman DS, de Reynies A, et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology (Baltimore, Md.)* 2007;45: 42–52.
 33. Hai H, Tamori A, Thuy LTT, et al. Polymorphisms in MICA, but not in DEPDC5, HCP5 or PNPLA3, are associated with chronic hepatitis C-related hepatocellular carcinoma. *Sci Rep* 2017;7:11912.
 34. Chen W, Wang M, Zhang Z, et al. Replication the association of 2q32.2-q32.3 and 14q32.11 with hepatocellular carcinoma. *Gene* 2015;561:63–7.
 35. Kim LH, Cheong HS, Namgoong S, et al. Replication of genome wide association studies on hepatocellular carcinoma susceptibility loci of STAT4 and HLA-DQ in a Korean population. *Infect Genetics Evol* 2015;33:72–6.
 36. Chen K, Shi W, Xin Z, et al. Replication of genome wide association studies on hepatocellular carcinoma susceptibility loci in a Chinese population. *PLoS One* 2013;8:e77315.
 37. Burza MA, Motta BM, Mancina RM, et al. DEPDC5 variants increase fibrosis progression in Europeans with chronic hepatitis C virus infection. *Hepatology (Baltimore, Md.)* 2016;63: 418–27.
 38. Yen YH, Tsai MC, Wu CK, et al. Association between PNPLA3 (rs738409 C>G) variant and hepatocellular carcinoma in Asian chronic hepatitis C patients: a longitudinal study. *J Formosan Med Assoc* 2017;117:833–40.
 39. Moritou Y, Ikeda F, Iwasaki Y, et al. Predictive impact of polymorphism of PNPLA3 on HCC development after interferon therapy in Japanese patients with chronic hepatitis C. *SpringerPlus* 2013;2:251.
 40. Khlaiphuengsin A, Kiatbumrung R, Payungporn S, et al. Association of PNPLA3 polymorphism with hepatocellular carcinoma development and prognosis in viral and non-viral chronic liver diseases. *Asian Pac J Cancer Prev* 2015;16: 8377–82.
 41. Trepo E, Nahon P, Bontempi G, et al. Association between the PNPLA3 (rs738409 C>G) variant and hepatocellular carcinoma: evidence from a meta-analysis of individual participant data. *Hepatology (Baltimore, Md.)* 2014;59:2170–7.
 42. Donati B, Dongiovanni P, Romeo S, et al. MBOAT7 rs641738 variant and hepatocellular carcinoma in non-cirrhotic individuals. *Sci Rep* 2017;7:4492.
 43. Pingitore P, Pirazzi C, Mancina RM, et al. Recombinant PNPLA3 protein shows triglyceride hydrolase activity and its I148M mutation results in loss of function. *Biochim Biophys Acta* 2014;1841:574–80.
 44. Perttila J, Huaman-Samanez C, Caron S, et al. PNPLA3 is regulated by glucose in human hepatocytes, and its I148M mutant slows down triglyceride hydrolysis. *Am J Physiol Endocrinol Metab* 2012;302:E1063–9.
 45. He S, McPhaul C, Li JZ, et al. A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. *J Biol Chem* 2010;285:6706–15.
 46. Pirazzi C, Adiels M, Burza MA, et al. Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro. *J Hepatol* 2012;57: 1276–82.
 47. Min HK, Sookoian S, Pirola CJ, et al. Metabolic profiling reveals that PNPLA3 induces widespread effects on metabolism beyond triacylglycerol remodeling in Huh-7 hepatoma cells. *Am J Physiol Gastrointest Liver Physiol* 2014;307: G66–76.
 48. Pirazzi C, Valenti L, Motta BM, et al. PNPLA3 has retinyl-palmitate lipase activity in human hepatic stellate cells. *Hum Mol Genet* 2014;23: 4077–85.
 49. Trepo E, Romeo S, Zucman-Rossi J, et al. PNPLA3 gene in liver diseases. *J Hepatol* 2016;65:399–412.
 50. Timpson NJ, Greenwood CMT, Soranzo N, et al. Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nat Rev Genet* 2018;19:110–24.

Supplementary tables and figures of Article 1

Supplementary table 1. List of SNPType assay from Fluidigm for genotyping.

| Gene | SNP | Assay ID | ASP1 allele | ASP2 allele | Allele specific primer 1 (ASP1) | Allele specific primer 2 (ASP2) | Locus specific primer (LSP) | Specific target amplification primer (STA) |
|------------------------|------------|----------------|-------------|-------------|----------------------------------|-----------------------------------|---|--|
| <i>PNPLA3</i> | rs738409 | GTA009201 3 | C | G | CCTTGGTATGTTTCCT GCTTCATG | CCTTGGTATGTTTCCT GCTTCATC | CCACGCCTCTGAAGGAAGG A | TCTCCTTTGCTTTCA CAGGC |
| <i>TM6SF2</i> | rs58542926 | GTA009655 5 | C | T | AAGAAGGCAGGCCT GATCTC | AAGAAGGCAGGCCT GATCTT | CCTGCACCATGGAAGGCAA | TGGCACCAGCAGGT AGG |
| <i>DEPDC5</i> | rs1012068 | GTA009655 1 | A | C | GGCTGCAATCAGGGG CTAT | GGCTGCAATCAGGGG GCTAG | CCAGGGAGGGCTGTAACAT TC | GCTCACAGGGAGG GTTT |
| <i>GRIK1</i> | rs455804 | GTA009690 1 | G | T | TGGAAAAATCTGTCA TAAACTAAGGCAC | GTGGAAAAATCTGTC ATAAACTAAGGCAA | TGTTTTGCTATTATTTTAAAGT TGAATGAATGTAGTTTAAA | ATTCACAGGGGCTT GCATG |
| <i>KIF1B</i> | rs17401966 | GTA005605 2 | A | G | CACATAGTGCCTCTA TGAGTCCA | CACATAGTGCCTCTA TGAGTCCG | GAAAATTTGAAAACCTCTA AGAACAACCTGACTCA | TCCAGCACTTAATG AAAACACATAGT |
| <i>STAT4</i> | rs7574865 | GTA003253 7 | G | T | AGTATGAAAAGTTGG TGACCAAATGTG | AGTATGAAAAGTTG GTGACCAAATGTT | CCCTGAAATTCCTGAAAT AAGATAACCAC | ATTACATGAGTGTG TATGCAGTAAAAGT |
| <i>MICA</i> | rs2596542 | GTA009655 8 | A | G | GTCTCCCAAAGAACA GCTACACA | TCTCCCAAAGAACA CTACACG | AATTGCTCTTACTCTTTTAT ATGACAACATAGAATTTTAT | CTCCACTACCACAG TGAAGATAAAG |
| <i>DLC1/LOC1019296</i> | rs2275959 | GTA009654 7 | C | T | CCCGAGCTGATTAC AGAAC | GCCCGAGCTGATTCA CAGAAT | GTAATTTATTGATAGCAAGG TTTGCATGCATTT | GGGATTCCCAGGAA GTCCA |
| <i>DDX18 region</i> | rs2551677 | GTA009650 2 | A | G | TGACACAGGATTTAA CTCGCT | TGACACAGGATTTAA CTCGCC | TCATCCCAGCAGCTTGAGAG T | GAGAATTCATCTAA ATGAGAGCCCT |

Supplementary table 2. Validation of association between SNPs from GWAS and risk of HCC in patients with chronic liver disease.

| Gene | (SNP) | Genotype | HCC/CLD (n=1018/2021) | OR (95%CI) | p-value | Adjusted OR (95%CI) | Adjusted p-value |
|----------------------------------|-------|----------|--------------------------|------------------|---------|------------------------|------------------|
| <i>DEPDC5</i> rs1012068 | | TT | 559/1049 | 1 | - | 1 | - |
| | | TG | 387/796 | 0.91 [0.78-1.07] | 0.27 | 0.95 [0.62-1.48] | 0.83 |
| | | GG | 74/171 | 0.81 [0.61-1.09] | 0.17 | 1.23 [0.80-1.88] | 0.34 |
| <i>KIF1B</i> rs17401966 | | AA | 499/1069 | 1 | - | 1 | - |
| | | AG | 429/791 | 1.16 [0.99-1.36] | 0.07 | 1.08 [0.85-1.36] | 0.55 |
| | | GG | 92/157 | 1.26 [0.95-1.66] | 0.11 | 1.09 [0.70-1.68] | 0.71 |
| <i>MICA</i> rs2596542 | | CC | 371/750 | 1 | - | 1 | - |
| | | CT | 466/912 | 1.03 [0.87-1.22] | 0.73 | 1.06 [0.83-1.37] | 0.63 |
| | | TT | 182/353 | 1.04 [0.84-1.30] | 0.74 | 0.94 [0.68-1.31] | 0.72 |
| <i>GRIK1</i> rs455804 | | CC | 672/1260 | 1 | - | 1 | - |
| | | CA | 302/655 | 0.86 [0.73-1.02] | 0.09 | 0.80 [0.48-1.34] | 0.40 |
| | | AA | 46/104 | 0.83 [0.58-1.19] | 0.33 | 0.86 [0.53-1.40] | 0.55 |
| <i>STAT4</i> rs7574865 | | GG | 633/1237 | 1 | - | 1 | - |
| | | GT | 342/664 | 1.13 [0.98-1.30] | 0.93 | 1.08 [0.84-1.37] | 0.56 |
| | | TT | 45/112 | 0.79 [0.55-1.12] | 0.22 | 1.05 [0.62-1.78] | 0.86 |
| <i>DLC1</i> rs2275959 | | GG | 272/546 | 1 | - | 1 | - |
| | | AG | 526/911 | 1.16 [0.97-1.39] | 0.12 | 1.10 [0.90-1.35] | 0.34 |
| | | AA | 222/382 | 1.17 [0.94-1.45] | 0.18 | 1.13 [0.88-1.45] | 0.34 |
| <i>DDX18 region</i> rs2551677 | | GG | 705/1241 | 1 | - | 1 | - |
| | | GA | 282/535 | 1.4 [1.17-1.68] | 0.41 | 1.00 [0.83-1.21] | 0.99 |
| | | AA | 33/63 | 0.92 [0.60-1.42] | 0.75 | 0.97 [0.60-1.57] | 0.90 |

CLD: chronic liver disease; HCC: hepatocellular carcinoma; SNP: single nucleotide polymorphism; OR: odds ratio; CI: confidence interval.

Analysis was performed using Fisher's exact test. Adjusted OR and P were adjusted by age, gender, etiology and severity of CLD. Allele at risk of the variants and significant p-values are in bold.

Supplementary Table 3: clinical features of prospective cohorts of patients with alcoholic related cirrhosis and HCV relative cirrhosis

| | Alcohol related cirrhosis (n=249) | HCV related cirrhosis (n=268) |
|--------------------------------|--|--------------------------------------|
| Age (years) | 57 +/- 10 | 56.3 +/- 0.74 |
| Gender (male) | 198 (79.5%) | 148 (55%) |
| Diabetes type 2 | 75/236 (32%) | 73/224 |
| BMI (kg/m2) | 29 +/- 5.9 | 28.9 +/- 0.41 |
| Platelets (103/mm3) | 137 +/- 4.4 | 135.7 +/- 3.6 |
| Albumin (g/l) | 36.6 +/- 0.4 | 41.2 +/- 0.33 |
| TP (%) | 64.4 +/- 1.29 | 83.1 +/- 0.9 |
| Total Bilirubine (ymol/l) | 23 +/- 4.4 | 16.24 +/- 0.65 |
| Child pugh score | 6.8 +/- 0.13 | 5.3 +/- 0.05 |
| GGT (x upper limit of normal) | 7.4 +/- 16.8 | 3.1 +/- 0.3 |
| AST (x upper limit of normal) | 2.18 +/- 0.09 | 2.8 +/- 0.37 |
| ALT (x upper limit of normal) | 1.43 +/- 0.05 | 3.6 +/- 0.66 |
| HCC incidence | 53 (21%) | 87 (32%) |
| Death or liver transplantation | 151 (60.6%) | 112 (42%) |

Mean with SEM and numbers (percentage)

BMI= Body mass index, GGT= gamma glutamyl transferase, AST= aspartate amino transferase, ALT= alanine amino transferase

Supplementary Table 4: univariate and multivariate analysis of variables associated with HCC occurrence in prospective cohort of alcohol related cirrhosis

| | Univariate analysis | | | Multivariate analysis | | |
|--------------------------------------|---------------------|------------|----------|-----------------------|------------|----------|
| | HR | CI95% | P value | HR | CI95% | P value |
| Age (years old) | 1.08 | 1.05-1.11 | 2.48E-07 | 1.09 | 1.05-1.12 | 2.84E-06 |
| Gender (male) | 8.00 | 0.001-0.19 | 0.004 | 9.68 | 1.32-70.84 | 0.03 |
| Diabetes type 2 | 2.02 | 1.18-3.48 | 0.01 | 1.21 | 0.65-2.28 | 0.55 |
| Body Mass Index (kg/m ²) | 1.10 | 1.94-32.92 | 6.28E-05 | 1.12 | 1.06-1.19 | 0.0001 |
| Platelets (/mm ³) | 1.00 | 0.99-1.00 | 0.45 | | | |
| Child Pugh score | 1.07 | 0.93-1.23 | 0.34 | | | |
| GGT (x upper limit of normal) | 0.96 | 0.91-1.01 | 0.10 | | | |
| AST (x upper limit of normal) | 0.95 | 0.77-1.18 | 0.65 | | | |
| ALT (x upper limit of normal) | 1.14 | 0.81-1.61 | 0.44 | | | |
| <i>TM6SF2 + PNPLA3</i> | | | | | | |
| 1 allele at risk | 1.09 | 0.53-2.25 | 0.81 | 2.31 | 0.97-5.50 | 0.06 |
| ≥ 2 alleles at risk | 3.12 | 1.58-6.17 | 0.001 | 4.38 | 2.00-9.6 | 0.0002 |

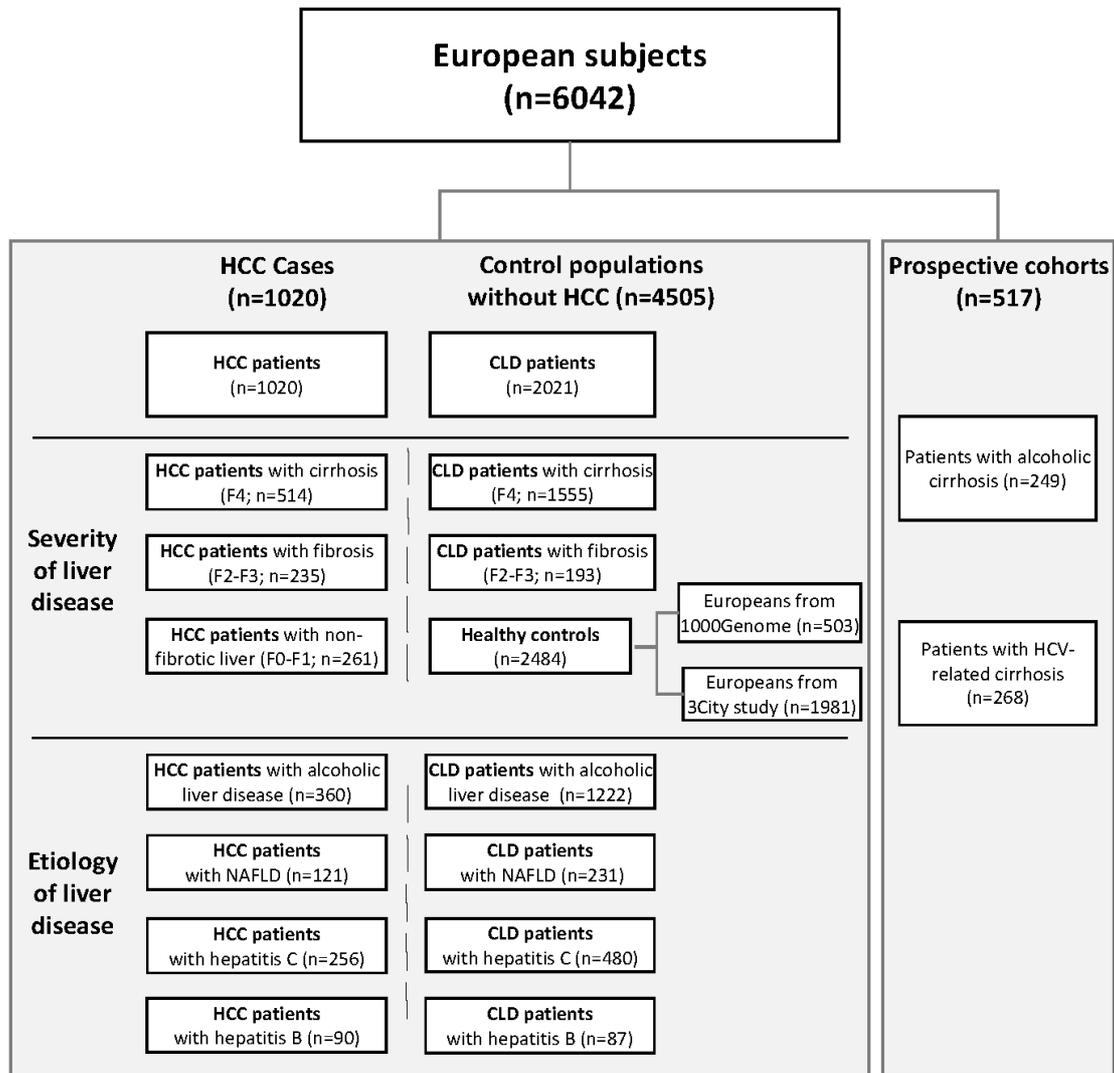
BMI: Body mass index, GGT: gamma glutamyl transferase, AST: aspartate amino transferase, ALT: alanine amino transferase

Supplementary Table 5. Association of *PNPLA3* rs738409 / *TM6SF2* rs58542926 with pathological and molecular features in HCC cases.

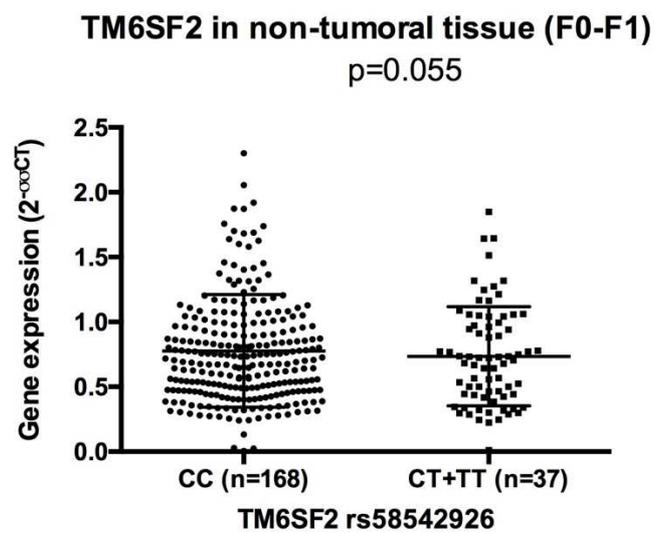
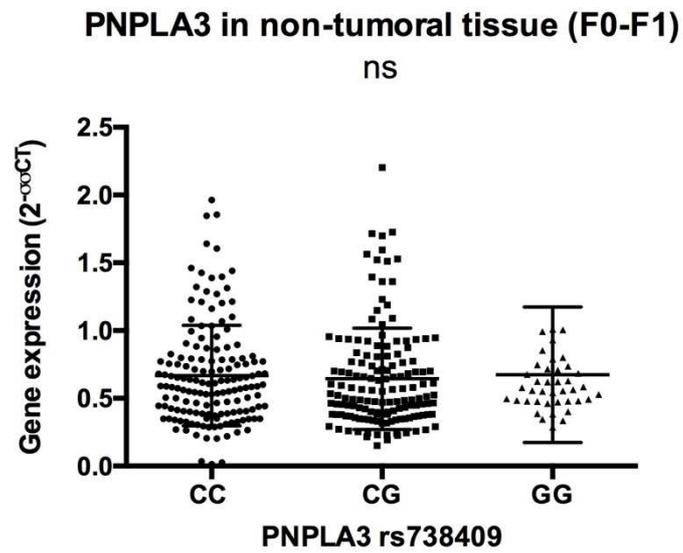
| Characteristics | Data available (n=563) | rs738409 | | | p-value | rs58542926 | | p-value |
|---------------------------------|------------------------|-----------|-----------|----------|-------------|------------|-----------|-------------|
| | | CC | CG | GG | | CC | CT+TT | |
| | | 246 (44%) | 241 (43%) | 75 (13%) | | 442 (79%) | 121 (21%) | |
| Tumor features | | | | | | | | |
| Largest nodule diameter (≤50mm) | 559 | 114 (47%) | 122 (51%) | 42 (56%) | 0.30 | 214 (49%) | 64 (53%) | 0.43 |
| Microscopic vascular invasion | 502 | 100 (49%) | 90 (46%) | 36 (56%) | 0.39 | 181 (50%) | 45 (46%) | 0.66 |
| Differentiation | | | | | | | | |
| good | 146 | 61 (27%) | 59 (26%) | 26 (38%) | | 116 (28%) | 30 (27%) | |
| medium | 308 | 140 (61%) | 132 (59%) | 35 (50%) | 0.34 | 239 (58%) | 69 (61%) | 0.88 |
| bad | 68 | 27 (12%) | 33 (15%) | 8 (12%) | | 54 (13%) | 14 (12%) | |
| 5 Gene Score (good prognosis) | 367 | 96 (59%) | 101 (65%) | 27 (57%) | 0.48 | 176 (60%) | 49 (58%) | 0.50 |
| G1G6 (G1-G3) | 393 | 55 (32%) | 56 (33%) | 16 (32%) | 0.78 | 106 (34%) | 22 (28%) | 0.34 |
| Major Somatic Mutation | | | | | | | | |
| <i>TERT</i> promoter | 403 | 93(53%) | 104 (61%) | 37 (69%) | 0.08 | 175 (55%) | 59 (70%) | 0.01 |
| <i>CTNNB1</i> | 345 | 55 (37%) | 52 (35%) | 15 (33%) | 0.90 | 100 (36%) | 22 (34%) | 0.97 |
| <i>TP53</i> | 346 | 33 (22%) | 20 (13%) | 8 (17%) | 0.14 | 47 (17%) | 14 (23%) | 0.34 |
| <i>AXIN1</i> | 322 | 9 (6%) | 6 (4%) | 4 (9%) | 0.45 | 15 (6%) | 4 (7%) | 1 |
| <i>ARID1A</i> | 294 | 14 (11%) | 15 (12%) | 10 (26%) | 0.05 | 36 (15%) | 3 (5%) | 0.09 |
| <i>ARID2</i> | 286 | 13 (10%) | 8 (7%) | 4 (10%) | 0.52 | 25 (11%) | 0 (0%) | 0.02 |
| <i>RPS6KA3</i> | 286 | 7 (6%) | 4 (3%) | 2 (5%) | 0.66 | 11 (5%) | 2 (4%) | 1 |
| <i>NFE2L2</i> | 293 | 7 (5%) | 3 (2%) | 1 (2%) | 0.42 | 9 (4%) | 2 (4%) | 1 |
| <i>KEAP1</i> | 276 | 3 (3%) | 5 (4%) | 1 (3%) | 0.74 | 7 (3%) | 2 (4%) | 1 |

Analysis was performed only in HCC patients with single tumor treated resection with comparison between distribution using Chi-square test.

Supplementary Figure 1



Supplementary Figure 2



Brief summarization of the current results:**Article 2: “A *HSD17B13* variant protects from hepatocellular carcinoma development in alcoholic liver disease”**

(Hepatology, 2019)

Jie Yang, Eric Trépo, Pierre Nahon, Qian Cao, Christophe Moreno, Eric Letouzé, Sandrine Imbeaud, Quentin Bayard, Thierry Gustot, Jacques Deviere, Paulette Bioulac-Sage, Julien Calderaro, Nathalie Ganne-Carrié, Alexis Laurent, Jean Frédéric Blanc, Erwan Guyot, Angela Sutton, Marianne Ziol, Jessica Zucman-Rossi*, Jean-Charles Nault*

Recently, a splicing variant of *HSD17B13* rs72613567, was identified as a new genetic variant associated with reduced risk of chronic liver disease, especially ALD, NAFLD and liver cirrhosis (Abul-Husn et al., 2018). Whether it could protect from the progression from CLD to HCC is currently unknown.

The goal of our study was to dissect the potential role of variants in *hydroxysteroid (HSD)* genes in liver carcinogenesis across varying etiologies and severities of underlying CLDs. By whole-exome sequencing, we investigated the association between 23 variants of *HSD* gene family and risk of HCC in an exploratory cohort (n=285). Frequency of TA allele of *HS17B13* rs72613567 was significantly lower in HCC patients than that in healthy populations. Then we included a cohort of 824 patients with HCC and 2206 patients without HCC but underlying CLD to validate the protective effect of *HSD17B13* variant. First, the protective TA allele of *HSD17B13* rs72613567 was observed less frequently in patients with ALD, NAFLD and CHC when compared with healthy populations. Then in patients with ALD, the TA allele of *HSD17B13* was significantly associated with a reduced risk of HCC, even after adjustment for age, sex, and fibrosis. In addition, it seems that TA allele could mitigate the risk of ALD-related HCC in carriers of at-risk G allele of *PNPLA3* rs738409.

Our study provided evidences that the *HSD17B13* rs72613567 loss-of-function variant protects from HCC development in patients with ALD.

A 17-Beta-Hydroxysteroid Dehydrogenase 13 Variant Protects From Hepatocellular Carcinoma Development in Alcoholic Liver Disease

Jie Yang,¹ Eric Trépo,^{1,2} Pierre Nahon,^{1,3,4} Qian Cao,⁵ Christophe Moreno,² Eric Letouzé,¹ Sandrine Imbeaud,¹ Quentin Bayard,¹ Thierry Gustot,² Jacques Deviere,² Paulette Bioulac-Sage,^{6,7} Julien Calderaro ,⁸ Nathalie Ganne-Carrié,^{1,3,4} Alexis Laurent,⁹ Jean Frédéric Blanc,¹⁰ Erwan Guyot,^{11,12} Angela Sutton,^{11,12} Marianne Ziol,^{1,4,13} Jessica Zucman-Rossi ,^{1,14*} and Jean-Charles Nault^{1,3,4*}

Recently, a loss of function variant (rs72613567) in 17-beta-hydroxysteroid dehydrogenase 13 (*HSD17B13*) has been identified as protective of nonalcoholic (NAFLD) and alcoholic liver disease (ALD). However, the role of this single-nucleotide polymorphism (SNP) in the development of hepatocellular carcinoma (HCC) is currently unknown. A total of 3,315 European patients with HCC (n=1,109) or without HCC, but with chronic liver disease (CLD; n=2,206), from four centers were analyzed either by whole-exome sequencing (WES; exploratory cohort, 285 HCC) or genotyped for *HSD17B13* rs72613567 (validation cohort, 824 HCC and all CLD). We included a control group of 33,337 healthy European individuals from the Exome Aggregation Consortium. We compared distribution of genotype using the chi-square test and logistic regression. In the exploratory cohort analyzed by WES, frequency of the TA allele of *HSD17B13* rs72613567 was significantly decreased in HCC patients compared to healthy controls ($P=1.52 \times 10^{-06}$). In the validation cohort, frequency of TA allele carriers was also decreased in patients with CLD and without HCC (39%) compared to healthy individuals (47%; $P < 0.0001$). The protective effect of the TA allele of *HSD17B13* rs72613567 was identified in patients with ALD (odds ratio [OR]=0.73; 95% confidence interval [CI], 0.65-0.82; $P < 0.0001$), NAFLD (OR=0.64; 95% CI, 0.49-0.83; $P=0.0007$), and hepatitis C (OR=0.71; 95% CI, 0.60-0.85; $P=0.0002$). In patients with ALD, the proportion of TA allele carriers with HCC was significantly lower (32%) than in CLD patients without HCC (40%), even after adjustment for age, sex, and fibrosis (OR=0.64; 95% CI, 0.46-0.87; $P=0.005$). **Conclusion:** The *HSD17B13* rs72613567 loss of function variant is protective of HCC development in patients with ALD. (HEPATOLOGY 2019;0:1-10).

