Hepatology Publish Ahead of Print
DOI:10.1097/HEP.0000000000000839

OPEN

Impact of Clonal Hematopoiesis of Indeterminate Potential on Hepatocellular Carcinoma in Individuals with Steatotic Liver Disease

Alfredo Marchetti[1,2]*, Serena Pelusi[3]*, Alessio Marella[1], Francesco Malvestiti[4], Antony Ricchiuti[1,2], Luisa Ronzoni[3], Marta Lionetti[2], Vittoria Moretti[3], Elisabetta Bugianesi[5], Luca Miele[6], Umberto Vespasiani-Gentilucci[7], Paola Dongiovanni[8], Alessandro Federico[9], Giorgio Soardo[10], Roberta D'Ambrosio[11], Misti V. McCain[12], Helen L. Reeves[12], Vincenzo La Mura[4,13], Daniele Prati[3], Niccolò Bolli[1,2§], Luca Valenti[3,4,14§], EPIDEMIC Study Investigators

[1] Department of Oncology and Hemato-Oncology, Università degli Studi di Milano, Milan, Italy
[2] Hematology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
[3] Transfusion Medicine Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
[4] Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy
[5] Department of Medical Sciences, Division of Gastroenterology, University of Turin, Turin, Italy
[6] Dipartimento Universitario Medicina e Chirurgia Traslazionale, Università Cattolica del Sacro Cuore, Rome, Italy; Area Medicina Interna, Gastroenterologia e Oncologia Medica, Fondazione Policlinico A. Gemelli IRCCS, Rome, Italy.
[7] Clinical Medicine and Hepatology Unit, Department of Medicine and Surgery, Campus Bio-Medico University of Rome, Rome, Italy.
[8] Medicine and Metabolic Diseases, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
[9] Department of Precision Medicine, University of Campania "Luigi Vanvitelli", Naples, Italy.
[10] Clinica Medica, Department of Medicine, European Excellence Center for Arterial Hypertension, University of Udine, Udine, Italy.
[11] Gastroenterology and Hepatology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy.
[12] Newcastle University Translational Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom
[13] General Medicine, Haemostasis and Thrombosis, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
[14] Biological Resource Center Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy.

* Equal contributors
§ Joint senior authors and correspondence address.

°EPIDEMIC-SERENA Study Investigators
Luca Valenti, Serena Pelusi, Francesco Malvestiti, Luisa Ronzoni, Sara Margarita, Cristiana Bianco, Alessandro Cherubini, Rossana Carpani, Vittoria Moretti, Daniele Prati, Paola Dongiovanni, Vittorio Borroni, Roberta D'Ambrosio, Valentina Vaira, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano
Chiara Rosso