The complex pathogenesis of hepatocellular carcinoma (HCC) is related to the high diversity of etiologies and severity of underlying liver diseases (LDs) leading to HCC development. Consequently, understanding the mechanisms of liver carcinogenesis requires the analysis of the full spectrum

Abbreviations: ALD, alcoholic liver disease; BMI, body mass index; CHB, chronic hepatitis B; CHC, chronic hepatitis; CI, confidence interval; CLD, chronic liver disease; ExAC, Exome Aggregation Consortium; GWAS, genome-wide association study; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; HSD17B13, 17-beta-hydroxysteroid dehydrogenase 13; NAFLD, nonalcoholic fatty liver disease; OR, odds ratio; PNPLA3, patatin-like phospholipase domain-containing 3; SNP, single-nucleotide polymorphism; TM6SF2, transmembrane 6 superfamily 2; WES, whole-exome sequencing.

Received December 25, 2018; accepted March 5, 2019.

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.30623/supinfo.

Supported by Institut National du Cancer (INCa) with the International Cancer Genome Consortium (ICGC LICA-FR project) and NoFLIC projects (PAIR HCC, INCa and ARC), INSERM with the "Cancer et Environnement" (plan Cancer), MUTHEC projects (INCa), and the HECAM project (BPI). The group is supported by the Ligue Nationale contre le Cancer (Equipe Labellisée), Labex OncoImmunology (investissement d'avenir), grant IREB, Coup d'Élan de la Fondation Bettencourt-Shueller, the SIRIC CARPEM, and Fondation Mérieux.

*These authors shared co-authorships.

© 2019 by the American Association for the Study of Liver Diseases.

View this article online at wileyonlinelibrary.com.

of the disease among the different etiologies and varying severities of LDs from normal liver to cirrhosis. Single-nucleotide polymorphisms (SNPs) have been associated with risk of developing cancer, and the genome-wide association study (GWAS) is a powerful tool to identify in an agnostic way predisposition factor to human diseases.⁽¹⁾ Several SNPs have been robustly associated with risk to develop HCC.⁽²⁻⁵⁾ A GWAS has identified a SNP of patatin-like phospholipase domain-containing 3 (*PNPLA3*) gene rs738409, which is associated with risk of development of nonalcoholic fatty liver disease (NAFLD).⁽⁶⁾ Additional studies have also reported an association with risk of developing steatosis, fibrosis, and cirrhosis attributed to high alcohol intake and also with risk of HCC on alcohol-related liver disease (ALD).⁽⁷⁻¹²⁾ Furthermore, transmembrane 6 superfamily 2 (*TM6SF2*) rs58542926 has also been associated with risk of NAFLD and ALD as well as HCC.^(10,11,13,14) The impact of *PNPLA3* and *TM6SF2*

in HCC development on chronic hepatitis C (CHC) or B (CHB) seems to be marginal.^(11,15,16) Interestingly, these genes are involved in lipid metabolism. For example, *PNPLA3* has a lipase activity, and the rs738409 *PNPLA3* I148M induces a loss of function of the protein and promotes triglyceride accumulation in hepatocytes.⁽¹⁷⁻¹⁹⁾ Recently, a variant (rs72613567) of 17-beta-hydroxysteroid dehydrogenase 13 (*HSD17B13*) has been identified as protective of NAFLD and ALD in a GWAS.⁽²⁰⁾ *HSD17B13* rs72613567 is a loss-of-function variant with an indel (insertion of A) leading to production of a truncated protein with a reduced enzymatic activity.^(20,21) Like *PNPLA3*, *HSD17B13* also locates to lipid droplets, and its expression was found to be 5.9-fold higher in NAFLD patients compared to normal liver samples.^(22,23) However, the role of this SNP in development of HCC is currently unknown. The aim of our study was to dissect the potential role of variants in hydroxysteroid genes in liver carcinogenesis

DOI 10.1002/hep.30623

Potential conflict of interest: J.F.B. reports personal fees from Bayer SP, Eli Lilly Oncology, Novartis, and BMS outside the submitted work. O. Seror received personal fees and nonfinancial support from Angiodynamics, Olympus, and Bayer Schering Pharma and received personal fees from GE as a consultant outside the submitted work. N. Ganne and P. Nabon received personal fees from Bayer Schering Pharma outside the submitted work.

ARTICLE INFORMATION:

From the ¹Cordeliers Research Center, Sorbonne University, Inserm, USPC, Paris Descartes University, Paris Diderot University, Paris 13 University, Functional Genomics of Solid Tumors Laboratory, Paris, France; ²Department of Gastroenterology, Hepatopancreatology and Digestive Oncology, University clinics of Bruxelles Erasme Hospital, Free University of Bruxelles, Bruxelles, Belgium; ³Liver unit, Jean Verdier Hospital, Paris-Seine-Saint-Denis University Hospital, Assistance-Publique Paris Hospitals, Bondy, France; ⁴Paris 13 University, Sorbonne Paris Cité, Paris, France; ⁵Department of Infectious Diseases, Zhongnan Hospital of Wuhan University, Wuhan, China; ⁶Bordeaux university, UMR1053 Bordeaux Research in Translational Oncology, BaRITOn, Bordeaux, France; ⁷Department of Pathology, Pellegrin Hospital, Bordeaux CHU, Bordeaux, France; ⁸Department of Pathology, Henri Mondor Hospital; Paris Est Créteil University, Inserm U955, Team 18, Mondor Institute for Biomedical Research, Créteil, France; ⁹Department of Hepatobiliary Surgery, Henri Mondor Hospital, Créteil; Paris Est Créteil University, Inserm U955, Team 18, Mondor Institute for Biomedical Research, Créteil, France; ¹⁰Department of Gastroenterology, Magellan Medical and Surgical Center, Haut-Lévêque hospital, CHU of Bordeaux, Bordeaux, France; ¹¹Department of Biochemistry, Jean Verdier Hospital, Paris-Seine-Saint-Denis University Hospital, Assistance-Publique Paris Hospitals, Bondy, France; ¹²INSERM U1148 LVTS, UFR SMBH, Paris 13 University, PRES Paris Sorbonne Cité, Bobigny, France; ¹³Biological Resource Center (BB-0033-00027), Paris-Seine-Saint-Denis University Hospital, Assistance-Publique Paris Hospitals, Bondy, France; ¹⁴European Georges Pompidou Hospital, Assistance-Publique Paris Hospitals, Paris, France.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO:

Jean-Charles Nault, M.D., Ph.D.
Service d'hépatologie, hôpital Jean Verdier
Avenue du 14 juillet
93140 Bondy, France
E-mail: naultjc@gmail.com
Tel.: +33 6 10 67 94 61
or

Jessica Zucman-Rossi, M.D., Ph.D.
INSERM U 1162, Génomique fonctionnelle des tumeurs solides
27 Rue Juliette Dodu
75010 Paris, France
E-mail: jessica.zucman-rossi@inserm.fr
Tel.: +33 1 53 72 51 66

across varying etiologies and severities of underlying LDs.

Patients and Methods

STUDY PARTICIPANTS

We enrolled a total of 3,315 patients with HCC (n = 1,109) or with chronic liver diseases (CLDs), but without HCC (n = 2,206), in four European centers (three in France and one in Belgium). Severity of CLD was assessed histologically and classified according to the METAVIR scoring system⁽²⁴⁾ (F0-F1 as nonfibrotic liver, F2-F3 as fibrotic liver, and F4 as cirrhotic liver). CHC was defined by positive anti-HCV (hepatitis C virus) serology with a positive viral load in PCR at diagnosis or at any time of the history of the patient. CHB was defined on the basis of hepatitis B surface antigen seropositivity. Chronic alcohol intake was defined as the main cause of LD in patients with alcohol consumption above 80 g per day in males and above 60 g per day in females for at least 10 years, but without chronic HCV or hepatitis B virus (HBV) infection. NAFLD was defined as the main cause of LD in patient with metabolic syndrome, but without high alcohol intake or viral infection. Diagnosis of HCC was based on histological analysis or non-invasive criteria in patients with cirrhosis, according to European Association for the Study of the Liver (EASL)/European Organization for Research and Treatment of Cancer (EORTC) Clinical Practice Guidelines.⁽²⁵⁾ We also used publically available data of 33,337 Healthy European individuals from the Exome Aggregation Consortium (ExAC) as a control group for patients with CLD (<http://exac.broadinstitute.org/>). Distribution of *HSD17B13* rs72613567 was the same in the ExAC population compared to the healthy European population of the 1000 Genomes Project.^(26,27) All individuals analyzed in this study were of European descent. The study was approved by the Ethics Committee (IRB HECAM IRB00003888), and all individuals provided written informed consent.

WHOLE-EXOME SEQUENCING

Description of analysis of whole-exome sequencing (WES) of nontumor liver of the exploratory cohort (n = 285 HCC patients) is detailed in the Supporting Information.

GENOTYPING

Among the 1,109 patients with HCC, 285 patients had genotypes determined using WES and 824 using custom SNP genotyping assays, whereas all the 2,206 CLD patients had the genotype determined using SNP genotyping assays. Genotyping for *HSD17B13* rs72613567 was undertaken in all patients by TaqMan custom SNP genotyping assays, according to the manufacturer's protocol. In brief, the forward primer sequence was GCTCTATTG GTGTTTTAGTATTTGGGTGTT; the reverse primer sequence was TGTTCCATCGTATATC AATATCTTTCTGAGACT; the reporter 1 (-VIC) sequence was CTGTGCTGTACTTACTTCT; and the reporter 2 (-FAM) sequence was TGCTGTAC TTAACTTCT. Genotyping success rate was >98%. Genotype frequencies for *HSD17B13* rs72613567 were in Hardy-Weinberg equilibrium (HWE) in healthy individuals and patients with CLD (online HWE calculator; <http://www.oege.org/software/hwe-mr-calc.shtml>).

STATISTICAL ANALYSES

To describe the study population, continuous variables are presented as mean with SD and categorical variables were reported as numbers and percentages. Variant allele frequency between patients included in the WES HCC cohort (n = 285) and European individuals from the ExAC database (n = 33,337) was compared using Fisher's exact test, and the odds ratio (OR) was derived for the presence of the alternate allele. To account for multiple testing, we used a Bonferroni correction. For the comparison of genotype distribution between CLD patients versus healthy control (ExAC) and between HCC patients versus CLD patients, we used the chi-square test in univariate analysis. Associations between *HSD17B13* rs72613567 variant and tumor features were analyzed by chi-square test. Statistical analyses were carried out using R statistical analysis software (version 3.2.0; <http://www.R-project.org/>). *P* values <0.05 were considered statistically significant.

Results

A total of 3,315 European patients including 1,109 with HCC (cases) and 2,206 with CLDs (controls)

were included in the study (Table 1; Fig. 1). We also included 33,337 healthy European individuals from the ExAC population.

First, we analyzed an exploratory cohort of 285 patients with WES of nontumor liver available (81% male; mean average, 64 years old; 35% with ALD, 21% with HCV, 11% with HBV, and 45% with cirrhosis; Table 1). We compared distribution of damaging germline variants of 20 hydroxysteroid genes (*HSD3B2*, *HSD3B1*, *HSD17B7*, *HSD11B1*, *HSD17B13*, *HSD17B11*, *HSD17B4*, *HSD17B8*, *HSD17B3*, *HSDL2*, *HSD17B7P2*, *HSD17B12*, *HSD17B6*, *HSD3B7*, *HSD11B2*, *HSD17B2*, *HSDL1*, *HSD17B1*, *HSD11B1L*, and *HSD17B14*) between the 285 HCC patients and 33,337 healthy patients of the ExAC population (Supporting Table S1). Among the 23 damaging variants of the hydrosteroid gene family identified, only the frequency of *HSD17B13* rs72613567 SNP was significantly different between HCC patients and healthy controls after Bonferroni correction (Bonferroni-corrected statistical significance threshold of $P = 2.17 \times 10^{-03}$), with a lower frequency of the TA allele in HCC patients compared to healthy individuals (OR = 0.40; 95% confidence interval [CI], 0.25-0.60; $P = 1.52 \times 10^{-06}$; Fig. 2; Supporting Table S1).

In a second step, we aimed to assess the role of *HSD17B13* rs72613567 in CLD and in liver

carcinogenesis using a validation cohort including 824 patients with HCC and 2,206 control patients with CLD, but without HCC, and genotyped for *HSD17B13* rs72613567. Patients with CLD and without HCC had a mean age of 55 years and 69% were males. Etiologies of CLD were mainly ALD (60%), CHC (24%) and NAFLD (11%). Fibrosis stages of nontumor livers were F0-F1 for 11% of patients, F2-F3 for 10% of patients, and F4 for 79% of patients according to METAVIR classification (Table 1). Patients with HCC had a mean age of 63 years and 83% were males (see Supporting Table S2 for HCC features). ALD was observed in 34%, CHC in 25%, NAFLD in 17%, and CHB in 10% of patients. HCC was developed in a nontumor liver with fibrosis classified F0-F1 in 24%, F2-F3 in 25%, and F4 in 51% of cases (Table 1).

First, we analyzed distribution of the *HSD17B13* rs72613567 genotype in 2,206 CLD patients without HCC compared to the 33,337 healthy European individuals from the ExAC population. Combination of T/TA and TA/TA genotypes (defining the TA allele carriers) was less frequent in patients with CLD (39%) compared to ExAC individuals (47%; $P < 0.0001$; Table 2; Fig. 3A). Among the different etiologies, the protective effect of the TA allele of *HSD17B13* rs72613567 was identified in patients with ALD (OR = 0.73; 95% CI, 0.65-0.82; $P < 0.0001$), NAFLD

TABLE 1. Description of the Population

| Variables | | Patients With CLD Without HCC (n = 2,206) | HCC Patients Combined Cohort (n = 1,109) | |
|-------------------------------|-----------------------------|--|---|--|
| | | | HCC patients Exploratory Cohort (n = 285) | HCC Patients Validation Cohort (n = 824) |
| Age (years)* | Years old | 55 ± 11 | 64.8 ± 11.1 | 63 ± 11 |
| Sex† | Male | 1,513 (69%) | 232 (81%) | 684 (83%) |
| Etiology†† | Alcohol | 1,332 (60%) | 101 (35%) | 273 (34%) |
| | HBV | 94 (4%) | 30 (11%) | 83 (10%) |
| | HCV | 538 (24%) | 59 (21%) | 202 (25%) |
| | NAFLD | 241 (11%) | 51 (18%) | 137 (17%) |
| | Without or other etiologies | 1 (0.05%) | 42 (15%) | 117 (14%) |
| Nontumor liver†† | F0-F1 | 245 (11%) | 112 (39%) | 194 (24%) |
| Fibrosis according to METAVIR | F2-F3 | 221 (10%) | 44 (16%) | 207 (25%) |
| | F4 | 1,724 (79%) | 129 (45%) | 418 (51%) |

*Mean ± SD.

†Number (percentage).

‡Data unavailable: etiology of 2 HCC patients in exploratory cohort and 12 HCC patients in validation cohort; fibrosis stage of 16 CLD patients and 5 HCC patients in validation cohort.

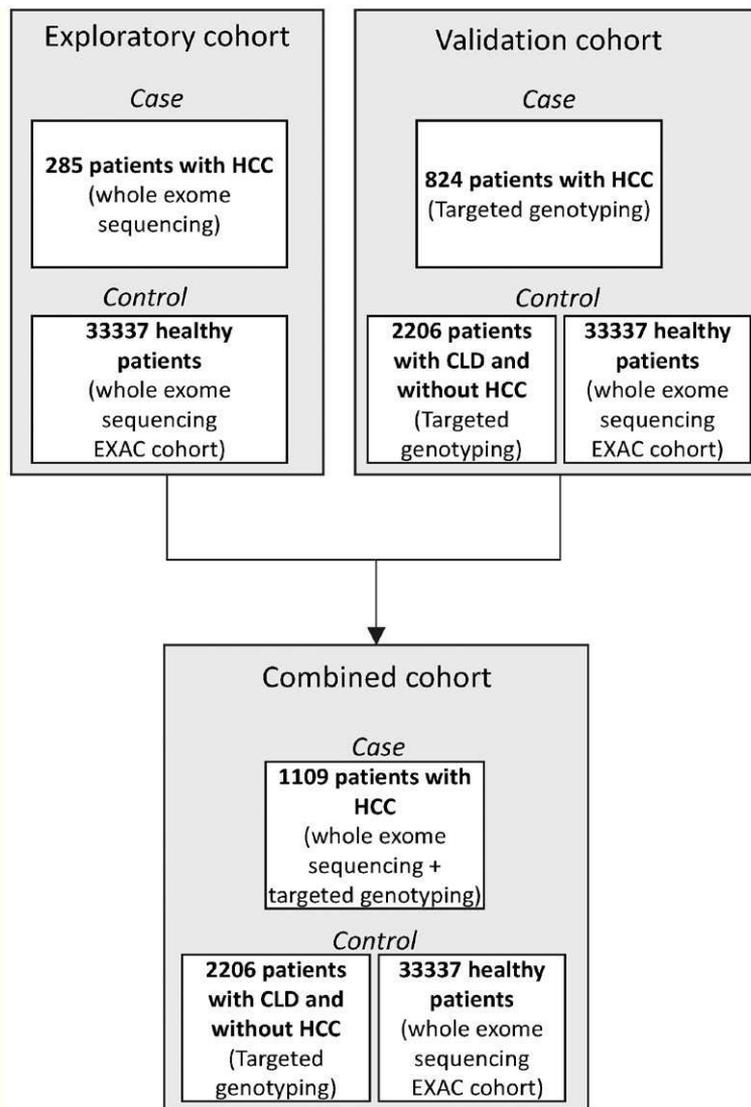


FIG. 1. Flow chart of the study.

(OR = 0.64; 95% CI, 0.49-0.83; $P = 0.0007$), and HCV (OR = 0.71; 95% CI, 0.60-0.85; $P = 0.0002$; Table 2; Fig. 3B).

Interestingly, across fibrotic stages, frequency of TA allele carriers of *HSD17B13* rs72613567 was significantly lower in patients with cirrhosis (38%) than in healthy controls (47%; $P < 0.0001$; Table 2; Fig. 3A) with a protective effect against cirrhosis mainly in patients with ALD, CHC infection, and NAFLD (Fig. 3C).

In addition, we compared the genotype distribution of *HSD17B13* rs72613567 in the validation cohort between patients with HCC ($n = 824$) and patients

with CLD, but without HCC ($n = 2,206$; Table 3). *HSD17B13* rs72613567 TA allele carriers were not significantly different between HCC cases and CLD controls in the validation cohort (Table 3). In contrast, frequency of *HSD17B13* rs72613567 TA allele carriers was significantly decreased in ALD patients with HCC (32%) compared to ALD patients without HCC (40%; $P = 0.02$; Table 3). The protective effect of TA allele remained significant after adjustment for age, sex, and fibrosis stage (OR = 0.64; 95% CI, 0.47-0.87; $P = 0.005$). No significant association between the *HSD17B13* SNP and HCC risk was identified in other etiologies (i.e., HCV, HBV, and NAFLD;

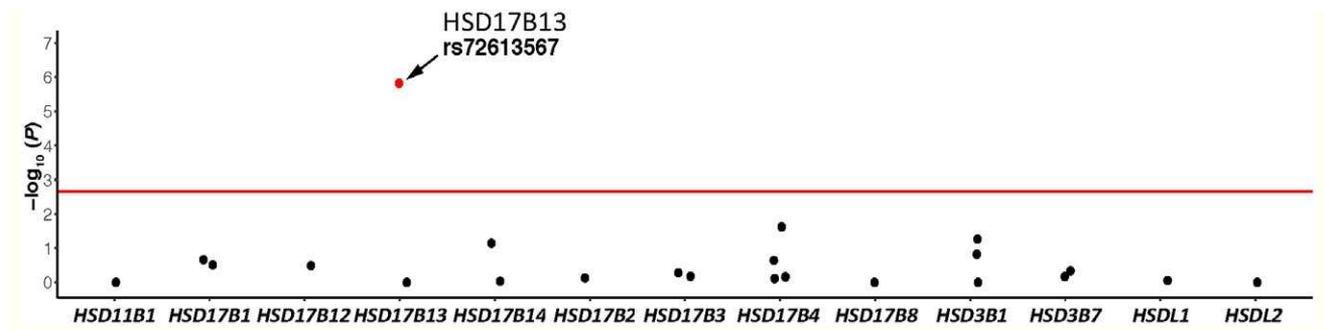


FIG. 2. Association of WES candidate damaging variants in hydroxysteroid genes with HCC in 285 HCC cases and 33,337 Europeans from the ExAC. *P* values are shown for candidate damaging variants that passed quality control in 20 hydroxysteroid genes. The Bonferroni-corrected statistical significance threshold ($P = 2.17 \times 10^{-3}$) is shown as a red line.

TABLE 2. Distribution of the *HSD17B13* rs72613567 Genotypes in Healthy Individuals and CLD Patients* Without HCC From the Validation Cohort

| Group | T/T | T/TA + TA/TA | T/TA + TA/TA vs. T/T | |
|----------------------------------|--------------|--------------|----------------------|------------------|
| | | | <i>P</i> Value | OR (95%CI) |
| Healthy population (ExAC) | 17,648 (53%) | 15,689 (47%) | — | — |
| All CLD patients | 1,342 (61%) | 864 (39%) | <0.0001 | 0.72 [0.66-0.79] |
| Alcohol-related CLD | 806 (60%) | 526 (40%) | <0.0001 | 0.73 [0.65-0.82] |
| NAFLD-related CLD | 154 (64%) | 87 (36%) | 0.0007 | 0.64 [0.49-0.83] |
| HCV-related CLD | 329 (61%) | 209 (39%) | 0.0002 | 0.71 [0.60-0.85] |
| HBV-related CLD | 52 (55%) | 42 (45%) | 0.64 | 0.91 [0.60-1.37] |
| CLD with cirrhosis (F4) | 1,063 (62%) | 661 (38%) | <0.0001 | 0.70 [0.63-0.77] |
| CLD with fibrosis (F2-F3) | 135 (61%) | 86 (39%) | 0.02 | 0.55 [0.55-0.94] |
| CLD on nonfibrotic liver (F0-F1) | 135 (55%) | 110 (45%) | 0.50 | 0.92 [0.71-1.18] |

Analysis was performed using the chi-square test. Bold indicates *P* values < 0.05.

*Data unavailable: fibrosis stage of 16 CLD patients.

Table 3). We did not find any difference in HCC features according to the *HSD17B13* rs72613567 genotype (Supporting Table S2). In the combined population (exploratory and validation cohorts) including 1,109 patients with HCC, *HSD17B13* TA allele carriers still had a decreased risk of HCC generally (univariate analysis, OR = 0.84; 95% CI = 0.73-0.98; *P* = 0.03 and multivariate analysis, OR = 0.77; 95% CI = 0.64-0.93; *P* = 0.006) and also in the subgroup of patients with ALD (univariate analysis, OR = 0.66; 95% CI = 0.51-0.84; *P* = 0.0008 and multivariate analysis, OR = 0.59; 95% CI = 0.44-0.79; *P* = 0.0004; Table 3; Fig. 3D). The protective effect *HSD17B13* TA allele was confirmed in the subgroups of patients with alcoholic-related cirrhosis (Fig. 3E).

Discussion

In this study, we showed that the loss-of-function variant rs72613567 of *HSD17B13* was protective against development of HCC in ALD. The original GWAS study was designed to identify genetic markers predisposing to ALD and NAFLD, but was not designed to assess the risk of HCC in patients with CLD.⁽²⁰⁾ By comparing a population of patients with CLD with a control group of healthy individuals, we confirmed that *HSD17B13* rs72613567 decreased the risk of ALD and NALFD. We also found a strong protective effect against ALD- and NAFLD-related cirrhosis, as already reported in the initial study. Based on the analysis of 88 patients, a recent report has

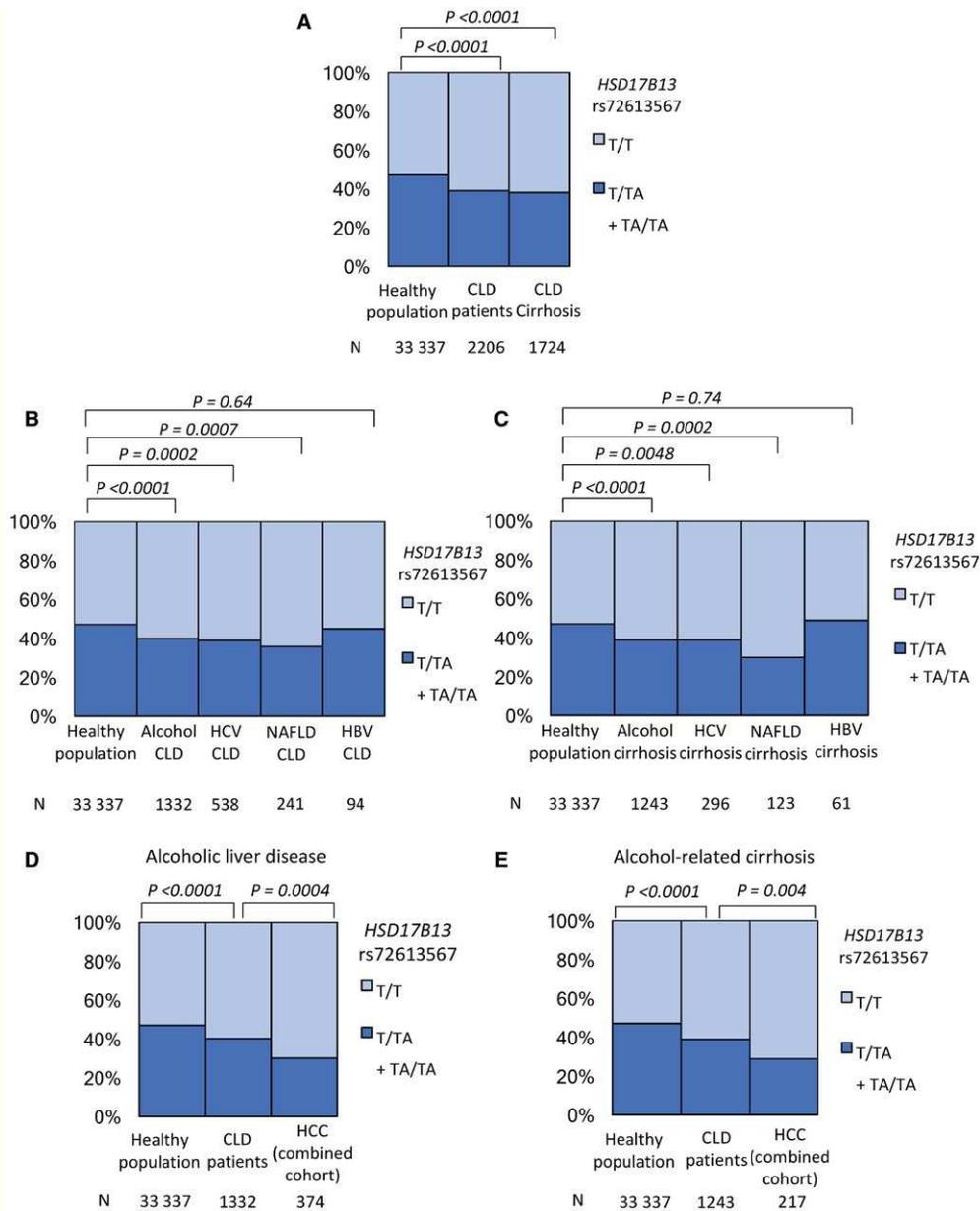


FIG. 3. Risk of cirrhosis and HCC according to the genotype distribution of *HSD17B13* rs72613567 in the combined population. We compared the genotype distribution of *HSD17B13* rs72613567 (A) between CLD patients and healthy population from ExAC consortium; (B) between CLD patients according to the different etiologies and healthy population from ExAC consortium; and (C) between CLD patients with cirrhosis according to the different etiologies and healthy population from ExAC consortium. Analysis was performed using the chi-square test. We compared the genotype distribution of *HSD17B13* rs72613567 (D) between CLD patients without HCC and HCC patients (combined cohort) and (E) between alcohol-related patients with cirrhosis without HCC and HCC patients (combined cohort). Analysis was performed using logistic regression adjusted for age, sex, and degree of fibrosis on the nontumor liver. Abbreviation: AL, alcohol-related cirrhosis.

suggested a potential protective effect of this SNP on severe fibrosis in CHC.⁽²⁸⁾ Our data also suggested a protective effect of *HSD17B13* rs72613567 on HCV-related liver disease, including HCV-related cirrhosis.

We also identified an association between HCC and a lower incidence of TA allele in an exploratory cohort of HCC patients analyzed by WES and compared to a healthy population. To bypass

TABLE 3. Risk of HCC According to HSD17B13 rs72613567 Genotypes

| Etiology | Genotype | CLD Patients* | HCC Patients* | Univariate | | Multivariate | |
|-------------------|------------|---------------|---------------|---------------|-------------------------|---------------|-------------------------|
| | | | | P Value | OR (95% CI) | P Value | OR (95% CI) |
| Validation cohort | | | | | | | |
| All | T/T | 1,342 (61%) | 527 (64%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 864 (39%) | 297 (36%) | 0.12 | 0.88 [0.74-1.03] | 0.07 | 0.83 [0.68-1.02] |
| Alcohol | T/T | 806 (60%) | 185 (68%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 526 (40%) | 88 (32%) | 0.02 | 0.73 [0.55-0.96] | 0.005 | 0.64 [0.47-0.87] |
| NAFLD | T/T | 154 (64%) | 86 (63%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 87 (36%) | 51 (37%) | 0.83 | 1.05 [0.68-1.62] | 0.89 | 0.96 [0.56-1.65] |
| HCV | T/T | 329 (61%) | 121 (60%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 209 (39%) | 81 (40%) | 0.76 | 1.05 [0.76-1.47] | 0.96 | 1.01 [0.71-1.44] |
| HBV | T/T | 52 (55%) | 51 (61%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 42 (45%) | 32 (39%) | 0.41 | 0.78 [0.43-1.42] | 0.23 | 0.67 [0.35-1.29] |
| Combined cohort | | | | | | | |
| All | T/T | 1,342 (61%) | 717 (64%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 864 (39%) | 389 (36%) | 0.03 | 0.84 [0.73-0.98] | 0.006 | 0.77 [0.64-0.93] |
| Alcohol | T/T | 806 (60%) | 262 (70%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 526 (40%) | 112 (30%) | 0.0008 | 0.66 [0.51-0.84] | 0.0004 | 0.59 [0.44-0.79] |
| NAFLD | T/T | 154 (64%) | 122 (65%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 87 (36%) | 65 (35%) | 0.77 | 0.94 [0.63-1.41] | 0.36 | 0.79 [0.48-1.30] |
| HCV | T/T | 329 (61%) | 156 (60%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 209 (39%) | 103 (40%) | 0.80 | 1.04 [0.78-1.41] | 0.95 | 1.01 [0.72-1.41] |
| HBV | T/T | 52 (55%) | 67 (59%) | — | 1 | — | 1 |
| | T/TA+TA/TA | 42 (45%) | 46 (41%) | 0.56 | 0.85 [0.49-1.48] | 0.40 | 0.77 [0.42-1.42] |

Univariate analysis was performed using the chi-square test, then adjusted by sex, age, stage of fibrosis, and etiology by logistic regression analysis (multivariate analysis). Significant P values and relative OR values are in bold style.

*Data unavailable: etiology of 2 HCC patients in exploratory cohort and 12 HCC patients in validation cohort; fibrosis stage of 5 HCC patients in validation cohort.

potential bias related to uses of the healthy population as control, we used a large validation cohort of HCC and control patients with CLD to assess the role of *HSD17B13* rs72613567 in liver carcinogenesis. We identified a protective effect of the T/TA and TA/TA genotypes against HCC development on ALD even after adjustment for confounding factors, such as age, sex, and fibrosis, in the nontumor liver. Interestingly, the *HSD17B13* rs72613567 TA allele remained protective when restraining the analysis to alcohol-related patients with cirrhosis. However, patients with alcohol-related HCC were more frequently male and were older than patients with ALD, but without HCC. Even if the protective effect of *HSD17B13* rs72613567 against HCC remained significant after adjustment for age and sex, a larger cohort of patients will be required to validate these results. Moreover, the low number of patients with HCC on CHB hampered a robust

conclusion on the effect of *HSD17B13* rs72613567 on liver carcinogenesis attributed to HBV.

Interestingly, a previous study showed that body mass index (BMI) could have an impact on the effect of the *HSD17B13* variant in NAFLD.⁽²¹⁾ Moreover, it is well known that BMI and diabetes type 2 could increase the risk of HCC development in patients with CLD.⁽²⁹⁾ Unfortunately, a large proportion of patients had missing values for these clinical variables which did not allow a robust analysis. Further studies are warranted to clarify this point.

More data will be useful to understand the consequence of loss of function of *HSD17B13* and how this could decrease the risk of progression to cirrhosis and HCC development.⁽³⁰⁾

Interestingly, several studies showed a relation between *HSD17B13* polymorphisms and abnormal liver test. *HSD17B13* rs72613567 TA variant was associated with decreased aminotransferase levels in an allele

dose-dependent manner.⁽²⁰⁾ Another study showed that the minor allele of *HSD17B13* rs10433879, that is, in a close proxy with *HSD17B13* rs72613567 ($R^2 = 0.985$; $P = 0.0024$), was associated with lower alanine aminotransferase in excessive drinkers.⁽³¹⁾

Moreover, several studies showed that *HSD17B13* expression was lower in HCC tumor tissue than in nontumor adjacent tissue.⁽³²⁻³⁵⁾ One study suggested that a low level of *HSD17B13* mRNA in nontumor liver was associated with a higher rate of tumor recurrence after liver resection of HCC, but these results need external validation.⁽³⁵⁾ Currently, the exact function of *HSD17B13* needs to be better explained. These data also suggest that therapeutic modulation by inhibition of *HSD17B13* could be a strategy to prevent progression of CLD and decrease HCC incidence in patients with ALD. Patients harboring the at-risk allele of *PNPLA3* rs738409 could be the ideal candidate for this chemopreventive strategy. More generally, our results underline the necessity of functional studies dedicated to understanding the potential interactions between the different constitutional variants leading to HCC development.

In conclusion, the rs72613567 loss-of-function variant in *HSD17B13* is protective against fibrosis progression and HCC development in ALD.

Acknowledgment: We thank Benedict Monteiro for the critical reviewing of the manuscript.

REFERENCES

- 1) Sud A, Kinnersley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives. *Nat Rev Cancer* 2017;17:692-704.
- 2) Zhu H, Wu J, Shen X. Genome-wide association study: new genetic insights into HBV/HCV-related hepatocellular carcinoma genomes. *Scand J Gastroenterol* 2017;52:209-215.
- 3) Zhang H, Zhai Y, Hu Z, Wu C, Qian J, Jia W, et al. Genome-wide association study identifies 1p36.22 as a new susceptibility locus for hepatocellular carcinoma in chronic hepatitis B virus carriers. *Nat Genet* 2010;42:755-758.
- 4) Kumar V, Kato N, Urabe Y, Takahashi A, Muroyama R, Hosono N, et al. Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma. *Nat Genet* 2011;43:455-458.
- 5) Jiang DK, Sun J, Cao G, Liu Y, Lin D, Gao YZ, et al. Genetic variants in *STAT4* and *HLA-DQ* genes confer risk of hepatitis B virus-related hepatocellular carcinoma. *Nat Genet* 2013;45:72-75.
- 6) Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in *PNPLA3* confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2008;40:1461-1465.
- 7) Guyot E, Sutton A, Rufat P, Laguillier C, Mansouri A, Moreau R, et al. *PNPLA3* rs738409, hepatocellular carcinoma occurrence and risk model prediction in patients with cirrhosis. *J Hepatol* 2013;58:312-318.
- 8) Buch S, Stickel F, Trepo E, Way M, Herrmann A, Nischalke HD, et al. A genome-wide association study confirms *PNPLA3* and identifies *TM6SF2* and *MBOAT7* as risk loci for alcohol-related cirrhosis. *Nat Genet* 2015;47:1443-1448.
- 9) Seko Y, Sumida Y, Tanaka S, Mori K, Taketani H, Ishiba H, et al. Development of hepatocellular carcinoma in Japanese patients with biopsy-proven non-alcoholic fatty liver disease: association between *PNPLA3* genotype and hepatocarcinogenesis/fibrosis progression. *Hepatol Res* 2017;47:1083-1092.
- 10) Stickel F, Buch S, Nischalke HD, Weiss KH, Gotthardt D, Fischer J, et al. Genetic variants in *PNPLA3* and *TM6SF2* predispose to the development of hepatocellular carcinoma in individuals with alcohol-related cirrhosis. *Am J Gastroenterol* 2018;113:1475-1483.
- 11) Yang J, Trepo E, Nahon P, Cao Q, Moreno C, Letouze E, et al. *PNPLA3* and *TM6SF2* variants as risk factors of hepatocellular carcinoma across various etiologies and severity of underlying liver diseases. *Int J Cancer* 2019;144:533-544.
- 12) Burza MA, Pirazzi C, Maglio C, Sjöholm K, Mancina RM, Svensson PA, et al. *PNPLA3* I148M (rs738409) genetic variant is associated with hepatocellular carcinoma in obese individuals. *Dig Liver Dis* 2012;44:1037-1041.
- 13) Kozlitina J, Smagris E, Stender S, Nordestgaard BG, Zhou HH, Tybjaerg-Hansen A, et al. Exome-wide association study identifies a *TM6SF2* variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2014;46:352-356.
- 14) Liu YL, Reeves HL, Burt AD, Tiniakos D, McPherson S, Leathart JB, et al. *TM6SF2* rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. *Nat Commun* 2014;5:4309.
- 15) Yen YH, Tsai MC, Wu CK, Chang KC, Hung CH, Chiu KW, et al. Association between *PNPLA3* (rs738409 C>G) variant and hepatocellular carcinoma in Asian chronic hepatitis C patients: a longitudinal study. *J Formos Med Assoc* 2018;117:833-840.
- 16) Nakaoka K, Hashimoto S, Kawabe N, Nitta Y, Murao M, Nakano T, et al. *PNPLA3* I148M associations with liver carcinogenesis in Japanese chronic hepatitis C patients. *Springerplus* 2015;4:83.
- 17) He S, McPhaul C, Li JZ, Garuti R, Kinch L, Grishin NV, et al. A sequence variation (I148M) in *PNPLA3* associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. *J Biol Chem* 2010;285:6706-6715.
- 18) Min HK, Sookoian S, Pirola CJ, Cheng J, Mirshahi F, Sanyal AJ. Metabolic profiling reveals that *PNPLA3* induces widespread effects on metabolism beyond triacylglycerol remodeling in Huh-7 hepatoma cells. *Am J Physiol Gastrointest Liver Physiol* 2014;307:G66-G76.
- 19) Pirazzi C, Valenti L, Motta BM, Pingitore P, Hedfalk K, Mancina RM, et al. *PNPLA3* has retinyl-palmitate lipase activity in human hepatic stellate cells. *Hum Mol Genet* 2014;23:4077-4085.
- 20) Abul-Husn NS, Cheng X, Li AH, Xin Y, Schurmann C, Stevis P, et al. A protein-truncating *HSD17B13* variant and protection from chronic liver disease. *N Engl J Med* 2018;378:1096-1106.
- 21) Pirola CJ, Garaycoechea M, Flichman D, Arrese M, San Martino J, Gazzi C, et al. Splice variant rs72613567 prevents worst histologic outcomes in patients with nonalcoholic fatty liver disease. *J Lipid Res* 2019;60:176-185.
- 22) Su W, Wang Y, Jia X, Wu W, Li L, Tian X, et al. Comparative proteomic study reveals 17beta-HSD13 as a pathogenic protein

- in nonalcoholic fatty liver disease. *Proc Natl Acad Sci U S A* 2014;111:11437-11442.
- 23) Ma Y, Belyaeva OV, Brown PM, Fujita K, Valles K, Karki S, et al. 17-Beta hydroxysteroid dehydrogenase 13 is a hepatic retinol dehydrogenase associated with histological features of non-alcoholic fatty liver disease. *HEPATOLOGY* 2019;69:1504-1519.
 - 24) Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology* 1996;24:289-293.
 - 25) European Association for the Study of the Liver; European Organization for Research and Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2012;56:908-943.
 - 26) 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. *Nature* 2010;467:1061-1073.
 - 27) 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature* 2015;526:68-74.
 - 28) About F, Abel L, Cobat A. HCV-associated liver fibrosis and HSD17B13. *N Engl J Med* 2018;379:1875-1876.
 - 29) N'Kontchou G, Paries J, Htar MT, Ganne-Carrie N, Costentin L, Grando-Lemaire V, et al. Risk factors for hepatocellular carcinoma in patients with alcoholic or viral C cirrhosis. *Clin Gastroenterol Hepatol* 2006;4:1062-1068.
 - 30) Ferenci P, Pfeiffenberger J, Stättermayer AF, Stauber RE, Willheim C, Weiss KH, et al. HSD17B13 truncated variant is associated with a mild hepatic phenotype in Wilson's Disease. *JHEP Reports* 2019. <https://doi.org/10.1016/j.jhepr.2019.02.007>.
 - 31) Whitfield JB, Zhu G, Madden PA, Montgomery GW, Heath AC, Martin NG. Biomarker and genomic risk factors for liver function test abnormality in hazardous drinkers. *Alcohol Clin Exp Res* 2019;43:473-482.
 - 32) Xing X, Huang Y, Wang S, Chi M, Zeng Y, Chen L, et al. Comparative analysis of primary hepatocellular carcinoma with single and multiple lesions by iTRAQ-based quantitative proteomics. *J Proteomics* 2015;128:262-271.
 - 33) Xu G, Yang F, Ding CL, Zhao LJ, Ren H, Zhao P, et al. Small nucleolar RNA 113-1 suppresses tumorigenesis in hepatocellular carcinoma. *Mol Cancer* 2014;13:216.
 - 34) Cotoi CG, Khorsandi SE, Plesea IE, Quaglia A. Whole-genome DASL gene expression profiling of hepatocellular carcinoma sub-populations isolated by laser microdissection on formalin-fixed and paraffin-embedded liver tissue samples. *Rom J Morphol Embryol* 2012;53:893-902.
 - 35) Chen J, Zhuo JY, Yang F, Liu ZK, Zhou L, Xie HY, et al. 17-beta-hydroxysteroid dehydrogenase 13 inhibits the progression and recurrence of hepatocellular carcinoma. *Hepatobiliary Pancreat Dis Int* 2018;17:220-226.

Author names in bold designate shared co-first authorship.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.30623/supinfo.

Supplementary materials of Article 2

Whole-Exome Sequencing

Sequence capture, enrichment, and elution from non-HCC liver counterpart genomic DNA from 285 patients were performed by IntegraGen (Evry, France) using Agilent in-solution enrichment with their biotinylated oligonucleotides probes library (SureSelect Human All-Exon kit v5+UTRs-75Mb and SureSelect Clinical Research Exome V2 - 64 Mb; Agilent Technologies, Santa Clara, CA) according to manufacturer's instruction. Eluted-enriched DNA sample was sequenced on an Illumina HiSeq 2000 sequencer as paired-end 75-bp reads. Image analysis and base calling was performed using Illumina Real Time Analysis Pipeline, version 1.14, with default parameters. Raw sequences were aligned to the human reference genome (hg19) using the Maximum Exact Matches algorithm in Burrows Wheeler Aligner (BWA) (1) of the Human genome using BWA. PCR duplicate removal was performed using Picard tools (broadinstitute.github.io/picard), local indel realignment and base quality recalibration with the Genome Analysis Software Kit (GATK) (2), variants were called using GATK HaplotypeCaller (2). Resulting high quality variants were annotated using the Variant Effect Predictor (VEP) toolset (3). We selected variants in the following 20 hydroxysteroid genes (*HSD3B2*, *HSD3B1*, *HSD17B7*, *HSD11B1*, *HSD17B13*, *HSD17B11*, *HSD17B4*, *HSD17B8*, *HSD17B3*, *HSDL2*, *HSD17B7P2*, *HSD17B12*, *HSD17B6*, *HSD3B7*, *HSD11B2*, *HSD17B2*, *HSDL1*, *HSD17B1*, *HSD11B1L*, *HSD17B14*) predicted to be damaging (nonsense, frameshift, splice or missense) by VEP. Samples for a given a variant with a depth of coverage < 8 or a genotype quality < 20 were removed as recommended (4), using a homemade python (version 2.7.6) script (Python Software Foundation. Python Language Reference, version 2.7.6 (<http://www.python.org>)). Only variants with a genotyping call rate > 95% were considered in the analysis.

1. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. [Internet]. *Bioinformatics* 2010;26:589–95. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20080505>
2. DePristo MA, Banks E, Poplin R, *et al*. A framework for variation discovery and genotyping using next-generation DNA sequencing data [Internet]. *Nat. Genet.* 2011;43:491–498. Available from: <http://www.nature.com/articles/ng.806>
3. McLaren W, Gil L, Hunt SE, *et al*. The Ensembl Variant Effect Predictor [Internet]. *Genome Biol.* 2016;17:122. Available from: <http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-0974-4>
4. Carson AR, Smith EN, Matsui H, *et al*. Effective filtering strategies to improve data quality from population-based whole exome sequencing studies [Internet]. *BMC Bioinformatics* 2014;15:125. Available from: <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-15-12>

Supplementary Table 1: Association of WES candidate damaging variants in hydroxysteroid genes with HCC

| Variant ID | Chromosome | Position | Reference allele | Alternate allele | Gene | Location | allele frequency in cases* | allele frequency in controls* | OR | 95% CI | P |
|-------------|------------|-----------|------------------|------------------|-----------------|---------------|----------------------------|-------------------------------|--------|-------------|--------------------------|
| rs72613567 | 4 | 88231392 | T | TA | <i>HSD17B13</i> | Splice region | 0.129 | 0.2682 | 0.40 | 0.25-0.60 | 1.52 x 10 ⁻⁰⁶ |
| rs1143650 | 5 | 118832244 | C | G | <i>HSD17B4</i> | Missense | 0.00495 | 0.00015 | 47.201 | 0.04-369.88 | 0.024 |
| rs1047303 | 1 | 120057246 | C | A | <i>HSD3B1</i> | Missense | 0.613 | 0.6743 | 0.75 | 0.55-1.01 | 0.054 |
| rs35299026 | 19 | 49318380 | G | A | <i>HSD17B14</i> | Missense | 0.0808 | 0.04963 | 1.64 | 0.92-2.74 | 0.072 |
| rs6205 | 1 | 120057002 | T | C | <i>HSD3B1</i> | Missense | 0.0099 | 0.00293 | 2.94 | 0.35-10.89 | 0.151 |
| rs147402365 | 17 | 40706596 | C | T | <i>HSD17B1</i> | Missense | 0.015 | 0.00878 | 1.90 | 0.39-5.66 | 0.215 |
| rs25640 | 5 | 118811533 | G | A | <i>HSD17B4</i> | Missense | 0.42 | 0.4579 | 0.84 | 0.62-1.12 | 0.228 |
| rs113690224 | 17 | 40706679 | C | T | <i>HSD17B1</i> | Splice region | 0.0204 | 0.01299 | 1.67 | 0.45-4.36 | 0.312 |
| rs11555762 | 11 | 43876698 | C | T | <i>HSD17B12</i> | Missense | 0.342 | 0.3054 | 1.16 | 0.85-1.56 | 0.322 |
| rs9938550 | 16 | 30999142 | A | G | <i>HSD3B7</i> | Missense | 0.663 | 0.6343 | 1.12 | 0.83-1.53 | 0.464 |
| rs2066479 | 9 | 98997810 | C | T | <i>HSD17B3</i> | Missense | 0.04 | 0.04975 | 0.74 | 0.32-1.49 | 0.526 |
| rs2066480 | 9 | 99064296 | C | T | <i>HSD17B3</i> | Missense | 0.0099 | 0.00739 | 1.32 | 0.16-4.85 | 0.667 |
| rs34212827 | 16 | 30999434 | T | C | <i>HSD3B7</i> | Missense | 0.0099 | 0.00756 | 1.24 | 0.15-4.54 | 0.679 |
| rs11539471 | 5 | 118860938 | T | C | <i>HSD17B4</i> | Missense | 0.0842 | 0.07623 | 1.10 | 0.63-1.81 | 0.691 |
| rs117437228 | 16 | 82124518 | A | G | <i>HSD17B2</i> | Missense | 0.00495 | 0.01136 | 0.41 | 0.01-2.31 | 0.738 |
| rs11205 | 5 | 118861713 | A | G | <i>HSD17B4</i> | Missense | 0.426 | 0.4066 | 1.04 | 0.78-1.39 | 0.775 |
| rs4378600 | 16 | 84158248 | G | C | <i>HSDL1</i> | Missense | 0.282 | 0.2922 | 0.97 | 0.70-1.32 | 0.877 |
| rs8110220 | 19 | 49339098 | T | C | <i>HSD17B14</i> | Splice region | 0.248 | 0.2396 | 1.01 | 0.72-1.40 | 0.935 |
| rs41280175 | 9 | 115167994 | A | T | <i>HSDL2</i> | Missense | 0.0099 | 0.01386 | 0.72 | 0.09-2.66 | 1 |
| rs144261994 | 1 | 120050210 | G | C | <i>HSD3B1</i> | Missense | 0 | 0.00026 | - | - | 1 |
| rs202219444 | 1 | 209879292 | G | A | <i>HSD11B1</i> | Splice region | 0 | 0.00343 | - | - | 1 |
| rs80182459 | 4 | 88235096 | CG | C | <i>HSD17B13</i> | Frameshift | 0 | 0.00045 | - | - | 1 |
| rs34491699 | 6 | 33173607 | A | G | <i>HSD17B8</i> | Splice region | 0 | 0.00066 | - | - | 1 |

*frequency of the alternate allele

Supplementary Table 2: Association of HSD17B13 rs72613567 with HCC features

| Tumor features | Data available | Patients with HCC | T/T | T/TA+TA/TA | p-value |
|-------------------------------|-----------------------|--------------------------|------------|-------------------|----------------|
| Alpha-fetoprotein >20UI | 694 | 252 (36%) | 165 (65%) | 87 (35%) | 0.93 |
| Main nodule > 5 cm | 814 | 406 (50%) | 264 (65%) | 142 (35%) | 0.75 |
| Unique nodule | 750 | 609 (81%) | 392 (64%) | 217 (36%) | 0.73 |
| Macroscopic vascular invasion | 660 | 110 (17%) | 77 (70%) | 33 (30%) | 0.19 |
| Microscopic vascular invasion | 662 | 343 (52%) | 227 (66%) | 116 (34%) | 0.35 |
| Differentiation | 735 | | | | |
| good | | 203 (28%) | 122 (60%) | 81 (40%) | |
| medium | | 425 (58%) | 282 (66%) | 143 (34%) | 0.30 |
| bad | | 107 (15%) | 70 (65%) | 37 (35%) | |

HCC features were available in a subset of patients with HCC (n=814/1109). Analysis was performed using Chi-square test.

Discussion

1. Interaction between environmental and genetic risk factors of HCC

1.1 Validation of SNPs identified in GWAS of HCC

Following the identification of SNPs linked with the susceptibility of HBV- or HCV-related HCC in GWAS, several groups investigated whether these variants were associated with disease progression and development of virus-related HCC. However, as presented in the introduction, these SNPs often failed to be validated in other GWAS and there are discrepancies in candidate-gene studies. Here in our study, we also tested 7 SNPs (*DEPDC5* rs1012068, *GRIK1* rs455804, *KIF1B* rs17401966, *STAT4* rs7574865, *MICA* rs2596542, *DLC1* rs2275959 and *DDX18* rs2551677) identified as associated with HCC risk in GWAS. First, we compared the genotype distributions of these SNPs between overall patients with HCC and overall patients with CLD and without HCC, but we did not identify any significant association even after adjustment for common confounders as age, gender, etiology and severity of underlying liver diseases. Considering the heterogeneity of underlying liver diseases, next we performed subgroup analysis within the same etiology as reported in the original GWAS. Among them, only *STAT4* rs7574865 was validated as associated with the risk of HBV-related HCC in European ancestry. The association was still significant after adjustment for age, gender and severity of liver disease. *STAT4* is a transcription factor and a critical regulator of immune responses, by transmitting signals from interleukin-12 and type I interferon (IFN- α or IFN- β) and could induce IFN- γ production (Nguyen et al., 2002; Thierfelder et al., 1996). IFN- γ is a pleiotropic cytokine that has a crucial role in host defense. Impaired activation of IFN- γ via *STAT4* may reduce its antiviral and antitumor activities (Horras, Lamb, & Mitchell, 2011; Saha, Jyothi Prasanna, Chandrasekar, & Nandi, 2010). In original GWAS, it was found that the at-risk G allele of *STAT4* rs7574865 was associated with a lower mRNA level of *STAT4*, compared to the T allele (Jiang et al., 2013). It provides a potential hypothesis how *STAT4* SNP contributes to a decreased host defense in antiviral activities and an increased risk of HBV-related HCC, but additional functional studies are required to elucidate this issue.

There are some possible reasons for the discrepant associations of other SNPs between our study and the original GWAS studies: 1) the differences in ethnicity between Asians

in all the GWAS and Europeans in our study; 2) the differences in term of control group (healthy population or patient with underlying chronic liver diseases but without adjustment on fibrosis in GWAS); 3) the limited number of patients with viral infection, especially HBV infection, in our study cohort that leads to a low statistic power. Therefore, further studies considering these points are warranted to clarify the association between these genetic variants and the risk of HCC across various populations.

1.2 Association between *PNPLA3/TM6SF2* variants and increased risk of alcohol-related HCC

Until now, rs738409 in *PNPLA3* is the best characterized and deeply investigated genetic variant tightly associated with risk of HCC. *PNPLA3* gene, containing 9 exons, is located on chromosome 22 (22q13.31). Rs738409 is a non-synonymous SNP with a switch from C to G which results in amino acid substitution of an isoleucine (I) to a methionine (M) at the position 148 of the coding sequence. In year 2008, it was firstly identified as an independent risk factor of hepatic fat accumulation in a genome-wide association study. The at-risk G allele of *PNPLA3* was more frequent in Hispanics population who are more susceptible to NAFLD than African-Americans and European-Americans, partly explaining the inter-ethnic heterogeneity of the risk of NAFLD. However, even though obesity and diabetes are two of the main risk factors of NAFLD, in the study, no association was observed between *PNPLA3* rs738409 and BMI or diabetes (Romeo et al., 2008).

Furthermore, *PNPLA3* rs738409 was not only associated with hepatic steatosis, but also strongly related with elevated serum ALT and AST level and progression to fibrosis or cirrhosis, indicating that *PNPLA3* rs738409 was an independent risk factor of severity of the liver disease (Kitamoto et al., 2013; Kotronen et al., 2009; Valenti, Al-Serri, et al., 2010; Valenti, Alisi, et al., 2010).

In 2014, an exome-wide association study identified a novel non-synonymous SNP in *TM6SF2* (rs58542926 c.449 C>T, p.Glu167Lys). It was associated with higher circulating level of ALT but lower level of very-low density lipoprotein (VLDL) and TG, which

confers a susceptibility to hepatic fat accumulation and NAFLD (Kozlitina et al., 2014). Similarly, *TM6SF2* SNP was associated with advanced hepatic fibrosis/cirrhosis independently of *PNPLA3* rs738409 and other potential confounding factors (Y. L. Liu, Reeves, et al., 2014).

Then whether *PNPLA3* rs738409 and *TM6SF2* rs58542926 act as genetic risk factors in liver injury induced by other chronic liver diseases was questioned. In the past ten years, numerous candidate-gene studies together with meta-analysis showed an independent and robust association between *PNPLA3* rs738409 and alcoholic liver disease (ALD), especially alcohol-related cirrhosis across ethnicities and in different geographical regions (Atkinson, Way, McQuillin, Morgan, & Thursz, 2017; Burza et al., 2014; Dutta, 2013; Falletti et al., 2011; Friedrich et al., 2014; Kolla et al., 2018; Liangpunsakul et al., 2018; Mancina et al., 2019; Nischalke et al., 2011; Rosendahl et al., 2012; Seth, Daly, Haber, & Day, 2010; Stickel et al., 2011; Tian, Stokowski, Kershenobich, Ballinger, & Hinds, 2010; Trepo, Gustot, et al., 2011; Y. Zhang et al., 2018). Remarkably, a genome-wide association study also confirmed *PNPLA3* rs738409 as a prominent risk loci for alcohol-related cirrhosis (Buch et al., 2015). Meanwhile, it identified *TM6SF2* rs58542926 as a new risk loci for alcohol-related cirrhosis. In our study, we did the validation analysis and found the same association consistent with these literatures. By comparing with European healthy controls from 1000Genome, we found that *PNPLA3* rs738409 was significantly associated with alcoholic and non-alcoholic associated chronic liver disease and liver cirrhosis. And patients with NAFLD harbored more frequently the at-risk TT+CT genotypes of *TM6SF2*.

Several studies were performed in European and Asian populations to assess the impact of *PNPLA3* rs738409 and *TM6SF2* rs58542926 on steatosis and fibrosis progression in patients with chronic hepatitis C. After adjustment for other confounders (age, sex, body mass index, alcohol consumption and diabetes), the association between *PNPLA3* rs738409 and the risk of steatosis and fibrosis progression remained but was less strong than that observed in ALD or NAFLD. Meta-analysis showed that the effect of *PNPLA3* rs738409 was more evident in European population than in Asians, but a controversy still exists (J. H. Fan et al., 2016; Manchiero et al., 2017; Nischalke et al., 2011; Trepo, Pradat, et al., 2011; Valenti et al., 2011; Yasui et al., 2015; H. Zhang et al., 2015). *TM6SF2* SNP

was also reported to impact the steatosis severity and associated with liver damage and fibrosis in patients with CHC, but only with the genotype 1 of HCV (Coppola et al., 2015; Milano et al., 2015; Petta et al., 2016). In univariate analysis of our study, the genotype distribution was not different between healthy population and patients with chronic hepatitis C or B. It could be explained by the difference between ethnicities, lack of adjustment for common confounders or by the negligible and indirect effect in CHC/CHB as reported in previous studies. Recently, several studies reported *PNPLA3* and *TM6SF2* SNPs as independent risk factors for fibrosis progression to cirrhosis in HIV/HCV co-infected patients (Jimenez-Sousa et al., 2016; Nunez-Torres et al., 2016; Sagnelli et al., 2016). It indicates that *PNPLA3* and *TM6SF2* SNPs might only be risk factors in a subgroup of patients with CHC. Anyhow, these studies on chronic liver diseases with different etiologies (NAFLD and ALD) suggest that *PNPLA3* and *TM6SF2* play a role in the common biological pathway leading to liver injury and cirrhosis.

We investigated the role *PNPLA3* rs738409 together with *TM6SF2* rs58542926 in the development of HCC. In our study, the comparison between overall patients with and without HCC but underlying CLDs showed that both *PNPLA3* and *TM6SF2* SNPs were independent risk factors of HCC development. In this multivariate analysis, we had adjusted for the well-known confounders like age, gender, different etiologies and severity of CLD besides BMI and diabetes due to lack of data. However, even though obesity and diabetes are two important risk factors of HCC, in the previous study, the association between *PNPLA3* rs738409 and BMI or diabetes was negligible (Romeo et al., 2008). And another study confirmed that insulin sensitivity was related to liver fat content independently of *PNPLA3* genotype status (Kotronen et al., 2009).

In line with a large number of retrospective case-control studies, our subgroup analysis according to etiologies validated that both *PNPLA3* rs738409 and *TM6SF2* rs58542926 were independent risk factors of HCC in patients with ALD. However, well-defined prospective study is of the gold standard to assess the association between genetic traits and the development of liver cancer in order to avoid bias of case/control study. In consist with a previous study (Guyot et al., 2013), the link between *PNPLA3* rs738409 and ALD-related HCC incidence was also well validated in our prospective cohort of ALD-related cirrhosis. And we confirmed for the first time in a prospective manner the link between

TM6SF2 rs58542926 and HCC occurrence. In contrast, the association between *PNPLA3* rs738409 and the risk of HCC in patients with HCV-related cirrhosis was controversial either in European or Asian populations (Balasus et al., 2016; Ezzikouri et al., 2014; Guyot et al., 2013; Hai et al., 2017; Nakaoka et al., 2015; Nischalke et al., 2011; Valenti et al., 2011). In our study, we did not identify a significant association neither in our case-control nor in prospective cohorts. Similarly, we did not observe an obvious impact of *TM6SF2* rs58542926 on HCC incidence in patients with HCV-related cirrhosis. The negative association with HCV-related HCC in our study together with controversial results reported in literature probably indicate that the two variants have very limited role in the pathogenesis and progression of HCV-related cirrhosis.

Since the two variants were originally identified as risk factors of NAFLD, the impact on NAFLD-related HCC was also assessed. However, there is not yet a well-defined conclusion about the association between *PNPLA3/ TM6SF2* SNPs and the risk of NAFLD-related HCC due to limited number of cases analyzed in a limited number of studies (Donati et al., 2017; Y. L. Liu, Patman, et al., 2014; Y. L. Liu, Reeves, et al., 2014; Seko et al., 2017). In our study, we observed that the frequency of at-risk genotypes of *PNPLA3* and *TM6SF2* SNPs were higher in HCC patients. However, due to a limited number of patients of NAFLD with or without HCC, the association didn't reach a statistic significance. Regarding the increasing incidence of HCC due to NAFLD, additional studies including a larger number of patients are required to fully investigate the role of *PNPLA3* rs738409 and *TM6SF2* rs58542926 in liver carcinogenesis due to NAFLD. Although the *PNPLA3* and *TM6SF2* SNPs were less studied in patients with CHB, there seems to be a lack of association with the risk of HBV-related HCC (Brouwer et al., 2015). But it would be informative to further confirm whether the two variants play also a role in carcinogenesis in patients with CHB.

Moreover, a subset of HCC developed in patients without liver cirrhosis and without any known risk factors of HCC. However, due to lack of effective predisposing markers, they were rarely enrolled into a surveillance program. And there are few data about the impact of *PNPLA3* variant on HCC developed on non-fibrotic liver. Only one study suggested this association in a case-control study including 25 HCC developed on non-cirrhotic liver (F0-F2) in patients with HCV or ALD (Donati et al., 2017). In our study, we

confirmed this association in a larger series of 212 patients with HCC developed on non-fibrotic liver (F0-F1), mostly without any etiologies of liver diseases, using two different control populations composed of a total of 2,484 healthy subjects. It suggested a direct role of *PNPLA3* in the liver carcinogenesis.

Normally, a complex trait is influenced by multiply genetic variants with gene-gene interactions. *PNPLA3* rs738409 and *TM6SF2* rs58542926 shared a similarity in their heritability with risk of liver steatosis, fibrosis, and progression to HCC in patients with ALD and NAFLD. However, no gene-gene interaction has ever been identified between them. In line with previous studies (Falleti et al., 2016; Stickel et al., 2018b), we also observed a progressive increase of HCC risk with increasing number of the at-risk alleles of *PNPLA3* and *TM6SF2* SNPs, probably due to an additive effect of these 2 SNPs on HCC risk.

1.3 Biological function of *PNPLA3/TM6SF2* on lipid metabolism and liver carcinogenesis

PNPLA3 is one of the nine members in patatin-like phospholipase domain-containing (*PNPLA*) family, a group of lipid-metabolizing enzymes. In human tissue, *PNPLA3* is highly expressed in the liver, followed by retina, skin, adipose tissue (Y. Huang et al., 2010; Pirazzi et al., 2014). The expression level is predominantly observed in hepatic stellate cells (HSCs) in human liver, whereas is predominately observed in hepatocytes in mice liver (Bruschi, Claudel, et al., 2017; Y. Huang et al., 2010; Pirazzi et al., 2014). Expression of *PNPLA3* is regulated by two key factors of the glycolytic and lipogenic pathways, indicating its implication in the metabolism of carbohydrates and lipids (Dubuquoy et al., 2011). Sterol regulatory element-binding proteins 1c (SREBP-1c) promotes accumulation of *PNPLA3* directly by activating *PNPLA3* transcription and indirectly by inhibiting *PNPLA3* degradation through the stimulation of fatty acid synthesis (Y. Huang et al., 2010). Regarding cellular organelles, it localizes principally on the membrane of lipid droplets (He et al., 2010). In hepatic stellate cells (HSCs), there is colocalization of *PNPLA3* with profibrogenic protein alpha-smooth muscle actin in biopsies of patients with NASH, indicating the potential role of *PNPLA3* in fibrosis progression in NAFLD (Bruschi, Claudel, et al., 2017).

Since the genetic variant of *PNPLA3* (rs738409) is robustly associated with an increased risk of liver steatosis, fibrosis progression, HCC development in ALD/NAFLD, several studies focus on *PNPLA3* function and try to understand the biological consequences of *PNPLA3* rs738409 in the liver.

In our study, we have questioned whether the different genotypes of *PNPLA3* rs738409 and *TM6SF2* rs58542926 were associated with a different level of transcription of their respective genes. We analyzed the expression of *PNPLA3* and *TM6SF2* in non-tumor tissues from 340 patients with HCC and we didn't find any association between mRNA expression level and different genotypes of *PNPLA3* rs738409 and *TM6SF2* rs58542926. Actually, several studies showed that simple genetic manipulations of *PNPLA3* either by knockout or overexpression strategies failed to influence the TG content (W. Chen, Chang, Li, & Chan, 2010; Lake et al., 2005). It suggests that the role of *PNPLA3* is probably not dependent on its expression level but on the modification of the protein itself. Consistent with the data observed in human, several studies showed that the *PNPLA3* rs738409 causes accumulation of hepatic fat. For example, expression of *PNPLA3* 148M increased TG accumulation compared to the wild type *PNPLA3*. Moreover, *PNPLA3* 148M could abolish the activity of wild type *PNPLA3* (He et al., 2010). Then, further study was performed to understand how 148M lead to *PNPLA3* loss-of-function statue. In mice model harboring the 148M variant, the posttranslational regulation of *PNPLA3* by fasting and refeeding was blunted. In addition, it was observed that inhibition of the proteasome (bortezomib) markedly increased *PNPLA3* levels in WT mice, coincident with the appearance of ubiquitylated forms of the protein. However, bortezomib did not increase *PNPLA3* levels in *PNPLA3*148M/M mice and only trace amounts of ubiquitylated *PNPLA3* were seen in these animals. It suggested that 148M variant disrupted the ubiquitylation and proteasomal degradation of *PNPLA3*, resulting in accumulation of dysfunctional *PNPLA3*, impaired mobilization of TG from LDs, and larger LDs (BasuRay, Smagris, Cohen, & Hobbs, 2017; Chamoun, Vacca, Parton, & Gruenberg, 2013; He et al., 2010). This observation explained also why simple overexpression of wild type *PNPLA3* in mice created no clear phenotype.

Regarding its robust association with the whole spectrum of liver disease, the key question relies on the mechanism involving in progression from steatosis to cirrhosis thus liver tumor. Firstly, it was found that *PNPLA3* 148M induced widespread effects on metabolism beyond its classical participation in TG remodeling. Overexpression of the *PNPLA3* 148M was associated with an increase in lactic acid, indicating a shift to anaerobic metabolism and mitochondrial dysfunction (Min et al., 2014). Second, the enrichment of long-chain fatty acids, such as n-6/n-3 polyunsaturated fatty acids (PUFAs), contributes to a proinflammatory condition (Araya et al., 2004). Strikingly, *PNPLA3* 148M was associated with increased n-6 PUFA in hepatic fat content (Santoro et al., 2012). It indicates a possible way how *PNPLA3* 148M induces liver inflammation. Then, regarding liver fibrosis, a study showed that *PNPLA3* 148M was required for the activation of HSCs, the key player in fibrosis development. *PNPLA3* 148M promotes the release of proinflammatory mediators that favor immune cell recruitment. Moreover, *PNPLA3* 148M expressing cells harbored a higher number of specific FA species and enhanced phosphorylation of JNK, resulting in PPAR γ inhibition and increased AP-1 transcriptional activity that fostered the profibrogenic features of HSCs (Bruschi, Claudel, et al., 2017). The **Figure-5** summarizes the potential biological mechanisms of *PNPLA3* rs738409 involved in the pathogenesis of liver disease progression. In our study with 563 patients with a unique HCC nodule, we searched for the link between pathological and molecular features of tumor and *PNPLA3* rs738409. We didn't notice a significant association with histological features of tumor, but patients with at-risk G allele of *PNPLA3* rs738409 were more likely to have a tumor with somatic mutation of *ARID1A*. However, the association is quite limited and the biological meaning behind remains unclear.

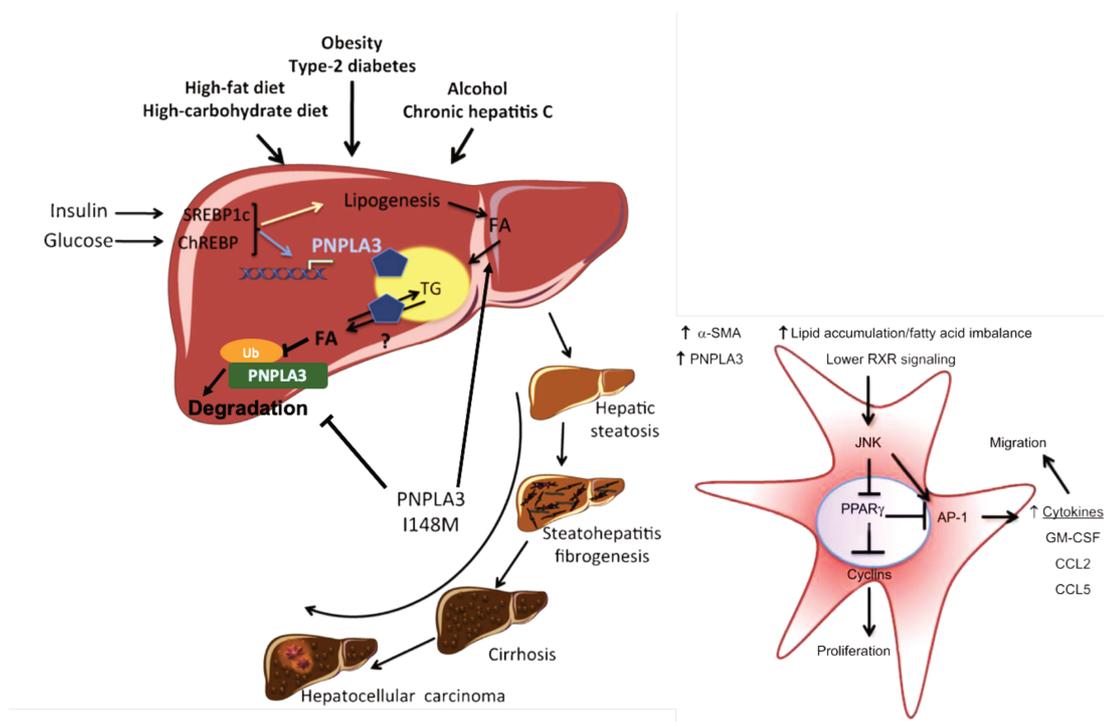


Figure 5 : The possible mechanism of PNPLA3 involved in the pathogenesis of liver disease progression.

PNPLA3 is highly expressed in human hepatocytes and hepatic stellate cells, involved in TG metabolism. Expression of PNPLA3 is regulated by glycolytic and lipogenic pathways. PNPLA3 148M could abolish the activity of wild type PNPLA3, increasing TG accumulation and disrupting the ubiquitylation and degradation on of PNPLA3. It induces liver inflammation and activates HSCs, the key player in fibrosis development, via enhancement of AP-1 and inhibition of PPAR γ . PNPLA3 I148M plays a role in the whole spectrum of liver disease, from hepatic steatosis to HCC. (Adopted and modified from (Bruschi, Tardelli, Claudel, & Trauner, 2017; Dubuquoy, Burnol, & Moldes, 2013))

TM6SF2 is expressed mainly in liver and intestine and localizes in endoplasmic reticulum and the ER-Golgi intermediate compartment. In vitro study showed that TM6SF2 is involved in lipid metabolism, particularly in secretion of TG-rich lipoprotein and affecting hepatic lipid droplet content and plasma lipid levels (Mahdessian et al., 2014). **Figure-6** showed the potential mechanism of TM6SF2 involved in lipid metabolism. *ATP binding cassette subfamily G member (ABCG) 5* and *ABCG8* are two genes promoting hepatic cholesterol elimination from the body through hepatobiliary secretion. CRISPR-Cas9 generated *TM6SF2* knockout mice altered their expression, indicating the role of *TM6SF2* in the hepatobiliary secretion of cholesterol (Y. Fan et al., 2016). These data confirmed the studies that showed the association between loss-of-function variant of *TM6SF2* rs58542926 and the accumulation of hepatic fat, a lower level of serum VLDL

and an increased risk of NAFLD. In contrast, wild type *TM6SF2* rs58542926 (CC) was associated with elevated plasma lipid content and increased risk of cardiovascular disease (Kozlitina et al., 2014; Wong, Wong, Tse, & Chan, 2014). Moreover, protein expression of *TM6SF2* was markedly decreased in the liver of NAFLD patients. And carriers with C allele of *TM6SF2* rs58542926 have 1.79-fold higher mRNA expression of *TM6SF2* than carriers with at-risk T allele. (Sookoian et al., 2015). In our study, no association between genotype and transcriptional level of non-tumor tissue was found in patients with HCC. Moreover, patients with at-risk TT + CT genotype of *TM6SF2* rs58542926 had HCC with higher rate of somatic mutations in *TERT* promoter (70% vs. 55%; $p = 0.01$), with lower rate of somatic mutations in *ARID2* (0% vs. 11%, $p = 0.02$) compared to patients with CC genotype. However, the biological significance of these weak associations is unknown. Consequently, further studies are warranted to uncover the biological function of *TM6SF2* in hepatic lipid metabolism and to provide new insights about the potential mechanism of its variant in disease progression and HCC occurrence.

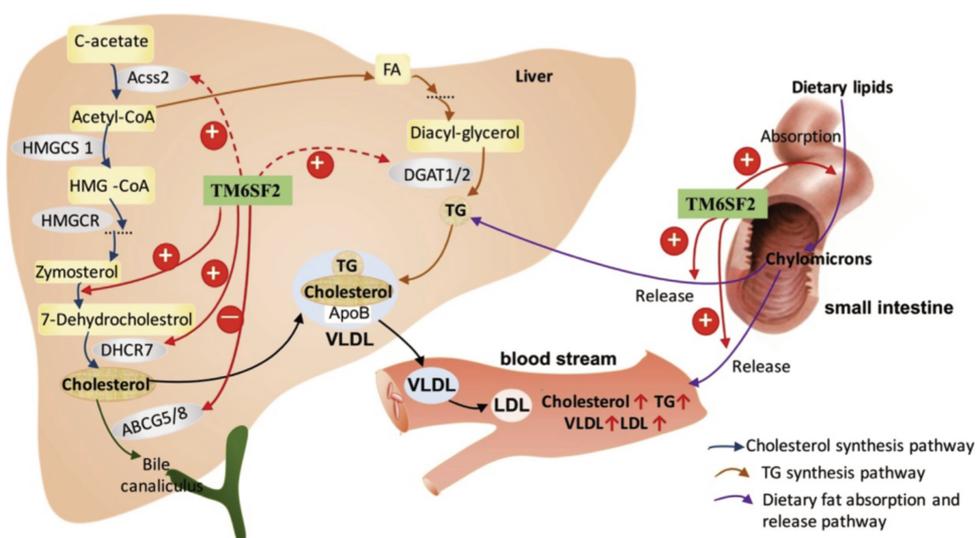


Figure 6 : The possible mechanism of *TM6SF2* involved in lipid metabolism.

TM6SF2 is mainly expressed in liver and intestine. In liver, it can promote expression of *DHCR7* that induces cholesterol biosynthesis. It could also decrease expression of *ABCG5/8* which are associated with TG and VLDL secretion. In small intestine, *TM6SF2* may promote dietary lipids absorption, increasing TG and VLDL in circulating system (T. T. Li et al., 2018).

To conclude, the association between *PNPLA3* and *TM6SF2* genetic variants and HCC development highlighted the role of lipid metabolism in liver tumorigenesis. However,

we need to admit that more data are warranted to better decipher their role in progression of liver disease and in HCC development.

1.4 Protective effect of *HSD17B13* variant on alcohol-related HCC risk and potential mechanism of action

HSD17B13 is abundant in liver tissue. Similar to *PNPLA3*, it is selectively expressed on the membrane of lipid droplets (Horiguchi, Araki, & Motojima, 2008; Su et al., 2014). Last year, an exome sequencing data of 46,544 individuals identified a SNP (rs72613567) in *HSD17B13*. Normally, *HSD17B13* rs72613567 with wild type T allele encoded transcript A, while this TA allele encoded transcript D, which was a truncated isoform with reduced enzymatic activity. This splicing variant was identified associated with reduced levels of ALT. It was also negatively associated with risk of ALD, NAFLD and especially the progression from simple steatosis to liver cirrhosis. (Abul-Husn et al., 2018). However, this variant didn't protect from simple steatosis alone. Another two SNPs, that are, in a close proxy with *HSD17B13* rs72613567 (rs6834314: $R^2=0.94$; and rs10433879: $R^2=0.985$), were also separately identified associated with lower level of liver enzyme (Chambers et al., 2011; Whitfield et al., 2019). Remarkably, the protective TA allele of rs72613567 was less frequent (0.27 vs 0.33) in Europeans than Asians. Another loss-of-function variant of *HSD17B13* (rs143404524) was associated with reduced risk of liver disease and was also less frequent in Hispanics than blacks, explaining the inter-ethnic heterogeneity of risk of liver disease (Kozlitina, Stender, Hobbs, & Cohen, 2018). In our study, we compared the genotype distribution of *HSD17B13* rs72613567 in patients with CLD versus healthy control, and validated the protective effect of *HSD17B13* rs72613567 from ALD, NAFLD and liver cirrhosis. Another two studies investigated the association between *HSD17B13* SNP and the histological features of NAFLD. They precisely confirmed that this splicing variant rs72613567 was associated with increased steatosis but decreased inflammation and fibrosis and prevented the progression of fibrosis in patients with NAFLD (Ma et al., 2019; Pirola et al., 2019). In line with another study (About, Abel, & Cobat, 2018), we also observed a protective effect against HCV-related liver disease, including HCV-related cirrhosis.

However, the associations between *HSD17B13* variants and HCC development were never fully investigated. Therefore, our study aimed to dissect the potential role of variants in *hydroxysteroid* genes in liver carcinogenesis. In the discovery study, by whole exome sequencing, we studied 23 predicted damaging variants of *HSDs* gene family in 285 patients with HCC. In comparison to healthy population, TA allele of *HSD17B13* rs72613567 was closely associated with a reduced risk of HCC development. As most of HCC develops on chronic cirrhosis along with different etiologies, using healthy population as a control may cause some bias. To avoid that, in validation study, we included a case-control cohort of patients with or without HCC but all underlying CLD. We identified that TA/T and TA/TA genotypes of *HSD17B13* rs72613567 protected from ALD-related HCC even after the adjustment for age, gender and fibrosis. The protective effect remained when we restrained the analysis only in subgroup of patients with ALD-related cirrhosis. Interestingly, a previous study showed that BMI could have an impact on the effect of the *HSD17B13* variant in NAFLD (Pirola et al., 2019). Moreover, it is well known that BMI and T2D could increase the risk of HCC development in patients with CLD (N'Kontchou et al., 2006). Unfortunately, a large proportion of patients had missing values for these clinical variables which did not allow a robust analysis. Further studies are warranted to clarify this point.

As mentioned in a previous study, the *HSD17B13* variant could mitigate the liver injury due to the at-risk G allele of *PNPLA3* rs738409 (Abul-Husn et al., 2018). Therefore, we investigated whether it could also mitigate the risk of HCC induced by the at-risk G allele of *PNPLA3* in patients with ALD with or without cirrhosis (shown in **Figure 7**). A protective effect of the TA allele of *HSD17B13* was significant in patients carrying wild type or one at-risk G allele of *PNPLA3* rs738409. In patient with GG genotype of *PNPLA3* rs738409, incidence of HCC was lower in patients harboring the protective TA allele than those with wild type *HSD17B13* rs72613567. However, the effect was no more statistically significant. It might be explained for two reasons. First, the number of patients with homozygous of minor allele was limited that led to a lack of statistical power. Second, the HCC risk with two G allele of *PNPLA3* rs738409 (odd ratio, OR=3.9[2.5-6.1]) was quite high whereas the protective effect of TA allele of *HSD17B13* (OR=0.6[0.4-0.8]) was quite moderate. Therefore, in order to better use these genetic variants as HCC predisposing markers, further prospective study is required.

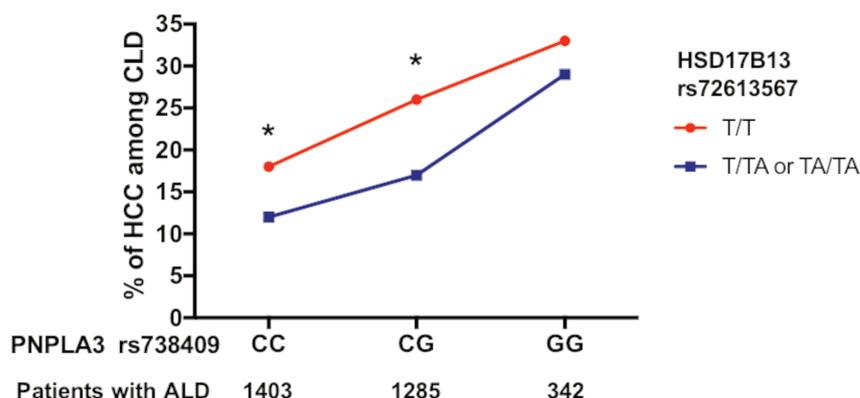


Figure 7 : Risk of ALD-related HCC according to the combination of at-risk G allele of *PNPLA3* rs738409 and the protective TA allele of *HSD17B13* rs72613567.

Analysis was performed using logistic regression to adjust for age, gender and degree of fibrosis on the non-tumor liver. * *P* Value < 0.05

In addition, more data will be useful to understand the biological function of *HSD17B13* and to understand how the genetic variant of *HSD17B13* could protect against CLDs in ALD/NALFD/CHC and progression to HCC in alcoholic liver diseases.

HSD17B13 belongs to an enzyme family (*HSD17Bs*) catalyzing the interconversion between the weak 17keto-hydroxysteroids and the highly active 17 β -hydroxysteroids. Among the fifteen family members, most of them regulate the metabolism of sex hormones, and some members of *HSD17Bs* are involved in fatty acid metabolism, cholesterol biosynthesis and bile acid production (Saloniemi, Jokela, Strauss, Pakarinen, & Poutanen, 2012). Unfortunately, little is known about function of *HSD17B13*. SREBP-1c, a protein regulated by insulin, saturated fatty acid and LXR agonists, could increase the expression of *HSD17B13*, which promotes SREBP-1 maturation in a feedback loop (**Figure-8**) (Su et al., 2019). It may partly explain why this loss-of-function variant of *HSD17B13* was associated with decreased mRNA expression of *PNPLA3* (Abul-Husn et al., 2018), whose translational expression is also directly regulated by SREBP-1c. However, we did not observe any association between *HSD17B13* rs72613567 and mRNA expression of *PNPLA3* in our study.

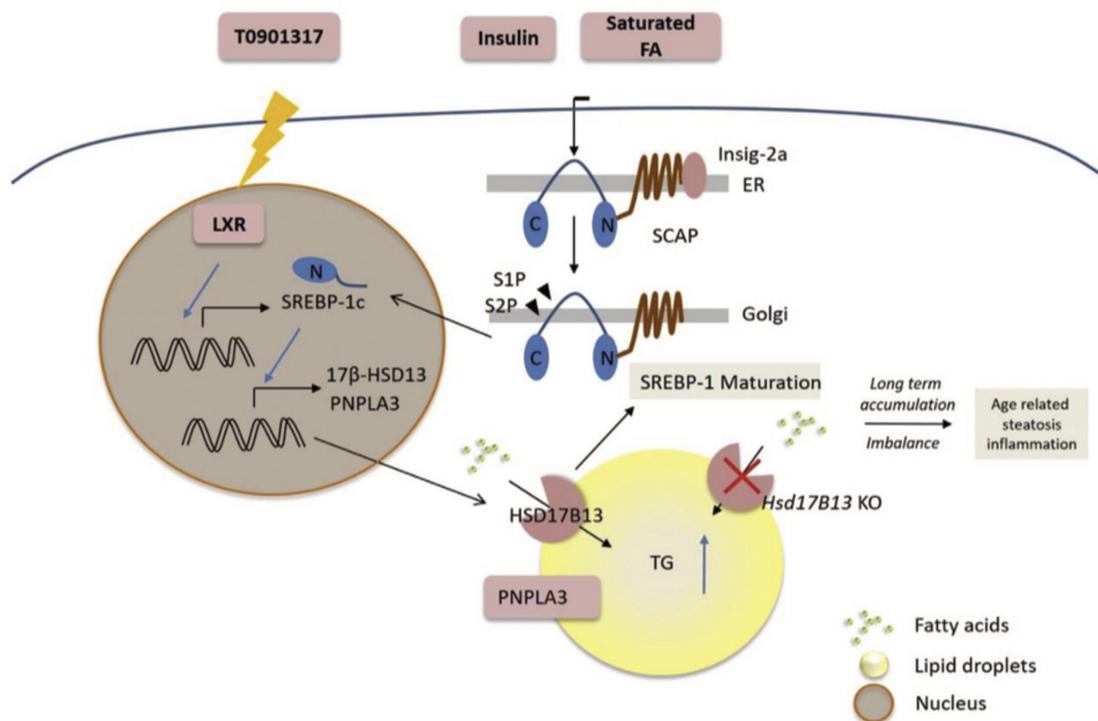


Figure 8 : Proposed mechanisms by which HSD17B13 is involved in hepatic lipid droplet homeostasis.

Lipid homeostasis is controlled by metabolic nuclear receptors, among which the major contributors are liver X receptors (LXRs) and SREBP-1c. HSD17B13 is a direct target gene of SREBP-1c. SREBP-1c induced by insulin, saturated fatty acid or LXR agonists such as T0901317, could increase expression of HSD17B13, which promotes SREBP-1 maturation in feedback loop (Su et al., 2019).

Using a comparative proteomic approach, LDs associated proteins were compared between patients with and without NAFLD. The authors suggested that HSD17B13 played an important role in the pathogenesis of NAFLD. Among the upregulated LD-associated proteins, 17 β -HSD13 represented a previously unidentified LD-associated protein significantly up-regulated in NAFLD. Overexpression of human HSD17B13 increased the number and size of LDs in cultured hepatocytes (Su et al., 2014). In contrast, HSD17B13 KO mice displayed impaired hepatic-lipid metabolism in mice, resulting in age related liver steatosis and inflammation at 9 months, indicating a complex interaction of HSD17B13 in lipogenesis (Adam et al., 2018). The discrepancy between the data observed in mice and in human remains to be better deciphered. There are some hypotheses: 1) HSD17B13 is an enzyme, whose substrates and metabolites are still unknown. The change in HSD17B13 activity might be more important to modulate liver lipid homeostasis. 2) Due to species difference, the mice knockout model could not fully

mimic the biological process in human. Therefore, further studies are required to clarify how expression and activity of HSD17B13 are modulated with fatty liver disease.

Some studies have been performed in order to understand the role of HSD17B13 in liver carcinogenesis. In comparison to non-tumor tissue, expression of *HSD17B13* was downregulated in tumor tissue of HCC (J. Chen et al., 2018; Cotoi, Khorsandi, Plesea, & Quaglia, 2012; X. Wang et al., 2019; Xing et al., 2015; G. Xu et al., 2014). A transcriptome analysis of HCC samples using RNAseq found that overexpression of *HSD17B13* was associated with a favorable prognosis (C. B. Yu et al., 2016), whereas, another study using GSE36376 dataset showed that a high expression of *HSD17B13* was associated with poor prognosis (X. Wang et al., 2019). We have to highlight that none of these two studies had precise which isoform of HSD17B13 has been analyzed. In the following studies, the isoform of *HSD17B13* should be considered. In our study, we investigated the association between genotype of *HSD17B13* rs72613567 and prognosis of HCC treated by liver resection. However, we didn't observe any significant association with prognosis. The discrepancy expression of *HSD17B13* in NAFLD and HCC suggests that it may play distinct roles at different stages of liver disease.

To conclude, increasing data showed a close association between genetic variants of *HSD17B13* and protection against liver disease and HCC development. Functional studies suggested that HSD17B13 could play a role in hepatic lipid metabolism. Enhance expression was observed in liver tissue of patients with NAFLD and mice with NAFLD, with an accumulation of hepatic fat content. However, to date, the mechanisms by which *HSD17B13* rs72613567 protects form CLD and HCC development need to be better understood.

2. Potential clinical applications

2.1 Identification of patients at risk of HCC development

Regarding the vast number of HCC diagnosed at advanced stages and the poor outcomes after the current available treatments, identification of high-risk population of HCC development is a major issue to improve HCC prevention and surveillance. However, due to the great complexity of liver carcinogenesis, the consideration of only traditional factors as predictors of HCC is insufficient for risk stratification. One of the way to refine risk prediction is taking into account individual genetic heterogeneity using SNP. Nowadays, with the popularization of GWAS, more and more HCC-associated genetic variants were identified, shedding light on HCC risk prediction.

Among all the SNPs ever identified, *PNPLA3* and *TM6SF2* SNPs are the most robust variants potentially useful as genetic markers. There are two reasons. First, their association with HCC is the most well investigated and validated. Normally, a variant often performs best in the population from which it was identified, and then its performance could decrease in validation cohort due to differences in genotype frequency, phenotypic effect sizes and diseases incidence (Wray, Goddard, & Visscher, 2008). For *PNPLA3* rs738409 and *TM6SF2* rs58542926, they were validated in numerous studies, including ours, as risk factors ALD, NAFLD and HCC in Western countries. Second, their effect on risk of HCC phenotype is relatively strong. Since the GWAS emphasis on common polymorphisms (minor allele frequency more than 5%), most of the variants associated with complex human diseases have quite a small odds ratio (often less than 1.5). In order to improve predictive accuracy, larger odd ratio is needed (Manolio, 2010). Consistent with other studies, we have showed that *PNPLA3* and *TM6SF2* variants were associated with ALD-related HCC with relatively strong odds ratios (GG=2.5~6.0, CG=1.5~3.0 vs. CC of *PNPLA3* rs738409; TT=2.7~25.9, CT=1.1~2.3 vs. CC of *TM6SF2* rs58542926) after adjustment for other important confounders.

Regarding that both *PNPLA3* and *TM6SF2* have a role in hepatic lipid metabolism and their SNPs share similar strong association in occurrence of ALD-related HCC, we performed a combination analysis in our retrospective study. We observed that patients

with ALD had more HCC if they carry more at-risk alleles of *PNPLA3* and *TM6SF2* SNPs. In our prospective cohort of patients with ALD-related cirrhosis, HCC incidence in patients with more than two at-risk alleles of *PNPLA3* rs738409 and *TM6SF2* rs58542926 are 4.4-fold higher compared with those harboring double wild type variants.

Then we tested the performance of these two genetic variants to predict HCC occurrence. The receive operator characteristic (ROC) curve and the increase in the area under the curve (AUC) are frequently used to test the value of a predictive factor to classify patients between cases and controls (Jakobsdottir, Gorin, Conley, Ferrell, & Weeks, 2009). We were aware that the common SNPs could not be useful alone to predict the occurrence of a common disease in clinical practice. Then two models were tested: clinical factors alone or clinical factors with genetic markers. In our retrospective and prospective cohorts, genetic markers improved only marginally the AUC of clinical factors (**Figure 10**). We need to admit that, although *PNPLA3* and *TM6SF2* SNPs have solid association with HCC risks, they only explain a small part of HCC heritability. According to the theory of Janssens et al, a much larger number of genetic variants than what we have to date are needed to achieve a high AUC (Janssens et al., 2006; Janssens et al., 2007). We could hypothesize that, as more risk variants will be identified, the predictive value of cumulative genotypes scores will increase (Jakobsdottir et al., 2009).

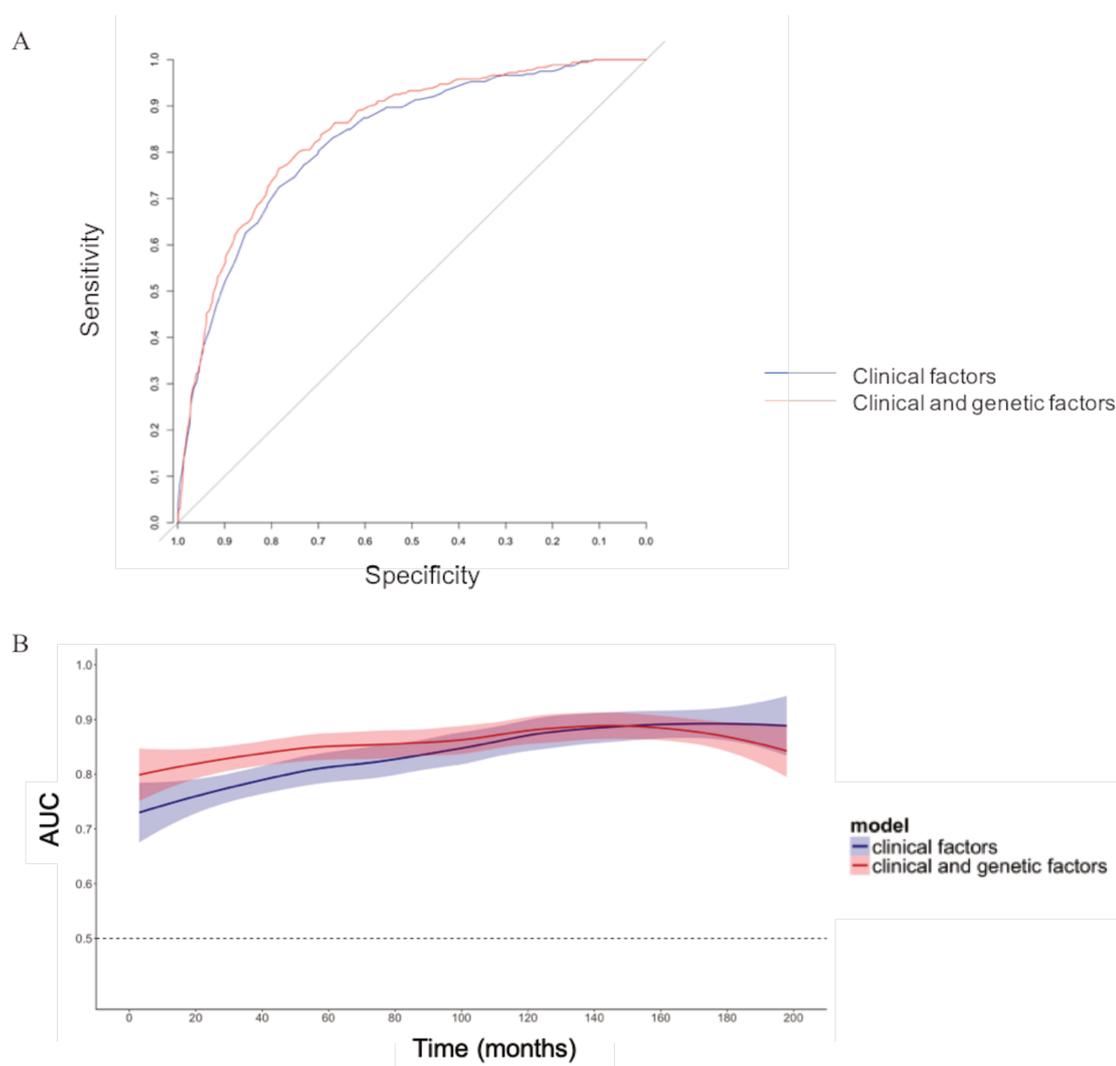


Figure 9 : HCC risk prediction with genetic variants of PNPLA3 and TM6SF2.

A) The AUC of HCC prediction model of clinical factors with and without genetic factors in patients with alcoholic liver disease in our retrospective study. The AUC for clinical factors (age + sex) only was 0.827 [95%CI: 0.803-0.851]; The AUC of clinical factors + genetic factors was 0.844 [95%CI: 0.821-0.867]. $P=0.03$; B) The time dependent AUC for HCC prediction model of clinical factors with and without genetic factors along the follow-up period in prospective cohort of patients with alcohol-related cirrhosis. Age, gender, BMI and diabetes are clinical factors included in prediction model.

Although it seems too early to apply these genetic variants to make an effective prediction of HCC in clinical practice, they may have a potential value for the patient reclassification. A marker with a modest effect on the AUC can sometimes improve risk classification (Cook, 2007). It is especially meaningful for patients who are initially classified to low HCC risk category. For example, patients with chronic liver diseases, such as NAFLD without liver cirrhosis, are outside the target population for HCC surveillance based on

the current guideline. Actually, those with metabolic disorders, together with at-risk genotypes of *PNPLA3* and *TM6SF2* SNPs might be reclassified into higher risk categories to develop HCC. This reclassification is of particular value in clinical decision making. Moreover, SNP could also be useful to stratify treatment. Among baseline (age, sex, BMI, and liver fat) matched patients under the same hypocaloric low-carbohydrate diet, weight loss was not that different but liver fat decreased by 45% in the group with *PNPLA3* 148M ($P < 0.001$) and by 18% in the group with wild type *PNPLA3* ($P < 0.01$) (Sevastianova et al., 2011). Then in clinic, patients with *PNPLA3* 148M could be priority suggested to undergo a special diet to correct and decrease the disease risk.

In addition, another important issue is the complex gene-gene and gene-environment interactions during the development of liver diseases. For example, the variant of *HSD17B13*, rs72613567, was reported to mitigate liver injury induced by the at-risk allele of *PNPLA3* rs738409, serving as a protective variant against liver disease. In our study, we also observed that it could mitigate the risk of ALD-related HCC in patients harboring at-risk allele of *PNPLA3* rs738409. However, more evidences are required to confirm the gene-gene interaction between *PNPLA3* and *HSD17B13* underlying the difficulty to capture all the “neutralized and diluted” associations. Along with the identification of more and more genetic variants with bigger effect, further experiments are required to clarify the physical and functional gene-gene interaction and gene-environment interaction.

Therefore, the future association studies should include much larger and well-phenotyped and prospectively followed-up cohorts to assess time-dependent exposition to risk factors and better assess the gene-environment interaction (Manolio, Bailey-Wilson, & Collins, 2006). Creation of predictive models integrating genetic and environmental/ clinical data will facilitate a better translation of biological information into clinical practice to target surveillance, chemoprevention and treatment in liver cancer (Trepo, Romeo, Zucman-Rossi, & Nahon, 2016).

2.2 Drug development for chemoprevention

Apart from the predictive effect, single nucleotide polymorphisms could be used to identify drugs to perform chemoprevention and identify patients who will respond to these treatments. The genetic background of an individual could have an influence on pharmacodynamics. For example, statins, a common drug to reduce cholesterol synthesis, could protect from the full spectrum of liver damage in patients with NASH. However, the beneficial effect was only seen in patients with wild type *PNPLA3* in one clinical study (Dongiovanni et al., 2015). Carriers of at risk G allele of *PNPLA3* had more risk of liver steatosis and injury. Some studies showed that compared with carriers of wild-type *PNPLA3*, patients with *PNPLA3* 148M will have a greater improvement of hepatic steatosis after a bariatric surgery, or loss of weight or an increase of vegetable consumption (Krawczyk et al., 2016; Nobili et al., 2014; Sevastianova et al., 2011). Pharmacogenomics is going to become increasingly more relevant in the future to tailor the diet, exercise regimen and pharmacological treatment in patients' management (Kovalic, Banerjee, Tran, Singal, & Satapathy, 2018).

In addition, identification of new genetic variants by genome-wide association studies are also of great importance to highlight new mechanisms involved in liver disease and liver carcinogenesis. Then, it brings new insights to identify new targets for drug development. The loss-of-function variant of *PNPLA3* alters the posttranslational modification of *PNPLA3*, protecting the protein from ubiquitylation and inducing dysfunctional *PNPLA3* accumulated on the surface of lipid droplets. It leads to impaired mobilization of triglycerides from the hepatic lipid droplets (BasuRay et al., 2017). Additional SNPs (*PNPLA3* rs2294918 E434K and *HSD17B13* rs72613567) mitigate the detrimental impact of the *PNPLA3* 148M mutant protein (Abul-Husn et al., 2018; Donati et al., 2016). It indicates that reducing expression of the *PNPLA3* 148M mutant protein may exert protective effects on NAFLD progression. In homozygous *PNPLA3* GG knock-in mutant mice, liver-targeted GalNAc3-conjugated antisense oligonucleotide (ASO)-mediated silencing of *PNPLA3* improved NAFLD features, including liver fibrosis, by suppressing the expression of *PNPLA3* 148M. Improvement of NAFLD activity score (NAS) was also seen in wild-type mice fed with a NASH- inducing diet, indicating that such a therapy may also be beneficial in carriers with wild type *PNPLA3* (Linden et al., 2019). This

preclinical study provides a proof of concept evidence for personalized medicine, by targeting a subgroup of patients with NASH (*PNPLA3* 148M allele carriers) using *PNPLA3* ASO therapy.

A splicing variant of *HSD17B13* induces a loss-of-function transcriptional isoform. In contrast with *PNPLA3* rs738409, this *HSD17B13* SNP has a protective effect against ALD, NAFLD, liver cirrhosis and progression to ALD-related HCC. It also suggests that therapeutic inhibition of *HSD17B13* could be a new strategy to prevent progression of CLD due to ALD/NAFLD/CHC and decrease HCC incidence in patients with ALD. Patients harboring the at-risk G allele of *PNPLA3* rs738409 could be the ideal candidate for this chemopreventive strategy. *HSD17Bs* are a group of multifunctional enzymes and a large number of *HSD17Bs* inhibitors have been described in recent years (Marchais-Oberwinkler et al., 2011). However, the exact function of *HSD17B13* remains unclear and our results underline the necessity of functional studies dedicated to understand the potential interactions between the different constitutional variants leading to HCC development. Currently, there is not yet inhibitor for *HSD17B13* to investigate its therapeutic effect. Nevertheless, emerging evidences suggest that *HSD17B13* could represent a new attractive target for drug development in chronic liver disease.

Conclusions and perspectives

Our studies investigated the association between candidate SNPs and the risk of HCC development in patients underlying chronic liver disease. Among them, single nucleotide polymorphisms of *PNPLA3* (rs738409) and *TM6SF2* (rs58542926) were confirmed as strong and independent risk markers of ALD-related HCC along with known HCC risk factors such as age, gender and severity of liver disease. The associations were further validated in a prospective manner in a cohort of patients with alcoholic cirrhosis. Strikingly, *PNPLA3* rs738409 was also associated with increased risk of HCC development on non-fibrotic livers, indicating a potential direct role involved in liver carcinogenesis. In addition, we discovered that a newly identified SNP of *HSD17B13* (rs72613567), that are protective from alcoholic and non-alcoholic liver disease, protects also against the progression from ALD to ALD-related HCC. We also observed that *HSD17B13* rs72613567 could even mitigate the HCC risk in patients with at-risk G allele of *PNPLA3*. Regarding all of our results, it suggests that *PNPLA3* rs738409, *TM6SF2* rs58542926 and *HSD17B13* rs58542926 are three genetic risk markers of ALD-related HCC.

SNPs genotyping could constitute a useful tool to assess the magnitude of the risk of HCC. However, more data are required before the use of these genetic markers in clinical practice to identify patients at high risk of HCC development. A good risk predictive model required more data to precise the genetic heritability of HCC. We advocate for prospective studies with larger patient sets and with well-documented data, using NGS methods. Genetic-based screening strategies may become feasible when host and environmental factors will be taken into account in complex models of “genomic risk prediction”.

Moreover, these current findings underline the heterogeneity of HCC pathogenesis and highlight key mechanisms of liver carcinogenesis. As all these three genes are involved in hepatic lipid metabolism, it suggests that NAFLD and ALD share several pathogenic mechanisms which lead to addictive risk of HCC development. Further experimental studies are warranted to clarify the involvement of lipid metabolism in HCC development. Despite these challenges, the genomic studies will undoubtedly bring us closer to personalized medicine in the field of HCC.

References

References

- Aben, N., Westerhuis, J. A., Song, Y., Kiers, H. A. L., Michaut, M., Smilde, A. K., & Wessels, L. F. A. (2018). iTOP: inferring the topology of omics data. *Bioinformatics*, *34*(17), i988-i996. doi:10.1093/bioinformatics/bty636
- About, F., Abel, L., & Cobat, A. (2018). HCV-Associated Liver Fibrosis and HSD17B13. *N Engl J Med*, *379*(19), 1875-1876. doi:10.1056/NEJMc1804638
- Abul-Husn, N. S., Cheng, X., Li, A. H., Xin, Y., Schurmann, C., Stevis, P., . . . Dewey, F. E. (2018). A Protein-Truncating HSD17B13 Variant and Protection from Chronic Liver Disease. *N Engl J Med*, *378*(12), 1096-1106. doi:10.1056/NEJMoa1712191
- Adam, M., Heikela, H., Sobolewski, C., Portius, D., Maki-Jouppila, J., Mehmood, A., . . . Poutanen, M. (2018). Hydroxysteroid (17beta) dehydrogenase 13 deficiency triggers hepatic steatosis and inflammation in mice. *FASEB J*, *32*(6), 3434-3447. doi:10.1096/fj.201700914R
- Ahn, S. M., Jang, S. J., Shim, J. H., Kim, D., Hong, S. M., Sung, C. O., . . . Kong, G. (2014). Genomic portrait of resectable hepatocellular carcinomas: implications of RB1 and FGF19 aberrations for patient stratification. *Hepatology*, *60*(6), 1972-1982. doi:10.1002/hep.27198
- Al-Qahtani, A., Al-Anazi, M., Viswan, N. A., Khalaf, N., Abdo, A. A., Sanai, F. M., . . . Al-Ahdal, M. (2012). Role of single nucleotide polymorphisms of KIF1B gene in HBV-associated viral hepatitis. *PLoS One*, *7*(9), e45128. doi:10.1371/journal.pone.0045128
- Al-Qahtani, A., Khalak, H. G., Alkuraya, F. S., Al-hamoudi, W., Alswat, K., Al Balwi, M. A., . . . Abdo, A. A. (2013). Genome-wide association study of chronic hepatitis B virus infection reveals a novel candidate risk allele on 11q22.3. *J Med Genet*, *50*(11), 725-732. doi:10.1136/jmedgenet-2013-101724
- Al-Qahtani, A. A., Al-Anazi, M., Abdo, A. A., Sanai, F. M., Al-Hamoudi, W., Alswat, K. A., . . . Al-Ahdal, M. N. (2013). Genetic variation at -1878 (rs2596542) in MICA gene region is associated with chronic hepatitis B virus infection in Saudi Arabian patients. *Exp Mol Pathol*, *95*(3), 255-258. doi:10.1016/j.yexmp.2013.08.005
- Altekruse, S. F., Henley, S. J., Cucinelli, J. E., & McGlynn, K. A. (2014). Changing hepatocellular carcinoma incidence and liver cancer mortality rates in the United States. *Am J Gastroenterol*, *109*(4), 542-553. doi:10.1038/ajg.2014.11
- An, P., Winkler, C., Guan, L., O'Brien, S. J., Zeng, Z., & Consortium, H. B. V. S. (2011). A common HLA-DPA1 variant is a major determinant of hepatitis B virus clearance in Han Chinese. *J Infect Dis*, *203*(7), 943-947. doi:10.1093/infdis/jiq154
- Araya, J., Rodrigo, R., Videla, L. A., Thielemann, L., Orellana, M., Pettinelli, P., & Poniachik, J. (2004).

- Increase in long-chain polyunsaturated fatty acid n - 6/n - 3 ratio in relation to hepatic steatosis in patients with non-alcoholic fatty liver disease. *Clin Sci (Lond)*, 106(6), 635-643. doi:10.1042/CS20030326
- Ascha, M. S., Hanouneh, I. A., Lopez, R., Tamimi, T. A., Feldstein, A. F., & Zein, N. N. (2010). The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *Hepatology*, 51(6), 1972-1978. doi:10.1002/hep.23527
- Atkinson, S. R., Way, M. J., McQuillin, A., Morgan, M. Y., & Thursz, M. R. (2017). Homozygosity for rs738409:G in PNPLA3 is associated with increased mortality following an episode of severe alcoholic hepatitis. *J Hepatol*, 67(1), 120-127. doi:10.1016/j.jhep.2017.01.018
- Augello, G., Balasus, D., Fusilli, C., Mazza, T., Emma, M. R., Giannitrapani, L., . . . Montalto, G. (2018). Association Between MICA Gene Variants and the Risk of Hepatitis C Virus-Induced Hepatocellular Cancer in a Sicilian Population Sample. *OMICS*, 22(4), 274-282. doi:10.1089/omi.2017.0215
- Balasus, D., Way, M., Fusilli, C., Mazza, T., Morgan, M. Y., Cervello, M., . . . Montalto, G. (2016). The association of variants in PNPLA3 and GRP78 and the risk of developing hepatocellular carcinoma in an Italian population. *Oncotarget*, 7(52), 86791-86802. doi:10.18632/oncotarget.13558
- BasuRay, S., Smagris, E., Cohen, J. C., & Hobbs, H. H. (2017). The PNPLA3 variant associated with fatty liver disease (I148M) accumulates on lipid droplets by evading ubiquitylation. *Hepatology*, 66(4), 1111-1124. doi:10.1002/hep.29273
- Bertot, L. C., & Adams, L. A. (2019). Trends in hepatocellular carcinoma due to non-alcoholic fatty liver disease. *Expert Rev Gastroenterol Hepatol*, 13(2), 179-187. doi:10.1080/17474124.2019.1549989
- Bertuccio, P., Turati, F., Carioli, G., Rodriguez, T., La Vecchia, C., Malvezzi, M., & Negri, E. (2017). Global trends and predictions in hepatocellular carcinoma mortality. *J Hepatol*, 67(2), 302-309. doi:10.1016/j.jhep.2017.03.011
- Bezemer, G., Van Gool, A. R., Verheij-Hart, E., Hansen, B. E., Lurie, Y., Esteban, J. I., . . . Group, D.-H. S. (2012). Long-term effects of treatment and response in patients with chronic hepatitis C on quality of life. An international, multicenter, randomized, controlled study. *BMC Gastroenterol*, 12, 11. doi:10.1186/1471-230X-12-11
- Bishayee, A. (2014). The role of inflammation and liver cancer. *Adv Exp Med Biol*, 816, 401-435. doi:10.1007/978-3-0348-0837-8_16
- Boyault, S., Rickman, D. S., de Reynies, A., Balabaud, C., Rebouissou, S., Jeannot, E., . . . Zucman-Rossi, J. (2007). Transcriptome classification of HCC is related to gene alterations and to new therapeutic

- targets. *Hepatology*, 45(1), 42-52. doi:10.1002/hep.21467
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 68(6), 394-424. doi:10.3322/caac.21492
- Brouwer, W. P., van der Meer, A. J., Boonstra, A., Pas, S. D., de Knegt, R. J., de Man, R. A., . . . Janssen, H. L. (2015). The impact of PNPLA3 (rs738409 C>G) polymorphisms on liver histology and long-term clinical outcome in chronic hepatitis B patients. *Liver Int*, 35(2), 438-447. doi:10.1111/liv.12695
- Bruix, J., Gores, G. J., & Mazzaferro, V. (2014). Hepatocellular carcinoma: clinical frontiers and perspectives. *Gut*, 63(5), 844-855. doi:10.1136/gutjnl-2013-306627
- Bruix, J., Qin, S., Merle, P., Granito, A., Huang, Y. H., Bodoky, G., . . . Investigators, R. (2017). Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*, 389(10064), 56-66. doi:10.1016/S0140-6736(16)32453-9
- Bruschi, F. V., Claudel, T., Tardelli, M., Caligiuri, A., Stulnig, T. M., Marra, F., & Trauner, M. (2017). The PNPLA3 I148M variant modulates the fibrogenic phenotype of human hepatic stellate cells. *Hepatology*, 65(6), 1875-1890. doi:10.1002/hep.29041
- Bruschi, F. V., Tardelli, M., Claudel, T., & Trauner, M. (2017). PNPLA3 expression and its impact on the liver: current perspectives. *Hepat Med*, 9, 55-66. doi:10.2147/HMER.S125718
- Buch, S., Stickel, F., Trepo, E., Way, M., Herrmann, A., Nischalke, H. D., . . . Hampe, J. (2015). A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. *Nat Genet*, 47(12), 1443-1448. doi:10.1038/ng.3417
- Burza, M. A., Molinaro, A., Attilia, M. L., Rotondo, C., Attilia, F., Ceccanti, M., . . . Ginanni Corradini, S. (2014). PNPLA3 I148M (rs738409) genetic variant and age at onset of at-risk alcohol consumption are independent risk factors for alcoholic cirrhosis. *Liver Int*, 34(4), 514-520. doi:10.1111/liv.12310
- Burza, M. A., Motta, B. M., Mancina, R. M., Pingitore, P., Pirazzi, C., Lepore, S. M., . . . Romeo, S. (2016). DEPDC5 variants increase fibrosis progression in Europeans with chronic hepatitis C virus infection. *Hepatology*, 63(2), 418-427. doi:10.1002/hep.28322
- Chalasani, N., Guo, X., Loomba, R., Goodarzi, M. O., Haritunians, T., Kwon, S., . . . Nonalcoholic Steatohepatitis Clinical Research, N. (2010). Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease. *Gastroenterology*, 139(5), 1567-1576, 1576 e1561-1566. doi:10.1053/j.gastro.2010.07.057

References

- Chambers, J. C., Zhang, W., Sehmi, J., Li, X., Wass, M. N., Van der Harst, P., . . . Kooner, J. S. (2011). Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *Nat Genet*, *43*(11), 1131-1138. doi:10.1038/ng.970
- Chamoun, Z., Vacca, F., Parton, R. G., & Gruenberg, J. (2013). PNPLA3/adiponutrin functions in lipid droplet formation. *Biol Cell*, *105*(5), 219-233. doi:10.1111/boc.201200036
- Chang, M. H., Chen, C. J., Lai, M. S., Hsu, H. M., Wu, T. C., Kong, M. S., . . . Chen, D. S. (1997). Universal hepatitis B vaccination in Taiwan and the incidence of hepatocellular carcinoma in children. Taiwan Childhood Hepatoma Study Group. *N Engl J Med*, *336*(26), 1855-1859. doi:10.1056/NEJM199706263362602
- Chayanupatkul, M., Omino, R., Mittal, S., Kramer, J. R., Richardson, P., Thrift, A. P., . . . Kanwal, F. (2017). Hepatocellular carcinoma in the absence of cirrhosis in patients with chronic hepatitis B virus infection. *J Hepatol*, *66*(2), 355-362. doi:10.1016/j.jhep.2016.09.013
- Chen, C. J., Yang, H. I., Iloeje, U. H., & Group, R.-H. S. (2009). Hepatitis B virus DNA levels and outcomes in chronic hepatitis B. *Hepatology*, *49*(5 Suppl), S72-84. doi:10.1002/hep.22884
- Chen, C. J., Yang, H. I., Su, J., Jen, C. L., You, S. L., Lu, S. N., . . . Group, R.-H. S. (2006). Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. *JAMA*, *295*(1), 65-73. doi:10.1001/jama.295.1.65
- Chen, J., Zhuo, J. Y., Yang, F., Liu, Z. K., Zhou, L., Xie, H. Y., . . . Zheng, S. S. (2018). 17-beta-hydroxysteroid dehydrogenase 13 inhibits the progression and recurrence of hepatocellular carcinoma. *Hepatobiliary Pancreat Dis Int*, *17*(3), 220-226. doi:10.1016/j.hbpd.2018.04.006
- Chen, J. D., Yang, H. I., Iloeje, U. H., You, S. L., Lu, S. N., Wang, L. Y., . . . Associated Liver Disease/Cancer in, H. B. V. S. G. (2010). Carriers of inactive hepatitis B virus are still at risk for hepatocellular carcinoma and liver-related death. *Gastroenterology*, *138*(5), 1747-1754. doi:10.1053/j.gastro.2010.01.042
- Chen, J. G., & Kensler, T. W. (2014). Changing rates for liver and lung cancers in Qidong, China. *Chem Res Toxicol*, *27*(1), 3-6. doi:10.1021/tx400313j
- Chen, K., Shi, W., Xin, Z., Wang, H., Zhu, X., Wu, X., . . . Liu, Y. (2013). Replication of genome wide association studies on hepatocellular carcinoma susceptibility loci in a Chinese population. *PLoS One*, *8*(10), e77315. doi:10.1371/journal.pone.0077315
- Chen, W., Chang, B., Li, L., & Chan, L. (2010). Patatin-like phospholipase domain-containing 3/adiponutrin deficiency in mice is not associated with fatty liver disease. *Hepatology*, *52*(3), 1134-1142. doi:10.1002/hep.23812
- Chen, W., Wang, M., Zhang, Z., Tang, H., Zuo, X., Meng, X., . . . Sun, L. (2015). Replication the association

- of 2q32.2-q32.3 and 14q32.11 with hepatocellular carcinoma. *Gene*, 561(1), 63-67. doi:10.1016/j.gene.2015.02.006
- Chiang, D. Y., Villanueva, A., Hoshida, Y., Peix, J., Newell, P., Minguez, B., . . . Llovet, J. M. (2008). Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. *Cancer Res*, 68(16), 6779-6788. doi:10.1158/0008-5472.CAN-08-0742
- Chisari, F. V., & Ferrari, C. (1995). Hepatitis B virus immunopathogenesis. *Annu Rev Immunol*, 13, 29-60. doi:10.1146/annurev.iy.13.040195.000333
- Clark, A., Gerlach, F., Tong, H., Hoan, N. X., Song le, H., Toan, N. L., . . . Velavan, T. P. (2013). A trivial role of STAT4 variant in chronic hepatitis B induced hepatocellular carcinoma. *Infect Genet Evol*, 18, 257-261. doi:10.1016/j.meegid.2013.05.025
- Colombo, M., de Franchis, R., Del Ninno, E., Sangiovanni, A., De Fazio, C., Tommasini, M., . . . Dioguardi, N. (1991). Hepatocellular carcinoma in Italian patients with cirrhosis. *N Engl J Med*, 325(10), 675-680. doi:10.1056/NEJM199109053251002
- Cook, N. R. (2007). Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*, 115(7), 928-935. doi:10.1161/CIRCULATIONAHA.106.672402
- Coppola, N., Rosa, Z., Cirillo, G., Stanzione, M., Macera, M., Boemio, A., . . . Miraglia del Giudice, E. (2015). TM6SF2 E167K variant is associated with severe steatosis in chronic hepatitis C, regardless of PNPLA3 polymorphism. *Liver Int*, 35(8), 1959-1963. doi:10.1111/liv.12781
- Cotoi, C. G., Khorsandi, S. E., Plesea, I. E., & Quaglia, A. (2012). Whole-genome DASL gene expression profiling of hepatocellular carcinoma sub-populations isolated by laser microdissection on formalin-fixed and paraffin-embedded liver tissue samples. *Rom J Morphol Embryol*, 53(4), 893-902.
- Coulouarn, C., Factor, V. M., & Thorgeirsson, S. S. (2008). Transforming growth factor-beta gene expression signature in mouse hepatocytes predicts clinical outcome in human cancer. *Hepatology*, 47(6), 2059-2067. doi:10.1002/hep.22283
- Cucchetti, A., Piscaglia, F., Cescon, M., Colecchia, A., Ercolani, G., Bolondi, L., & Pinna, A. D. (2013). Cost-effectiveness of hepatic resection versus percutaneous radiofrequency ablation for early hepatocellular carcinoma. *J Hepatol*, 59(2), 300-307. doi:10.1016/j.jhep.2013.04.009
- de La Coste, A., Romagnolo, B., Billuart, P., Renard, C. A., Buendia, M. A., Soubrane, O., . . . Perret, C. (1998). Somatic mutations of the beta-catenin gene are frequent in mouse and human hepatocellular carcinomas. *Proc Natl Acad Sci U S A*, 95(15), 8847-8851. doi:10.1073/pnas.95.15.8847
- Degasperi, E., & Colombo, M. (2016). Distinctive features of hepatocellular carcinoma in non-alcoholic

- fatty liver disease. *Lancet Gastroenterol Hepatol*, 1(2), 156-164. doi:10.1016/S2468-1253(16)30018-8
- Deutsch, M., Papadopoulos, N., Hadziyannis, E. S., & Koskinas, J. (2013). Clinical characteristics, spontaneous clearance and treatment outcome of acute hepatitis C: a single tertiary center experience. *Saudi J Gastroenterol*, 19(2), 81-85. doi:10.4103/1319-3767.108479
- Di Tommaso, L., Franchi, G., Park, Y. N., Fiamengo, B., Destro, A., Morengi, E., . . . Roncalli, M. (2007). Diagnostic value of HSP70, glypican 3, and glutamine synthetase in hepatocellular nodules in cirrhosis. *Hepatology*, 45(3), 725-734. doi:10.1002/hep.21531
- Donati, B., Dongiovanni, P., Romeo, S., Meroni, M., McCain, M., Miele, L., . . . Valenti, L. (2017). MBOAT7 rs641738 variant and hepatocellular carcinoma in non-cirrhotic individuals. *Sci Rep*, 7(1), 4492. doi:10.1038/s41598-017-04991-0
- Donati, B., Motta, B. M., Pingitore, P., Meroni, M., Pietrelli, A., Alisi, A., . . . Valenti, L. (2016). The rs2294918 E434K variant modulates patatin-like phospholipase domain-containing 3 expression and liver damage. *Hepatology*, 63(3), 787-798. doi:10.1002/hep.28370
- Donato, F., Tagger, A., Gelatti, U., Parrinello, G., Boffetta, P., Albertini, A., . . . Nardi, G. (2002). Alcohol and hepatocellular carcinoma: the effect of lifetime intake and hepatitis virus infections in men and women. *Am J Epidemiol*, 155(4), 323-331. doi:10.1093/aje/155.4.323
- Dongiovanni, P., Petta, S., Mannisto, V., Mancina, R. M., Pipitone, R., Karja, V., . . . Valenti, L. (2015). Statin use and non-alcoholic steatohepatitis in at risk individuals. *J Hepatol*, 63(3), 705-712. doi:10.1016/j.jhep.2015.05.006
- Dorfman, J. D., Schulick, R., Choti, M. A., Geschwind, J. F., Kamel, I., Torbenson, M., & Thuluvath, P. J. (2007). Differences in characteristics of patients with and without known risk factors for hepatocellular carcinoma in the United States. *World J Gastroenterol*, 13(5), 781-784. doi:10.3748/wjg.v13.i5.781
- Dubuquoy, C., Burnol, A. F., & Moldes, M. (2013). PNPLA3, a genetic marker of progressive liver disease, still hiding its metabolic function? *Clin Res Hepatol Gastroenterol*, 37(1), 30-35. doi:10.1016/j.clinre.2012.06.014
- Dubuquoy, C., Robichon, C., Lasnier, F., Langlois, C., Dugail, I., Fougelle, F., . . . Moldes, M. (2011). Distinct regulation of adiponutrin/PNPLA3 gene expression by the transcription factors ChREBP and SREBP1c in mouse and human hepatocytes. *J Hepatol*, 55(1), 145-153. doi:10.1016/j.jhep.2010.10.024
- Dutta, A. K. (2013). Genetic factors affecting susceptibility to alcoholic liver disease in an Indian population. *Ann Hepatol*, 12(6), 901-907.

- El-Serag, H. B. (2012). Epidemiology of viral hepatitis and hepatocellular carcinoma. *Gastroenterology*, *142*(6), 1264-1273 e1261. doi:10.1053/j.gastro.2011.12.061
- El-Serag, H. B., Hampel, H., & Javadi, F. (2006). The association between diabetes and hepatocellular carcinoma: a systematic review of epidemiologic evidence. *Clin Gastroenterol Hepatol*, *4*(3), 369-380. doi:10.1016/j.cgh.2005.12.007
- El-Serag, H. B., & Mason, A. C. (2000). Risk factors for the rising rates of primary liver cancer in the United States. *Arch Intern Med*, *160*(21), 3227-3230.
- European Association For The Study Of The, L., European Organisation For, R., & Treatment Of, C. (2012). EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol*, *56*(4), 908-943. doi:10.1016/j.jhep.2011.12.001
- European Association for the Study of the Liver. Electronic address, e. e. e., & European Association for the Study of the, L. (2018). EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol*, *69*(1), 182-236. doi:10.1016/j.jhep.2018.03.019
- Ezzikouri, S., Alaoui, R., Tazi, S., Nadir, S., Elmdaghri, N., Pineau, P., & Benjelloun, S. (2014). The adiponutrin I148M variant is a risk factor for HCV-associated liver cancer in North-African patients. *Infect Genet Evol*, *21*, 179-183. doi:10.1016/j.meegid.2013.11.005
- Falleti, E., Cussigh, A., Cmet, S., Fabris, C., & Toniutto, P. (2016). PNPLA3 rs738409 and TM6SF2 rs58542926 variants increase the risk of hepatocellular carcinoma in alcoholic cirrhosis. *Dig Liver Dis*, *48*(1), 69-75. doi:10.1016/j.dld.2015.09.009
- Falleti, E., Fabris, C., Cmet, S., Cussigh, A., Bitetto, D., Fontanini, E., . . . Toniutto, P. (2011). PNPLA3 rs738409C/G polymorphism in cirrhosis: relationship with the aetiology of liver disease and hepatocellular carcinoma occurrence. *Liver Int*, *31*(8), 1137-1143. doi:10.1111/j.1478-3231.2011.02534.x
- Falleti, E., Fabris, C., Vandelli, C., Colletta, C., Cussigh, A., Smirne, C., . . . Pirisi, M. (2010). Genetic polymorphisms of interleukin-6 modulate fibrosis progression in mild chronic hepatitis C. *Hum Immunol*, *71*(10), 999-1004. doi:10.1016/j.humimm.2010.06.006
- Fan, J. H., Xiang, M. Q., Li, Q. L., Shi, H. T., & Guo, J. J. (2016). PNPLA3 rs738409 Polymorphism Associated with Hepatic Steatosis and Advanced Fibrosis in Patients with Chronic Hepatitis C Virus: A Meta-Analysis. *Gut Liver*, *10*(3), 456-463. doi:10.5009/gnl15261
- Fan, Y., Lu, H., Guo, Y., Zhu, T., Garcia-Barrio, M. T., Jiang, Z., . . . Chen, Y. E. (2016). Hepatic Transmembrane 6 Superfamily Member 2 Regulates Cholesterol Metabolism in Mice. *Gastroenterology*, *150*(5), 1208-1218. doi:10.1053/j.gastro.2016.01.005
- Fargion, S., Fracanzani, A. L., Piperno, A., Braga, M., D'Alba, R., Ronchi, G., & Fiorelli, G. (1994).

- Prognostic factors for hepatocellular carcinoma in genetic hemochromatosis. *Hepatology*, 20(6), 1426-1431.
- Fattovich, G., Stroffolini, T., Zagni, I., & Donato, F. (2004). Hepatocellular carcinoma in cirrhosis: incidence and risk factors. *Gastroenterology*, 127(5 Suppl 1), S35-50.
- Fornier, A., Vilana, R., Ayuso, C., Bianchi, L., Sole, M., Ayuso, J. R., . . . Bruix, J. (2008). Diagnosis of hepatic nodules 20 mm or smaller in cirrhosis: Prospective validation of the noninvasive diagnostic criteria for hepatocellular carcinoma. *Hepatology*, 47(1), 97-104. doi:10.1002/hep.21966
- Friedrich, K., Wannhoff, A., Kattner, S., Brune, M., Hov, J. R., Weiss, K. H., . . . Gotthardt, D. N. (2014). PNPLA3 in end-stage liver disease: alcohol consumption, hepatocellular carcinoma development, and transplantation-free survival. *J Gastroenterol Hepatol*, 29(7), 1477-1484. doi:10.1111/jgh.12540
- Ganem, D., & Prince, A. M. (2004). Hepatitis B virus infection--natural history and clinical consequences. *N Engl J Med*, 350(11), 1118-1129. doi:10.1056/NEJMra031087
- Ge, D., Fellay, J., Thompson, A. J., Simon, J. S., Shianna, K. V., Urban, T. J., . . . Goldstein, D. B. (2009). Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature*, 461(7262), 399-401. doi:10.1038/nature08309
- Germani, G., Pleguezuelo, M., Gurusamy, K., Meyer, T., Isgro, G., & Burroughs, A. K. (2010). Clinical outcomes of radiofrequency ablation, percutaneous alcohol and acetic acid injection for hepatocellular carcinoma: a meta-analysis. *J Hepatol*, 52(3), 380-388. doi:10.1016/j.jhep.2009.12.004
- Goffredo, M., Caprio, S., Feldstein, A. E., D'Adamo, E., Shaw, M. M., Pierpont, B., . . . Santoro, N. (2016). Role of TM6SF2 rs58542926 in the pathogenesis of nonalcoholic pediatric fatty liver disease: A multiethnic study. *Hepatology*, 63(1), 117-125. doi:10.1002/hep.28283
- Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I. B., . . . Zucman-Rossi, J. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet*, 44(6), 694-698. doi:10.1038/ng.2256
- Gunes, C., & Rudolph, K. L. (2013). The role of telomeres in stem cells and cancer. *Cell*, 152(3), 390-393. doi:10.1016/j.cell.2013.01.010
- Guyot, E., Sutton, A., Rufat, P., Laguillier, C., Mansouri, A., Moreau, R., . . . Nahon, P. (2013). PNPLA3 rs738409, hepatocellular carcinoma occurrence and risk model prediction in patients with cirrhosis. *J Hepatol*, 58(2), 312-318. doi:10.1016/j.jhep.2012.09.036
- Guzman, G., Brunt, E. M., Petrovic, L. M., Chejfec, G., Layden, T. J., & Cotler, S. J. (2008). Does

- nonalcoholic fatty liver disease predispose patients to hepatocellular carcinoma in the absence of cirrhosis? *Arch Pathol Lab Med*, 132(11), 1761-1766. doi:10.1043/1543-2165-132.11.1761
- Gwack, J., Park, S. K., Lee, E. H., Park, B., Choi, Y., & Yoo, K. Y. (2011). Hepatitis B vaccination and liver cancer mortality reduction in Korean children and adolescents. *Asian Pac J Cancer Prev*, 12(9), 2205-2208.
- Hai, H., Tamori, A., Thuy, L. T. T., Yoshida, K., Hagihara, A., Kawamura, E., . . . Kawada, N. (2017). Polymorphisms in MICA, but not in DEPDC5, HCP5 or PNPLA3, are associated with chronic hepatitis C-related hepatocellular carcinoma. *Sci Rep*, 7(1), 11912. doi:10.1038/s41598-017-10363-5
- Han, H. S., Shehta, A., Ahn, S., Yoon, Y. S., Cho, J. Y., & Choi, Y. (2015). Laparoscopic versus open liver resection for hepatocellular carcinoma: Case-matched study with propensity score matching. *J Hepatol*, 63(3), 643-650. doi:10.1016/j.jhep.2015.04.005
- Hasegawa, K., Kokudo, N., Makuuchi, M., Izumi, N., Ichida, T., Kudo, M., . . . Matsuyama, Y. (2013). Comparison of resection and ablation for hepatocellular carcinoma: a cohort study based on a Japanese nationwide survey. *J Hepatol*, 58(4), 724-729. doi:10.1016/j.jhep.2012.11.009
- Hassan, M. M., Spitz, M. R., Thomas, M. B., Curley, S. A., Patt, Y. Z., Vauthey, J. N., . . . Li, D. (2009). The association of family history of liver cancer with hepatocellular carcinoma: a case-control study in the United States. *J Hepatol*, 50(2), 334-341. doi:10.1016/j.jhep.2008.08.016
- He, S., McPhaul, C., Li, J. Z., Garuti, R., Kinch, L., Grishin, N. V., . . . Hobbs, H. H. (2010). A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. *J Biol Chem*, 285(9), 6706-6715. doi:10.1074/jbc.M109.064501
- Hellier, S., Frodsham, A. J., Hennig, B. J., Klenerman, P., Knapp, S., Ramaley, P., . . . Hill, A. V. (2003). Association of genetic variants of the chemokine receptor CCR5 and its ligands, RANTES and MCP-2, with outcome of HCV infection. *Hepatology*, 38(6), 1468-1476. doi:10.1016/j.hep.2003.09.027
- Herrero, J. I., Sangro, B., Pardo, F., Quiroga, J., Inarrairaegui, M., Rotellar, F., . . . Prieto, J. (2008). Liver transplantation in patients with hepatocellular carcinoma across Milan criteria. *Liver Transpl*, 14(3), 272-278. doi:10.1002/lt.21368
- Horiguchi, Y., Araki, M., & Motojima, K. (2008). 17beta-Hydroxysteroid dehydrogenase type 13 is a liver-specific lipid droplet-associated protein. *Biochem Biophys Res Commun*, 370(2), 235-238. doi:10.1016/j.bbrc.2008.03.063
- Horras, C. J., Lamb, C. L., & Mitchell, K. A. (2011). Regulation of hepatocyte fate by interferon-gamma. *Cytokine Growth Factor Rev*, 22(1), 35-43. doi:10.1016/j.cytogfr.2011.01.001

- Hoshida, Y., Nijman, S. M., Kobayashi, M., Chan, J. A., Brunet, J. P., Chiang, D. Y., . . . Golub, T. R. (2009). Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res*, *69*(18), 7385-7392. doi:10.1158/0008-5472.CAN-09-1089
- Hsu, I. C., Metcalf, R. A., Sun, T., Welsh, J. A., Wang, N. J., & Harris, C. C. (1991). Mutational hotspot in the p53 gene in human hepatocellular carcinomas. *Nature*, *350*(6317), 427-428. doi:10.1038/350427a0
- Hu, Z., Liu, Y., Zhai, X., Dai, J., Jin, G., Wang, L., . . . Shen, H. (2013). New loci associated with chronic hepatitis B virus infection in Han Chinese. *Nat Genet*, *45*(12), 1499-1503. doi:10.1038/ng.2809
- Huang, M., Pan, Y., Liu, J., Qi, F., Wen, J., Xie, K., . . . Dai, J. (2014). A genetic variant at KIF1B predicts clinical outcome of HBV-related hepatocellular carcinoma in Chinese. *Cancer Epidemiol*, *38*(5), 608-612. doi:10.1016/j.canep.2014.07.012
- Huang, Y., He, S., Li, J. Z., Seo, Y. K., Osborne, T. F., Cohen, J. C., & Hobbs, H. H. (2010). A feed-forward loop amplifies nutritional regulation of PNPLA3. *Proc Natl Acad Sci U S A*, *107*(17), 7892-7897. doi:10.1073/pnas.1003585107
- Iloeje, U. H., Yang, H. I., Su, J., Jen, C. L., You, S. L., Chen, C. J., . . . Associated Liver Disease/Cancer-In, H. B. V. S. G. (2006). Predicting cirrhosis risk based on the level of circulating hepatitis B viral load. *Gastroenterology*, *130*(3), 678-686. doi:10.1053/j.gastro.2005.11.016
- Ishizawa, T., Hasegawa, K., Aoki, T., Takahashi, M., Inoue, Y., Sano, K., . . . Makuuchi, M. (2008). Neither multiple tumors nor portal hypertension are surgical contraindications for hepatocellular carcinoma. *Gastroenterology*, *134*(7), 1908-1916. doi:10.1053/j.gastro.2008.02.091
- Jakobsdottir, J., Gorin, M. B., Conley, Y. P., Ferrell, R. E., & Weeks, D. E. (2009). Interpretation of genetic association studies: markers with replicated highly significant odds ratios may be poor classifiers. *PLoS Genet*, *5*(2), e1000337. doi:10.1371/journal.pgen.1000337
- Janssens, A. C., Aulchenko, Y. S., Elefante, S., Borsboom, G. J., Steyerberg, E. W., & van Duijn, C. M. (2006). Predictive testing for complex diseases using multiple genes: fact or fiction? *Genet Med*, *8*(7), 395-400. doi:10.109701.gim.0000229689.18263.f4
- Janssens, A. C., Moonesinghe, R., Yang, Q., Steyerberg, E. W., van Duijn, C. M., & Khoury, M. J. (2007). The impact of genotype frequencies on the clinical validity of genomic profiling for predicting common chronic diseases. *Genet Med*, *9*(8), 528-535. doi:10.1097GIM.0b013e31812eece0
- Jiang, D. K., Sun, J., Cao, G., Liu, Y., Lin, D., Gao, Y. Z., . . . Yu, L. (2013). Genetic variants in STAT4 and HLA-DQ genes confer risk of hepatitis B virus-related hepatocellular carcinoma. *Nat Genet*, *45*(1), 72-75. doi:10.1038/ng.2483
- Jimenez-Sousa, M. A., Berenguer, J., Garcia-Alvarez, M., Gutierrez-Rivas, M., Aldamiz-Echevarria, T.,

- Tejerina, F., . . . Resino, S. (2016). Impact of patatin-like phospholipase domain-containing 3 gene polymorphism (rs738409) on severity of liver disease in HIV/hepatitis C virus-coinfected patients. *AIDS*, *30*(3), 465-470. doi:10.1097/QAD.0000000000000908
- Jutavijittum, P., Yousukh, A., Saysanasongkham, B., Samounry, B., Samounry, K., Toriyama, K., . . . Masaki, T. (2016). High Rate of Hepatitis B Virus Mother-to-Child Transmission in Lao People's Democratic Republic. *Southeast Asian J Trop Med Public Health*, *47*(2), 214-218.
- Kamatani, Y., Wattanapokayakit, S., Ochi, H., Kawaguchi, T., Takahashi, A., Hosono, N., . . . Matsuda, K. (2009). A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. *Nat Genet*, *41*(5), 591-595. doi:10.1038/ng.348
- Kao, J. H., Chen, P. J., Lai, M. Y., & Chen, D. S. (2002). Genotypes and clinical phenotypes of hepatitis B virus in patients with chronic hepatitis B virus infection. *J Clin Microbiol*, *40*(4), 1207-1209. doi:10.1128/jcm.40.4.1207-1209.2002
- Kaposi-Novak, P., Lee, J. S., Gomez-Quiroz, L., Coulouarn, C., Factor, V. M., & Thorgeirsson, S. S. (2006). Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype. *J Clin Invest*, *116*(6), 1582-1595. doi:10.1172/JCI27236
- Kawamura, Y., Arase, Y., Ikeda, K., Seko, Y., Imai, N., Hosaka, T., . . . Kumada, H. (2012). Large-scale long-term follow-up study of Japanese patients with non-alcoholic Fatty liver disease for the onset of hepatocellular carcinoma. *Am J Gastroenterol*, *107*(2), 253-261. doi:10.1038/ajg.2011.327
- Kim, L. H., Cheong, H. S., Namgoong, S., Kim, J. O., Kim, J. H., Park, B. L., . . . Kim, Y. J. (2015). Replication of genome wide association studies on hepatocellular carcinoma susceptibility loci of STAT4 and HLA-DQ in a Korean population. *Infect Genet Evol*, *33*, 72-76. doi:10.1016/j.meegid.2015.04.013
- Kim, Y. J., Kim, H. Y., Lee, J. H., Yu, S. J., Yoon, J. H., Lee, H. S., . . . Shin, H. D. (2013). A genome-wide association study identified new variants associated with the risk of chronic hepatitis B. *Hum Mol Genet*, *22*(20), 4233-4238. doi:10.1093/hmg/ddt266
- Kitamoto, T., Kitamoto, A., Yoneda, M., Hyogo, H., Ochi, H., Nakamura, T., . . . Hotta, K. (2013). Genome-wide scan revealed that polymorphisms in the PNPLA3, SAMM50, and PARVB genes are associated with development and progression of nonalcoholic fatty liver disease in Japan. *Hum Genet*, *132*(7), 783-792. doi:10.1007/s00439-013-1294-3
- Kolla, B. P., Schneekloth, T. D., Biernacka, J., Shah, V., Lazaridis, K. N., Geske, J., & Karpyak, V. (2018). PNPLA3 Association with Alcoholic Liver Disease in a Cohort of Heavy Drinkers. *Alcohol*, *53*(4), 357-360. doi:10.1093/alcalc/agy007
- Kotronen, A., Johansson, L. E., Johansson, L. M., Roos, C., Westerbacka, J., Hamsten, A., . . . Yki-Jarvinen,

- H. (2009). A common variant in PNPLA3, which encodes adiponutrin, is associated with liver fat content in humans. *Diabetologia*, *52*(6), 1056-1060. doi:10.1007/s00125-009-1285-z
- Kovalic, A. J., Banerjee, P., Tran, Q. T., Singal, A. K., & Satapathy, S. K. (2018). Genetic and Epigenetic Culprits in the Pathogenesis of Nonalcoholic Fatty Liver Disease. *J Clin Exp Hepatol*, *8*(4), 390-402. doi:10.1016/j.jceh.2018.04.001
- Kozlitina, J., Smagris, E., Stender, S., Nordestgaard, B. G., Zhou, H. H., Tybjaerg-Hansen, A., . . . Cohen, J. C. (2014). Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet*, *46*(4), 352-356. doi:10.1038/ng.2901
- Kozlitina, J., Stender, S., Hobbs, H. H., & Cohen, J. C. (2018). HSD17B13 and Chronic Liver Disease in Blacks and Hispanics. *N Engl J Med*, *379*(19), 1876-1877. doi:10.1056/NEJMc1804027
- Krawczyk, M., Jimenez-Aguero, R., Alustiza, J. M., Emparanza, J. I., Perugorria, M. J., Bujanda, L., . . . Banales, J. M. (2016). PNPLA3 p.I148M variant is associated with greater reduction of liver fat content after bariatric surgery. *Surg Obes Relat Dis*, *12*(10), 1838-1846. doi:10.1016/j.soard.2016.06.004
- Kudo, M., Finn, R. S., Qin, S., Han, K. H., Ikeda, K., Piscaglia, F., . . . Cheng, A. L. (2018). Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet*, *391*(10126), 1163-1173. doi:10.1016/S0140-6736(18)30207-1
- Kumar, V., Kato, N., Urabe, Y., Takahashi, A., Muroyama, R., Hosono, N., . . . Matsuda, K. (2011). Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma. *Nat Genet*, *43*(5), 455-458. doi:10.1038/ng.809
- Kumar, V., Yi Lo, P. H., Sawai, H., Kato, N., Takahashi, A., Deng, Z., . . . Matsuda, K. (2012). Soluble MICA and a MICA variation as possible prognostic biomarkers for HBV-induced hepatocellular carcinoma. *PLoS One*, *7*(9), e44743. doi:10.1371/journal.pone.0044743
- Kushner, T., Serper, M., & Kaplan, D. E. (2015). Delta hepatitis within the Veterans Affairs medical system in the United States: Prevalence, risk factors, and outcomes. *J Hepatol*, *63*(3), 586-592. doi:10.1016/j.jhep.2015.04.025
- Lake, A. C., Sun, Y., Li, J. L., Kim, J. E., Johnson, J. W., Li, D., . . . Gimeno, R. E. (2005). Expression, regulation, and triglyceride hydrolase activity of Adiponutrin family members. *J Lipid Res*, *46*(11), 2477-2487. doi:10.1194/jlr.M500290-JLR200
- Lange, C. M., Bibert, S., Dufour, J. F., Celleraï, C., Cerny, A., Heim, M. H., . . . Swiss Hepatitis, C. C. S. G. (2013). Comparative genetic analyses point to HCP5 as susceptibility locus for HCV-associated hepatocellular carcinoma. *J Hepatol*, *59*(3), 504-509. doi:10.1016/j.jhep.2013.04.032

- Larsson, S. C., & Wolk, A. (2007). Overweight, obesity and risk of liver cancer: a meta-analysis of cohort studies. *Br J Cancer*, *97*(7), 1005-1008. doi:10.1038/sj.bjc.6603932
- Lee, J. S., Heo, J., Libbrecht, L., Chu, I. S., Kaposi-Novak, P., Calvisi, D. F., . . . Thorgeirsson, S. S. (2006). A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat Med*, *12*(4), 410-416. doi:10.1038/nm1377
- Lee, Y. H., Bae, S. C., & Song, G. G. (2014). Meta-analysis of associations between the peroxisome proliferator-activated receptor-gamma Pro12Ala polymorphism and susceptibility to nonalcoholic fatty liver disease, rheumatoid arthritis, and psoriatic arthritis. *Genet Test Mol Biomarkers*, *18*(5), 341-348. doi:10.1089/gtmb.2013.0503
- Lencioni, R., de Baere, T., Soulen, M. C., Rilling, W. S., & Geschwind, J. F. (2016). Lipiodol transarterial chemoembolization for hepatocellular carcinoma: A systematic review of efficacy and safety data. *Hepatology*, *64*(1), 106-116. doi:10.1002/hep.28453
- Li, J., Yang, D., He, Y., Wang, M., Wen, Z., Liu, L., . . . Lin, J. (2011). Associations of HLA-DP variants with hepatitis B virus infection in southern and northern Han Chinese populations: a multicenter case-control study. *PLoS One*, *6*(8), e24221. doi:10.1371/journal.pone.0024221
- Li, S., Qian, J., Yang, Y., Zhao, W., Dai, J., Bei, J. X., . . . Zhou, W. (2012). GWAS identifies novel susceptibility loci on 6p21.32 and 21q21.3 for hepatocellular carcinoma in chronic hepatitis B virus carriers. *PLoS Genet*, *8*(7), e1002791. doi:10.1371/journal.pgen.1002791
- Li, T. T., Li, T. H., Peng, J., He, B., Liu, L. S., Wei, D. H., . . . Tang, Z. H. (2018). TM6SF2: A novel target for plasma lipid regulation. *Atherosclerosis*, *268*, 170-176. doi:10.1016/j.atherosclerosis.2017.11.033
- Li, Y., Si, L., Zhai, Y., Hu, Y., Hu, Z., Bei, J. X., . . . Zhou, G. (2016). Genome-wide association study identifies 8p21.3 associated with persistent hepatitis B virus infection among Chinese. *Nat Commun*, *7*, 11664. doi:10.1038/ncomms11664
- Li, Y., Zhai, Y., Song, Q., Zhang, H., Cao, P., Ping, J., . . . Zhou, G. (2018). Genome-Wide Association Study Identifies a New Locus at 7q21.13 Associated with Hepatitis B Virus-Related Hepatocellular Carcinoma. *Clin Cancer Res*, *24*(4), 906-915. doi:10.1158/1078-0432.CCR-17-2537
- Liangpunsakul, S., Beaudoin, J. J., Shah, V. H., Puri, P., Sanyal, A. J., Kamath, P. S., . . . Chalasani, N. P. (2018). Interaction between the patatin-like phospholipase domain-containing protein 3 genotype and coffee drinking and the risk for acute alcoholic hepatitis. *Hepatol Commun*, *2*(1), 29-34. doi:10.1002/hep4.1123
- Liao, Y., Cai, B., Li, Y., Chen, J., Tao, C., Huang, H., & Wang, L. (2014). Association of HLA-DP/DQ and

- STAT4 polymorphisms with HBV infection outcomes and a mini meta-analysis. *PLoS One*, 9(11), e111677. doi:10.1371/journal.pone.0111677
- Lin, C. L., & Kao, J. H. (2015). Hepatitis B virus genotypes and variants. *Cold Spring Harb Perspect Med*, 5(5), a021436. doi:10.1101/cshperspect.a021436
- Lin, C. W., Lin, C. C., Mo, L. R., Chang, C. Y., Perng, D. S., Hsu, C. C., . . . Yang, S. S. (2013). Heavy alcohol consumption increases the incidence of hepatocellular carcinoma in hepatitis B virus-related cirrhosis. *J Hepatol*, 58(4), 730-735. doi:10.1016/j.jhep.2012.11.045
- Lin, Y., Shi, C. Y., Li, B., Soo, B. H., Mohammed-Ali, S., Wee, A., . . . Chan, S. H. (1996). Tumour suppressor p53 and Rb genes in human hepatocellular carcinoma. *Ann Acad Med Singapore*, 25(1), 22-30.
- Linden, D., Ahnmark, A., Pingitore, P., Ciociola, E., Ahlstedt, I., Andreasson, A. C., . . . Romeo, S. (2019). Pnpla3 silencing with antisense oligonucleotides ameliorates nonalcoholic steatohepatitis and fibrosis in Pnpla3 I148M knock-in mice. *Mol Metab*, 22, 49-61. doi:10.1016/j.molmet.2019.01.013
- Liu, W., Ma, N., Zhao, D., Gao, X., Zhang, X., Yang, L., & Liu, D. (2019). Correlation between the DEPDC5 rs1012068 polymorphism and the risk of HBV-related hepatocellular carcinoma. *Clin Res Hepatol Gastroenterol*. doi:10.1016/j.clinre.2018.12.005
- Liu, Y. L., Patman, G. L., Leathart, J. B., Piguat, A. C., Burt, A. D., Dufour, J. F., . . . Anstee, Q. M. (2014). Carriage of the PNPLA3 rs738409 C>G polymorphism confers an increased risk of non-alcoholic fatty liver disease associated hepatocellular carcinoma. *J Hepatol*, 61(1), 75-81. doi:10.1016/j.jhep.2014.02.030
- Liu, Y. L., Reeves, H. L., Burt, A. D., Tiniakos, D., McPherson, S., Leathart, J. B., . . . Anstee, Q. M. (2014). TM6SF2 rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. *Nat Commun*, 5, 4309. doi:10.1038/ncomms5309
- Llovet, J. M., Bru, C., & Bruix, J. (1999). Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis*, 19(3), 329-338. doi:10.1055/s-2007-1007122
- Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J. F., . . . Group, S. I. S. (2008). Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med*, 359(4), 378-390. doi:10.1056/NEJMoa0708857
- Llovet, J. M., Zucman-Rossi, J., Pikarsky, E., Sangro, B., Schwartz, M., Sherman, M., & Gores, G. (2016). Hepatocellular carcinoma. *Nat Rev Dis Primers*, 2, 16018. doi:10.1038/nrdp.2016.18
- Lok, A. S., Sterling, R. K., Everhart, J. E., Wright, E. C., Hoefs, J. C., Di Bisceglie, A. M., . . . Group, H.-C. T. (2010). Des-gamma-carboxy prothrombin and alpha-fetoprotein as biomarkers for the early

- detection of hepatocellular carcinoma. *Gastroenterology*, 138(2), 493-502. doi:10.1053/j.gastro.2009.10.031
- Loomba, R., Rivera, M. M., McBurney, R., Park, Y., Haynes-Williams, V., Rehermann, B., . . . Heller, T. (2011). The natural history of acute hepatitis C: clinical presentation, laboratory findings and treatment outcomes. *Aliment Pharmacol Ther*, 33(5), 559-565. doi:10.1111/j.1365-2036.2010.04549.x
- Ludwig, J., Viggiano, T. R., McGill, D. B., & Oh, B. J. (1980). Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clin Proc*, 55(7), 434-438.
- Ma, Y., Belyaeva, O. V., Brown, P. M., Fujita, K., Valles, K., Karki, S., . . . Rotman, Y. (2019). 17-Beta Hydroxysteroid Dehydrogenase 13 Is a Hepatic Retinol Dehydrogenase Associated With Histological Features of Nonalcoholic Fatty Liver Disease. *Hepatology*, 69(4), 1504-1519. doi:10.1002/hep.30350
- Macaluso, F. S., Maida, M., & Petta, S. (2015). Genetic background in nonalcoholic fatty liver disease: A comprehensive review. *World J Gastroenterol*, 21(39), 11088-11111. doi:10.3748/wjg.v21.i39.11088
- Mahdessian, H., Taxiarchis, A., Popov, S., Silveira, A., Franco-Cereceda, A., Hamsten, A., . . . van't Hooft, F. (2014). TM6SF2 is a regulator of liver fat metabolism influencing triglyceride secretion and hepatic lipid droplet content. *Proc Natl Acad Sci U S A*, 111(24), 8913-8918. doi:10.1073/pnas.1323785111
- Mancebo, A., Gonzalez-Dieguez, M. L., Cadahia, V., Varela, M., Perez, R., Navascues, C. A., . . . Rodriguez, M. (2013). Annual incidence of hepatocellular carcinoma among patients with alcoholic cirrhosis and identification of risk groups. *Clin Gastroenterol Hepatol*, 11(1), 95-101. doi:10.1016/j.cgh.2012.09.007
- Manchiero, C., Nunes, A., Magri, M. C., Dantas, B. P., Mazza, C. C., Barone, A. A., & Tengan, F. M. (2017). The rs738409 polymorphism of the PNPLA3 gene is associated with hepatic steatosis and fibrosis in Brazilian patients with chronic hepatitis C. *BMC Infect Dis*, 17(1), 780. doi:10.1186/s12879-017-2887-6
- Mancina, R. M., Ferri, F., Farcomeni, A., Molinaro, A., Maffongelli, A., Mischitelli, M., . . . Ginanni Corradini, S. (2019). A two gene-based risk score predicts alcoholic cirrhosis development in males with at-risk alcohol consumption. *Appl Clin Genet*, 12, 1-10. doi:10.2147/TACG.S187922
- Manolio, T. A. (2010). Genomewide association studies and assessment of the risk of disease. *N Engl J Med*, 363(2), 166-176. doi:10.1056/NEJMra0905980
- Manolio, T. A., Bailey-Wilson, J. E., & Collins, F. S. (2006). Genes, environment and the value of

- prospective cohort studies. *Nat Rev Genet*, 7(10), 812-820. doi:10.1038/nrg1919
- Marchais-Oberwinkler, S., Henn, C., Moller, G., Klein, T., Negri, M., Oster, A., . . . Adamski, J. (2011). 17beta-Hydroxysteroid dehydrogenases (17beta-HSDs) as therapeutic targets: protein structures, functions, and recent progress in inhibitor development. *J Steroid Biochem Mol Biol*, 125(1-2), 66-82. doi:10.1016/j.jsbmb.2010.12.013
- Marcos, M., Gomez-Munuera, M., Pastor, I., Gonzalez-Sarmiento, R., & Laso, F. J. (2009). Tumor necrosis factor polymorphisms and alcoholic liver disease: a HuGE review and meta-analysis. *Am J Epidemiol*, 170(8), 948-956. doi:10.1093/aje/kwp236
- Marcos, M., Pastor, I., Chamorro, A. J., Ciria-Abad, S., Gonzalez-Sarmiento, R., & Laso, F. J. (2011). Meta-analysis: glutathione-S-transferase allelic variants are associated with alcoholic liver disease. *Aliment Pharmacol Ther*, 34(10), 1159-1172. doi:10.1111/j.1365-2036.2011.04862.x
- Marrero, J. A., Fontana, R. J., Fu, S., Conjeevaram, H. S., Su, G. L., & Lok, A. S. (2005). Alcohol, tobacco and obesity are synergistic risk factors for hepatocellular carcinoma. *J Hepatol*, 42(2), 218-224. doi:10.1016/j.jhep.2004.10.005
- Marrero, J. A., Hussain, H. K., Nghiem, H. V., Umar, R., Fontana, R. J., & Lok, A. S. (2005). Improving the prediction of hepatocellular carcinoma in cirrhotic patients with an arterially-enhancing liver mass. *Liver Transpl*, 11(3), 281-289. doi:10.1002/lt.20357
- Matsuura, K., Sawai, H., Ikeo, K., Ogawa, S., Iio, E., Isogawa, M., . . . Japanese Genome-Wide Association Study Group for Viral, H. (2017). Genome-Wide Association Study Identifies TLL1 Variant Associated With Development of Hepatocellular Carcinoma After Eradication of Hepatitis C Virus Infection. *Gastroenterology*, 152(6), 1383-1394. doi:10.1053/j.gastro.2017.01.041
- Mazzaferro, V., Llovet, J. M., Miceli, R., Bhoori, S., Schiavo, M., Mariani, L., . . . Metroticket Investigator Study, G. (2009). Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. *Lancet Oncol*, 10(1), 35-43. doi:10.1016/S1470-2045(08)70284-5
- Mbarek, H., Ochi, H., Urabe, Y., Kumar, V., Kubo, M., Hosono, N., . . . Matsuda, K. (2011). A genome-wide association study of chronic hepatitis B identified novel risk locus in a Japanese population. *Hum Mol Genet*, 20(19), 3884-3892. doi:10.1093/hmg/ddr301
- McConathy, W. J., Gesquiere, J. C., Bass, H., Tartar, A., Fruchart, J. C., & Wang, C. S. (1992). Inhibition of lipoprotein lipase activity by synthetic peptides of apolipoprotein C-III. *J Lipid Res*, 33(7), 995-1003.
- McGuire, S. (2016). World Cancer Report 2014. Geneva, Switzerland: World Health Organization, International Agency for Research on Cancer, WHO Press, 2015. *Adv Nutr*, 7(2), 418-419.

doi:10.3945/an.116.012211

- McMahon, B. J., Alberts, S. R., Wainwright, R. B., Bulkow, L., & Lanier, A. P. (1990). Hepatitis B-related sequelae. Prospective study in 1400 hepatitis B surface antigen-positive Alaska native carriers. *Arch Intern Med*, *150*(5), 1051-1054.
- Meffert, P. J., Repp, K. D., Volzke, H., Weiss, F. U., Homuth, G., Kuhn, J. P., . . . Aghdassi, A. A. (2018). The PNPLA3 SNP rs738409:G allele is associated with increased liver disease-associated mortality but reduced overall mortality in a population-based cohort. *J Hepatol*, *68*(4), 858-860. doi:10.1016/j.jhep.2017.11.038
- Milano, M., Aghemo, A., Mancina, R. M., Fischer, J., Dongiovanni, P., De Nicola, S., . . . Valenti, L. (2015). Transmembrane 6 superfamily member 2 gene E167K variant impacts on steatosis and liver damage in chronic hepatitis C patients. *Hepatology*, *62*(1), 111-117. doi:10.1002/hep.27811
- Min, H. K., Sookoian, S., Pirola, C. J., Cheng, J., Mirshahi, F., & Sanyal, A. J. (2014). Metabolic profiling reveals that PNPLA3 induces widespread effects on metabolism beyond triacylglycerol remodeling in Huh-7 hepatoma cells. *Am J Physiol Gastrointest Liver Physiol*, *307*(1), G66-76. doi:10.1152/ajpgi.00335.2013
- Mittal, S., & El-Serag, H. B. (2013). Epidemiology of hepatocellular carcinoma: consider the population. *J Clin Gastroenterol*, *47* Suppl, S2-6. doi:10.1097/MCG.0b013e3182872f29
- Modin, L., Arshad, A., Wilkes, B., Benselin, J., Lloyd, C., Irving, W. L., & Kelly, D. A. (2019). Epidemiology and natural history of hepatitis C virus infection among children and young people. *J Hepatol*, *70*(3), 371-378. doi:10.1016/j.jhep.2018.11.013
- Moosavy, S. H., Davoodian, P., Nazarnezhad, M. A., Nejatizadeh, A., Eftekhar, E., & Mahboobi, H. (2017). Epidemiology, transmission, diagnosis, and outcome of Hepatitis C virus infection. *Electron Physician*, *9*(10), 5646-5656. doi:10.19082/5646
- Morgan, R. L., Baack, B., Smith, B. D., Yartel, A., Pitasi, M., & Falck-Ytter, Y. (2013). Eradication of hepatitis C virus infection and the development of hepatocellular carcinoma: a meta-analysis of observational studies. *Ann Intern Med*, *158*(5 Pt 1), 329-337. doi:10.7326/0003-4819-158-5-201303050-00005
- Morgan, T. R., Mandayam, S., & Jamal, M. M. (2004). Alcohol and hepatocellular carcinoma. *Gastroenterology*, *127*(5 Suppl 1), S87-96.
- Muhlbauer, M., Bosserhoff, A. K., Hartmann, A., Thasler, W. E., Weiss, T. S., Herfarth, H., . . . Hellerbrand, C. (2003). A novel MCP-1 gene polymorphism is associated with hepatic MCP-1 expression and severity of HCV-related liver disease. *Gastroenterology*, *125*(4), 1085-1093.
- Munirajan, A. K., Ando, K., Mukai, A., Takahashi, M., Suenaga, Y., Ohira, M., . . . Nakagawara, A. (2008).

- KIF1Bbeta functions as a haploinsufficient tumor suppressor gene mapped to chromosome 1p36.2 by inducing apoptotic cell death. *J Biol Chem*, 283(36), 24426-24434. doi:10.1074/jbc.M802316200
- N'Kontchou, G., Paries, J., Htar, M. T., Ganne-Carrie, N., Costentin, L., Grando-Lemaire, V., . . . Beaugrand, M. (2006). Risk factors for hepatocellular carcinoma in patients with alcoholic or viral C cirrhosis. *Clin Gastroenterol Hepatol*, 4(8), 1062-1068. doi:10.1016/j.cgh.2006.05.013
- Nahon, P., & Nault, J. C. (2017). Constitutional and functional genetics of human alcohol-related hepatocellular carcinoma. *Liver Int*, 37(11), 1591-1601. doi:10.1111/liv.13419
- Nakaoka, K., Hashimoto, S., Kawabe, N., Nitta, Y., Muraio, M., Nakano, T., . . . Yoshioka, K. (2015). PNPLA3 I148M associations with liver carcinogenesis in Japanese chronic hepatitis C patients. *Springerplus*, 4, 83. doi:10.1186/s40064-015-0870-5
- Nault, J. C., Mallet, M., Pilati, C., Calderaro, J., Bioulac-Sage, P., Laurent, C., . . . Zucman-Rossi, J. (2013). High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. *Nat Commun*, 4, 2218. doi:10.1038/ncomms3218
- Nguyen, K. B., Watford, W. T., Salomon, R., Hofmann, S. R., Pien, G. C., Morinobu, A., . . . Biron, C. A. (2002). Critical role for STAT4 activation by type 1 interferons in the interferon-gamma response to viral infection. *Science*, 297(5589), 2063-2066. doi:10.1126/science.1074900
- Niederau, C., Fischer, R., Purschel, A., Stremmel, W., Haussinger, D., & Strohmeyer, G. (1996). Long-term survival in patients with hereditary hemochromatosis. *Gastroenterology*, 110(4), 1107-1119.
- Nischalke, H. D., Berger, C., Luda, C., Berg, T., Muller, T., Grunhage, F., . . . Spengler, U. (2011). The PNPLA3 rs738409 148M/M genotype is a risk factor for liver cancer in alcoholic cirrhosis but shows no or weak association in hepatitis C cirrhosis. *PLoS One*, 6(11), e27087. doi:10.1371/journal.pone.0027087
- Nobili, V., Liccardo, D., Bedogni, G., Salvatori, G., Gnani, D., Bersani, I., . . . Raponi, M. (2014). Influence of dietary pattern, physical activity, and I148M PNPLA3 on steatosis severity in at-risk adolescents. *Genes Nutr*, 9(3), 392. doi:10.1007/s12263-014-0392-8
- Novikova, M. V., Khromova, N. V., & Kopnin, P. B. (2017). Components of the Hepatocellular Carcinoma Microenvironment and Their Role in Tumor Progression. *Biochemistry (Mosc)*, 82(8), 861-873. doi:10.1134/S0006297917080016
- Nowak, A., Giger, R. S., & Krayenbuehl, P. A. (2018). Higher age at diagnosis of hemochromatosis is the strongest predictor of the occurrence of hepatocellular carcinoma in the Swiss hemochromatosis cohort: A prospective longitudinal observational study. *Medicine (Baltimore)*, 97(42), e12886. doi:10.1097/MD.00000000000012886

- Nunez-Torres, R., Macias, J., Mancebo, M., Frias, M., Dolci, G., Tellez, F., . . . Group, H. E.-e. S. (2016). The PNPLA3 Genetic Variant rs738409 Influences the Progression to Cirrhosis in HIV/Hepatitis C Virus Coinfected Patients. *PLoS One*, *11*(12), e0168265. doi:10.1371/journal.pone.0168265
- Oda, K., Uto, H., Mawatari, S., & Ido, A. (2015). Clinical features of hepatocellular carcinoma associated with nonalcoholic fatty liver disease: a review of human studies. *Clin J Gastroenterol*, *8*(1), 1-9. doi:10.1007/s12328-014-0548-5
- Ogurtsova, K., da Rocha Fernandes, J. D., Huang, Y., Linnenkamp, U., Guariguata, L., Cho, N. H., . . . Makaroff, L. E. (2017). IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract*, *128*, 40-50. doi:10.1016/j.diabres.2017.03.024
- Omata, M., Yoshida, H., & Shiratori, Y. (2005). Prevention of hepatocellular carcinoma and its recurrence in chronic hepatitis C patients by interferon therapy. *Clin Gastroenterol Hepatol*, *3*(10 Suppl 2), S141-143.
- Orito, E., Mizokami, M., Sakugawa, H., Michitaka, K., Ishikawa, K., Ichida, T., . . . Iino, S. (2001). A case-control study for clinical and molecular biological differences between hepatitis B viruses of genotypes B and C. Japan HBV Genotype Research Group. *Hepatology*, *33*(1), 218-223. doi:10.1053/jhep.2001.20532
- Pais, R., Fartoux, L., Goumard, C., Scatton, O., Wendum, D., Rosmorduc, O., & Ratziu, V. (2017). Temporal trends, clinical patterns and outcomes of NAFLD-related HCC in patients undergoing liver resection over a 20-year period. *Aliment Pharmacol Ther*, *46*(9), 856-863. doi:10.1111/apt.14261
- Park, B. K., Park, Y. N., Ahn, S. H., Lee, K. S., Chon, C. Y., Moon, Y. M., . . . Han, K. H. (2007). Long-term outcome of chronic hepatitis B based on histological grade and stage. *J Gastroenterol Hepatol*, *22*(3), 383-388. doi:10.1111/j.1440-1746.2007.04857.x
- Park, J. W., Chen, M., Colombo, M., Roberts, L. R., Schwartz, M., Chen, P. J., . . . Sherman, M. (2015). Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE Study. *Liver Int*, *35*(9), 2155-2166. doi:10.1111/liv.12818
- Patin, E., Kotalik, Z., Guergnon, J., Bibert, S., Nalpas, B., Jouanguy, E., . . . French, A. H. C. E. P. G. S. G. (2012). Genome-wide association study identifies variants associated with progression of liver fibrosis from HCV infection. *Gastroenterology*, *143*(5), 1244-1252 e1212. doi:10.1053/j.gastro.2012.07.097
- Petersen, K. F., Dufour, S., Hariri, A., Nelson-Williams, C., Foo, J. N., Zhang, X. M., . . . Shulman, G. I. (2010). Apolipoprotein C3 gene variants in nonalcoholic fatty liver disease. *N Engl J Med*, *362*(12), 1082-1089. doi:10.1056/NEJMoa0907295

References

- Petta, S., Maida, M., Grimaudo, S., Pipitone, R. M., Macaluso, F. S., Cabibi, D., . . . Craxi, A. (2016). TM6SF2 rs58542926 is not associated with steatosis and fibrosis in large cohort of patients with genotype 1 chronic hepatitis C. *Liver Int*, *36*(2), 198-204. doi:10.1111/liv.12918
- Pinero, F., Mendizabal, M., Ridruejo, E., Herz Wolff, F., Ameigeiras, B., Anders, M., . . . Lalrean. (2019). Treatment with direct-acting antivirals for HCV decreases but does not eliminate the risk of hepatocellular carcinoma. *Liver Int*. doi:10.1111/liv.14041
- Pirazzi, C., Valenti, L., Motta, B. M., Pingitore, P., Hedfalk, K., Mancina, R. M., . . . Romeo, S. (2014). PNPLA3 has retinyl-palmitate lipase activity in human hepatic stellate cells. *Hum Mol Genet*, *23*(15), 4077-4085. doi:10.1093/hmg/ddu121
- Pirola, C. J., Garaycochea, M., Flichman, D., Arrese, M., San Martino, J., Gazzi, C., . . . Sookoian, S. (2019). Splice variant rs72613567 prevents worst histologic outcomes in patients with nonalcoholic fatty liver disease. *J Lipid Res*, *60*(1), 176-185. doi:10.1194/jlr.P089953
- Pirola, C. J., & Sookoian, S. (2015). The dual and opposite role of the TM6SF2-rs58542926 variant in protecting against cardiovascular disease and conferring risk for nonalcoholic fatty liver: A meta-analysis. *Hepatology*, *62*(6), 1742-1756. doi:10.1002/hep.28142
- Popper, H., & Uenfriend, S. (1970). Hepatic fibrosis. Correlation of biochemical and morphologic investigations. *Am J Med*, *49*, 707-721.
- Powell, E. E., Edwards-Smith, C. J., Hay, J. L., Clouston, A. D., Crawford, D. H., Shorthouse, C., . . . Jonsson, J. R. (2000). Host genetic factors influence disease progression in chronic hepatitis C. *Hepatology*, *31*(4), 828-833. doi:10.1053/he.2000.6253
- Rauch, A., Kutalik, Z., Descombes, P., Cai, T., Di Iulio, J., Mueller, T., . . . Swiss, H. I. V. C. S. (2010). Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology*, *138*(4), 1338-1345, 1345 e1331-1337. doi:10.1053/j.gastro.2009.12.056
- Romeo, S., Kozlitina, J., Xing, C., Pertsemlidis, A., Cox, D., Pennacchio, L. A., . . . Hobbs, H. H. (2008). Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet*, *40*(12), 1461-1465. doi:10.1038/ng.257
- Rosendahl, J., Tonjes, A., Schleinitz, D., Kovacs, P., Wiegand, J., Ruffert, C., . . . Drenth, J. P. (2012). A common variant of PNPLA3 (p.I148M) is not associated with alcoholic chronic pancreatitis. *PLoS One*, *7*(1), e29433. doi:10.1371/journal.pone.0029433
- Sagnelli, C., Merli, M., Uberti-Foppa, C., Hasson, H., Grandone, A., Cirillo, G., . . . Sagnelli, E. (2016). TM6SF2 E167K variant predicts severe liver fibrosis for human immunodeficiency/hepatitis C virus co-infected patients, and severe steatosis only for a non-3 hepatitis C virus genotype. *World*

- J Gastroenterol*, 22(38), 8509-8518. doi:10.3748/wjg.v22.i38.8509
- Saha, B., Jyothi Prasanna, S., Chandrasekar, B., & Nandi, D. (2010). Gene modulation and immunoregulatory roles of interferon gamma. *Cytokine*, 50(1), 1-14. doi:10.1016/j.cyto.2009.11.021
- Salameh, H., Raff, E., Erwin, A., Seth, D., Nischalke, H. D., Falleti, E., . . . Singal, A. K. (2015). PNPLA3 Gene Polymorphism Is Associated With Predisposition to and Severity of Alcoholic Liver Disease. *Am J Gastroenterol*, 110(6), 846-856. doi:10.1038/ajg.2015.137
- Saloniemi, T., Jokela, H., Strauss, L., Pakarinen, P., & Poutanen, M. (2012). The diversity of sex steroid action: novel functions of hydroxysteroid (17beta) dehydrogenases as revealed by genetically modified mouse models. *J Endocrinol*, 212(1), 27-40. doi:10.1530/JOE-11-0315
- Sanchez-Tapias, J. M., Costa, J., Mas, A., Bruguera, M., & Rodes, J. (2002). Influence of hepatitis B virus genotype on the long-term outcome of chronic hepatitis B in western patients. *Gastroenterology*, 123(6), 1848-1856. doi:10.1053/gast.2002.37041
- Santi, V., Trevisani, F., Gramenzi, A., Grignaschi, A., Mirici-Cappa, F., Del Poggio, P., . . . Italian Liver Cancer, G. (2010). Semiannual surveillance is superior to annual surveillance for the detection of early hepatocellular carcinoma and patient survival. *J Hepatol*, 53(2), 291-297. doi:10.1016/j.jhep.2010.03.010
- Santoro, N., Savoye, M., Kim, G., Marotto, K., Shaw, M. M., Pierpont, B., & Caprio, S. (2012). Hepatic fat accumulation is modulated by the interaction between the rs738409 variant in the PNPLA3 gene and the dietary omega6/omega3 PUFA intake. *PLoS One*, 7(5), e37827. doi:10.1371/journal.pone.0037827
- Sanyal, A., Poklepovic, A., Moyneur, E., & Barghout, V. (2010). Population-based risk factors and resource utilization for HCC: US perspective. *Curr Med Res Opin*, 26(9), 2183-2191. doi:10.1185/03007995.2010.506375
- Schlisio, S., Kenchappa, R. S., Vredeveld, L. C., George, R. E., Stewart, R., Greulich, H., . . . Kaelin, W. G., Jr. (2008). The kinesin KIF1Bbeta acts downstream from EglN3 to induce apoptosis and is a potential 1p36 tumor suppressor. *Genes Dev*, 22(7), 884-893. doi:10.1101/gad.1648608
- Schulze, K., Imbeaud, S., Letouze, E., Alexandrov, L. B., Calderaro, J., Rebouissou, S., . . . Zucman-Rossi, J. (2015). Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet*, 47(5), 505-511. doi:10.1038/ng.3252
- Schweitzer, A., Horn, J., Mikolajczyk, R. T., Krause, G., & Ott, J. J. (2015). Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. *Lancet*, 386(10003), 1546-1555. doi:10.1016/S0140-6736(15)61412-X

- Seko, Y., Sumida, Y., Tanaka, S., Mori, K., Taketani, H., Ishiba, H., . . . Itoh, Y. (2017). Development of hepatocellular carcinoma in Japanese patients with biopsy-proven non-alcoholic fatty liver disease: Association between PNPLA3 genotype and hepatocarcinogenesis/fibrosis progression. *Hepatol Res*, *47*(11), 1083-1092. doi:10.1111/hepr.12840
- Seth, D., Daly, A. K., Haber, P. S., & Day, C. P. (2010). Patatin-like phospholipase domain containing 3: a case in point linking genetic susceptibility for alcoholic and nonalcoholic liver disease. *Hepatology*, *51*(4), 1463-1465. doi:10.1002/hep.23606
- Sevastianova, K., Kotronen, A., Gastaldelli, A., Perttila, J., Hakkarainen, A., Lundbom, J., . . . Yki-Jarvinen, H. (2011). Genetic variation in PNPLA3 (adiponutrin) confers sensitivity to weight loss-induced decrease in liver fat in humans. *Am J Clin Nutr*, *94*(1), 104-111. doi:10.3945/ajcn.111.012369
- Shin, J. G., Cheong, H. S., Lee, K., Ju, B. G., Lee, J. H., Yu, S. J., . . . Shin, H. D. (2017). Identification of novel OCT4 genetic variant associated with the risk of chronic hepatitis B in a Korean population. *Liver Int*, *37*(3), 354-361. doi:10.1111/liv.13245
- Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA Cancer J Clin*, *68*(1), 7-30. doi:10.3322/caac.21442
- Singal, A., Volk, M. L., Waljee, A., Salgia, R., Higgins, P., Rogers, M. A., & Marrero, J. A. (2009). Meta-analysis: surveillance with ultrasound for early-stage hepatocellular carcinoma in patients with cirrhosis. *Aliment Pharmacol Ther*, *30*(1), 37-47. doi:10.1111/j.1365-2036.2009.04014.x
- Singal, A. G., Manjunath, H., Yopp, A. C., Beg, M. S., Marrero, J. A., Gopal, P., & Waljee, A. K. (2014). The effect of PNPLA3 on fibrosis progression and development of hepatocellular carcinoma: a meta-analysis. *Am J Gastroenterol*, *109*(3), 325-334. doi:10.1038/ajg.2013.476
- Sookoian, S., Castano, G. O., Scian, R., Mallardi, P., Fernandez Gianotti, T., Burgueno, A. L., . . . Pirola, C. J. (2015). Genetic variation in transmembrane 6 superfamily member 2 and the risk of nonalcoholic fatty liver disease and histological disease severity. *Hepatology*, *61*(2), 515-525. doi:10.1002/hep.27556
- Sookoian, S., & Pirola, C. J. (2011). Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease. *Hepatology*, *53*(6), 1883-1894. doi:10.1002/hep.24283
- Sopipong, W., Tangkijvanich, P., Payungporn, S., Posuwan, N., & Poovorawan, Y. (2013). The KIF1B (rs17401966) single nucleotide polymorphism is not associated with the development of HBV-related hepatocellular carcinoma in Thai patients. *Asian Pac J Cancer Prev*, *14*(5), 2865-2869.
- Speliotes, E. K., Yerges-Armstrong, L. M., Wu, J., Hernaez, R., Kim, L. J., Palmer, C. D., . . . Consortium, G. (2011). Genome-wide association analysis identifies variants associated with nonalcoholic fatty

- liver disease that have distinct effects on metabolic traits. *PLoS Genet*, 7(3), e1001324. doi:10.1371/journal.pgen.1001324
- Stickel, F., Buch, S., Lau, K., Meyer zu Schwabedissen, H., Berg, T., Ridinger, M., . . . Hampe, J. (2011). Genetic variation in the PNPLA3 gene is associated with alcoholic liver injury in caucasians. *Hepatology*, 53(1), 86-95. doi:10.1002/hep.24017
- Stickel, F., Buch, S., Nischalke, H. D., Weiss, K. H., Gotthardt, D., Fischer, J., . . . Hampe, J. (2018a). Correction: Genetic Variants in PNPLA3 and TM6SF2 Predispose to the Development of Hepatocellular Carcinoma in Individuals With Alcohol-Related Cirrhosis. *Am J Gastroenterol*, 113(7), 1099. doi:10.1038/s41395-018-0118-4
- Stickel, F., Buch, S., Nischalke, H. D., Weiss, K. H., Gotthardt, D., Fischer, J., . . . Hampe, J. (2018b). Genetic variants in PNPLA3 and TM6SF2 predispose to the development of hepatocellular carcinoma in individuals with alcohol-related cirrhosis. *Am J Gastroenterol*, 113(10), 1475-1483. doi:10.1038/s41395-018-0041-8
- Su, W., Mao, Z., Liu, Y., Zhang, X., Zhang, W., Gustafsson, J. A., & Guan, Y. (2019). Role of HSD17B13 in the liver physiology and pathophysiology. *Mol Cell Endocrinol*, 489, 119-125. doi:10.1016/j.mce.2018.10.014
- Su, W., Wang, Y., Jia, X., Wu, W., Li, L., Tian, X., . . . Guan, Y. (2014). Comparative proteomic study reveals 17beta-HSD13 as a pathogenic protein in nonalcoholic fatty liver disease. *Proc Natl Acad Sci U S A*, 111(31), 11437-11442. doi:10.1073/pnas.1410741111
- Sung, H., Siegel, R. L., Torre, L. A., Pearson-Stuttard, J., Islami, F., Fedewa, S. A., . . . Jemal, A. (2019). Global patterns in excess body weight and the associated cancer burden. *CA Cancer J Clin*, 69(2), 88-112. doi:10.3322/caac.21499
- Sung, J. J., Tsoi, K. K., Wong, V. W., Li, K. C., & Chan, H. L. (2008). Meta-analysis: Treatment of hepatitis B infection reduces risk of hepatocellular carcinoma. *Aliment Pharmacol Ther*, 28(9), 1067-1077. doi:10.1111/j.1365-2036.2008.03816.x
- Suppiah, V., Moldovan, M., Ahlenstiel, G., Berg, T., Weltman, M., Abate, M. L., . . . George, J. (2009). IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet*, 41(10), 1100-1104. doi:10.1038/ng.447
- Szabo, E., Lotz, G., Paska, C., Kiss, A., & Schaff, Z. (2003). Viral hepatitis: new data on hepatitis C infection. *Pathol Oncol Res*, 9(4), 215-221. doi:PAOR.2003.9.4.0215
- Tamaki, N., Kurosaki, M., Higuchi, M., Takada, H., Nakakuki, N., Yasui, Y., . . . Izumi, N. (2015). Genetic Polymorphisms of IL28B and PNPLA3 Are Predictive for HCV Related Rapid Fibrosis Progression and Identify Patients Who Require Urgent Antiviral Treatment with New Regimens.

- PLoS One*, 10(9), e0137351. doi:10.1371/journal.pone.0137351
- Tanaka, Y., Hanada, K., Mizokami, M., Yeo, A. E., Shih, J. W., Gojobori, T., & Alter, H. J. (2002). A comparison of the molecular clock of hepatitis C virus in the United States and Japan predicts that hepatocellular carcinoma incidence in the United States will increase over the next two decades. *Proc Natl Acad Sci U S A*, 99(24), 15584-15589. doi:10.1073/pnas.242608099
- Tanaka, Y., Nishida, N., Sugiyama, M., Kurosaki, M., Matsuura, K., Sakamoto, N., . . . Mizokami, M. (2009). Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet*, 41(10), 1105-1109. doi:10.1038/ng.449
- Thierfelder, W. E., van Deursen, J. M., Yamamoto, K., Tripp, R. A., Sarawar, S. R., Carson, R. T., . . . Ihle, J. N. (1996). Requirement for Stat4 in interleukin-12-mediated responses of natural killer and T cells. *Nature*, 382(6587), 171-174. doi:10.1038/382171a0
- Thomas, R., Thio, C. L., Apps, R., Qi, Y., Gao, X., Marti, D., . . . Carrington, M. (2012). A novel variant marking HLA-DP expression levels predicts recovery from hepatitis B virus infection. *J Virol*, 86(12), 6979-6985. doi:10.1128/JVI.00406-12
- Tian, C., Stokowski, R. P., Kershenovich, D., Ballinger, D. G., & Hinds, D. A. (2010). Variant in PNPLA3 is associated with alcoholic liver disease. *Nat Genet*, 42(1), 21-23. doi:10.1038/ng.488
- Tong, H. V., Toan, N. L., Song, L. H., Bock, C. T., Kreamsner, P. G., & Velavan, T. P. (2013). Hepatitis B virus-induced hepatocellular carcinoma: functional roles of MICA variants. *J Viral Hepat*, 20(10), 687-698. doi:10.1111/jvh.12089
- Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. *CA Cancer J Clin*, 65(2), 87-108. doi:10.3322/caac.21262
- Trepo, E., Gustot, T., Degre, D., Lemmers, A., Verset, L., Demetter, P., . . . Moreno, C. (2011). Common polymorphism in the PNPLA3/adiponutrin gene confers higher risk of cirrhosis and liver damage in alcoholic liver disease. *J Hepatol*, 55(4), 906-912. doi:10.1016/j.jhep.2011.01.028
- Trepo, E., Guyot, E., Ganne-Carrie, N., Degre, D., Gustot, T., Franchimont, D., . . . Moreno, C. (2012). PNPLA3 (rs738409 C>G) is a common risk variant associated with hepatocellular carcinoma in alcoholic cirrhosis. *Hepatology*, 55(4), 1307-1308. doi:10.1002/hep.25518
- Trepo, E., Nahon, P., Bontempi, G., Valenti, L., Falleti, E., Nischalke, H. D., . . . Deltenre, P. (2014). Association between the PNPLA3 (rs738409 C>G) variant and hepatocellular carcinoma: Evidence from a meta-analysis of individual participant data. *Hepatology*, 59(6), 2170-2177. doi:10.1002/hep.26767
- Trepo, E., Pradat, P., Potthoff, A., Momozawa, Y., Quertinmont, E., Gustot, T., . . . Moreno, C. (2011). Impact of patatin-like phospholipase-3 (rs738409 C>G) polymorphism on fibrosis progression

- and steatosis in chronic hepatitis C. *Hepatology*, 54(1), 60-69. doi:10.1002/hep.24350
- Trepo, E., Romeo, S., Zucman-Rossi, J., & Nahon, P. (2016). PNPLA3 gene in liver diseases. *J Hepatol*, 65(2), 399-412. doi:10.1016/j.jhep.2016.03.011
- Trevisani, F., Frigerio, M., Santi, V., Grignaschi, A., & Bernardi, M. (2010). Hepatocellular carcinoma in non-cirrhotic liver: a reappraisal. *Dig Liver Dis*, 42(5), 341-347. doi:10.1016/j.dld.2009.09.002
- Trinchet, J. C., Chaffaut, C., Bourcier, V., Degos, F., Henrion, J., Fontaine, H., . . . Groupe d'Etude et de Traitement du Carcinome, H. (2011). Ultrasonographic surveillance of hepatocellular carcinoma in cirrhosis: a randomized trial comparing 3- and 6-month periodicities. *Hepatology*, 54(6), 1987-1997. doi:10.1002/hep.24545
- Tzartzeva, K., Obi, J., Rich, N. E., Parikh, N. D., Marrero, J. A., Yopp, A., . . . Singal, A. G. (2018). Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis. *Gastroenterology*, 154(6), 1706-1718 e1701. doi:10.1053/j.gastro.2018.01.064
- Ulveling, D., Le Clerc, S., Cobat, A., Labib, T., Noirel, J., Laville, V., . . . French, A. H. C. E. P. G. S. G. (2016). A new 3p25 locus is associated with liver fibrosis progression in human immunodeficiency virus/hepatitis C virus-coinfected patients. *Hepatology*, 64(5), 1462-1472. doi:10.1002/hep.28695
- Urabe, Y., Ochi, H., Kato, N., Kumar, V., Takahashi, A., Muroyama, R., . . . Matsuda, K. (2013). A genome-wide association study of HCV-induced liver cirrhosis in the Japanese population identifies novel susceptibility loci at the MHC region. *J Hepatol*, 58(5), 875-882. doi:10.1016/j.jhep.2012.12.024
- Valenti, L., Al-Serri, A., Daly, A. K., Galmozzi, E., Rametta, R., Dongiovanni, P., . . . Day, C. P. (2010). Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. *Hepatology*, 51(4), 1209-1217. doi:10.1002/hep.23622
- Valenti, L., Alisi, A., Galmozzi, E., Bartuli, A., Del Menico, B., Alterio, A., . . . Nobili, V. (2010). I148M patatin-like phospholipase domain-containing 3 gene variant and severity of pediatric nonalcoholic fatty liver disease. *Hepatology*, 52(4), 1274-1280. doi:10.1002/hep.23823
- Valenti, L., Rumi, M., Galmozzi, E., Aghemo, A., Del Menico, B., De Nicola, S., . . . Fargion, S. (2011). Patatin-like phospholipase domain-containing 3 I148M polymorphism, steatosis, and liver damage in chronic hepatitis C. *Hepatology*, 53(3), 791-799. doi:10.1002/hep.24123
- Vandenbulcke, H., Moreno, C., Colle, I., Knebel, J. F., Francque, S., Serste, T., . . . Deltenre, P. (2016). Alcohol intake increases the risk of HCC in hepatitis C virus-related compensated cirrhosis: A prospective study. *J Hepatol*, 65(3), 543-551. doi:10.1016/j.jhep.2016.04.031
- Varela-Rey, M., Woodhoo, A., Martinez-Chantar, M. L., Mato, J. M., & Lu, S. C. (2013). Alcohol, DNA

- methylation, and cancer. *Alcohol Res*, 35(1), 25-35.
- Vermehren, J., Lotsch, J., Susser, S., Wicker, S., Berger, A., Zeuzem, S., . . . Doehring, A. (2012). A common HLA-DPA1 variant is associated with hepatitis B virus infection but fails to distinguish active from inactive Caucasian carriers. *PLoS One*, 7(3), e32605. doi:10.1371/journal.pone.0032605
- Vernon, G., Baranova, A., & Younossi, Z. M. (2011). Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther*, 34(3), 274-285. doi:10.1111/j.1365-2036.2011.04724.x
- Wallace, M. C., Preen, D. B., Short, M. W., Adams, L. A., & Jeffrey, G. P. (2019). Hepatocellular carcinoma in Australia 1982-2014: Increasing incidence and improving survival. *Liver Int*, 39(3), 522-530. doi:10.1111/liv.13966
- Wang, F. S. (2003). Current status and prospects of studies on human genetic alleles associated with hepatitis B virus infection. *World J Gastroenterol*, 9(4), 641-644. doi:10.3748/wjg.v9.i4.641
- Wang, J., Guo, X., Wu, P., Song, J., Ye, C., Yu, S., . . . Dong, W. (2013). Association between the Pro12Ala polymorphism of PPAR-gamma gene and the non-alcoholic fatty liver disease: a meta-analysis. *Gene*, 528(2), 328-334. doi:10.1016/j.gene.2013.07.014
- Wang, X., Liao, X., Yang, C., Huang, K., Yu, T., Yu, L., . . . Peng, T. (2019). Identification of prognostic biomarkers for patients with hepatocellular carcinoma after hepatectomy. *Oncol Rep*, 41(3), 1586-1602. doi:10.3892/or.2019.6953
- Ward, J. W. (2013). The hidden epidemic of hepatitis C virus infection in the United States: occult transmission and burden of disease. *Top Antivir Med*, 21(1), 15-19.
- Whitfield, J. B., Zhu, G., Madden, P. A. F., Montgomery, G. W., Heath, A. C., & Martin, N. G. (2019). Biomarker and Genomic Risk Factors for Liver Function Test Abnormality in Hazardous Drinkers. *Alcohol Clin Exp Res*, 43(3), 473-482. doi:10.1111/acer.13949
- Wichajarn, K., Kosalaraksa, P., & Wiangnon, S. (2008). Incidence of hepatocellular carcinoma in children in Khon Kaen before and after national hepatitis B vaccine program. *Asian Pac J Cancer Prev*, 9(3), 507-509.
- Wong, V. W., Wong, G. L., Tse, C. H., & Chan, H. L. (2014). Prevalence of the TM6SF2 variant and non-alcoholic fatty liver disease in Chinese. *J Hepatol*, 61(3), 708-709. doi:10.1016/j.jhep.2014.04.047
- Wray, N. R., Goddard, M. E., & Visscher, P. M. (2008). Prediction of individual genetic risk of complex disease. *Curr Opin Genet Dev*, 18(3), 257-263. doi:10.1016/j.gde.2008.07.006
- Xing, X., Huang, Y., Wang, S., Chi, M., Zeng, Y., Chen, L., . . . Liu, J. (2015). Comparative analysis of primary hepatocellular carcinoma with single and multiple lesions by iTRAQ-based quantitative

- proteomics. *J Proteomics*, 128, 262-271. doi:10.1016/j.jprot.2015.08.007
- Xu, G., Yang, F., Ding, C. L., Zhao, L. J., Ren, H., Zhao, P., . . . Qi, Z. T. (2014). Small nucleolar RNA 113-1 suppresses tumorigenesis in hepatocellular carcinoma. *Mol Cancer*, 13, 216. doi:10.1186/1476-4598-13-216
- Xu, R., Tao, A., Zhang, S., Deng, Y., & Chen, G. (2015). Association between patatin-like phospholipase domain containing 3 gene (PNPLA3) polymorphisms and nonalcoholic fatty liver disease: a HuGE review and meta-analysis. *Sci Rep*, 5, 9284. doi:10.1038/srep09284
- Yang, H. I., Lu, S. N., Liaw, Y. F., You, S. L., Sun, C. A., Wang, L. Y., . . . Taiwan Community-Based Cancer Screening Project, G. (2002). Hepatitis B e antigen and the risk of hepatocellular carcinoma. *N Engl J Med*, 347(3), 168-174. doi:10.1056/NEJMoa013215
- Yang, J. D., Kim, W. R., Coelho, R., Mettler, T. A., Benson, J. T., Sanderson, S. O., . . . Roberts, L. R. (2011). Cirrhosis is present in most patients with hepatitis B and hepatocellular carcinoma. *Clin Gastroenterol Hepatol*, 9(1), 64-70. doi:10.1016/j.cgh.2010.08.019
- Yao, F. Y., Ferrell, L., Bass, N. M., Watson, J. J., Bacchetti, P., Venook, A., . . . Roberts, J. P. (2001). Liver transplantation for hepatocellular carcinoma: expansion of the tumor size limits does not adversely impact survival. *Hepatology*, 33(6), 1394-1403. doi:10.1053/jhep.2001.24563
- Yasui, K., Hashimoto, E., Komorizono, Y., Koike, K., Arii, S., Imai, Y., . . . Welfare of, J. (2011). Characteristics of patients with nonalcoholic steatohepatitis who develop hepatocellular carcinoma. *Clin Gastroenterol Hepatol*, 9(5), 428-433; quiz e450. doi:10.1016/j.cgh.2011.01.023
- Yasui, K., Kawaguchi, T., Shima, T., Mitsuyoshi, H., Seki, K., Sendo, R., . . . Okanoue, T. (2015). Effect of PNPLA3 rs738409 variant (I148 M) on hepatic steatosis, necroinflammation, and fibrosis in Japanese patients with chronic hepatitis C. *J Gastroenterol*, 50(8), 887-893. doi:10.1007/s00535-014-1018-z
- Yu, C. B., Zhu, L. Y., Wang, Y. G., Li, F., Zhang, X. Y., & Dai, W. J. (2016). Systemic transcriptome analysis of hepatocellular carcinoma. *Tumour Biol*, 37(10), 13323-13331. doi:10.1007/s13277-016-5286-5
- Yu, R. B., Hong, X., Ding, W. L., Tan, Y. F., Zhang, Y. X., Sun, N. X., . . . Ge, D. F. (2008). The association between the genetic polymorphism of HLA-DQA1, DQB1, and DRB1 and serum alanine aminotransferase levels in chronic hepatitis C in the Chinese population. *J Gastroenterol Hepatol*, 23(9), 1394-1402. doi:10.1111/j.1440-1746.2007.05215.x
- Yuan, X., Waterworth, D., Perry, J. R., Lim, N., Song, K., Chambers, J. C., . . . Mooser, V. (2008). Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. *Am J Hum Genet*, 83(4), 520-528. doi:10.1016/j.ajhg.2008.09.012

- Zain, S. M., Mohamed, Z., & Mohamed, R. (2015). Common variant in the glucokinase regulatory gene rs780094 and risk of nonalcoholic fatty liver disease: a meta-analysis. *J Gastroenterol Hepatol*, 30(1), 21-27. doi:10.1111/jgh.12714
- Zaltron, S., Spinetti, A., Biasi, L., Baiguera, C., & Castelli, F. (2012). Chronic HCV infection: epidemiological and clinical relevance. *BMC Infect Dis*, 12 Suppl 2, S2. doi:10.1186/1471-2334-12-S2-S2
- Zavattari, P., Perra, A., Menegon, S., Kowalik, M. A., Petrelli, A., Angioni, M. M., . . . Columbano, A. (2015). Nrf2, but not beta-catenin, mutation represents an early event in rat hepatocarcinogenesis. *Hepatology*, 62(3), 851-862. doi:10.1002/hep.27790
- Zhang, B., & Yang, B. (1999). Combined alpha fetoprotein testing and ultrasonography as a screening test for primary liver cancer. *J Med Screen*, 6(2), 108-110. doi:10.1136/jms.6.2.108
- Zhang, B. H., Yang, B. H., & Tang, Z. Y. (2004). Randomized controlled trial of screening for hepatocellular carcinoma. *J Cancer Res Clin Oncol*, 130(7), 417-422. doi:10.1007/s00432-004-0552-0
- Zhang, H., Chen, L., Xin, Y., Lou, Y., Liu, Y., & Xuan, S. (2014). Apolipoprotein c3 gene polymorphisms are not a risk factor for developing non-alcoholic Fatty liver disease: a meta-analysis. *Hepat Mon*, 14(10), e23100. doi:10.5812/hepatmon.23100
- Zhang, H., Xue, L., Chen, L., Jiang, S., Xin, Y., & Xuan, S. (2015). A Meta-Analysis of the Association Between the I148M Variant of Patatin-Like Phospholipase Domain Containing 3 Gene and the Presence of Chronic Hepatitis C. *Hepat Mon*, 15(11), e31987. doi:10.5812/hepatmon.31987
- Zhang, H., Zhai, Y., Hu, Z., Wu, C., Qian, J., Jia, W., . . . Zhou, G. (2010). Genome-wide association study identifies 1p36.22 as a new susceptibility locus for hepatocellular carcinoma in chronic hepatitis B virus carriers. *Nat Genet*, 42(9), 755-758. doi:10.1038/ng.638
- Zhang, L., You, W., Zhang, H., Peng, R., Zhu, Q., Yao, A., . . . Wu, J. (2015). PNPLA3 polymorphisms (rs738409) and non-alcoholic fatty liver disease risk and related phenotypes: a meta-analysis. *J Gastroenterol Hepatol*, 30(5), 821-829. doi:10.1111/jgh.12889
- Zhang, S., Wu, H., Wu, X., Lian, W., Wang, Y., Zhang, X., & Peng, X. (2015). Association between PNPLA3 rs738409 polymorphisms and risk of hepatocellular carcinoma and its development in patients with cirrhosis: a meta-analysis. *Int J Clin Exp Med*, 8(5), 6638-6649.
- Zhang, Y., Guo, T., Yang, F., Mao, Y., Li, L., Liu, C., . . . Huang, J. (2018). Single-nucleotide rs738409 polymorphisms in the PNPLA3 gene are strongly associated with alcoholic liver disease in Han Chinese males. *Hepatol Int*, 12(5), 429-437. doi:10.1007/s12072-018-9889-3
- Zhong, R., Tian, Y., Liu, L., Qiu, Q., Wang, Y., Rui, R., . . . Li, H. (2012). HBV-related hepatocellular carcinoma susceptibility gene KIF1B is not associated with development of chronic hepatitis B.

PLoS One, 7(2), e28839. doi:10.1371/journal.pone.0028839

Zucman-Rossi, J., Villanueva, A., Nault, J. C., & Llovet, J. M. (2015). Genetic Landscape and Biomarkers of Hepatocellular Carcinoma. *Gastroenterology*, 149(5), 1226-1239 e1224. doi:10.1053/j.gastro.2015.05.061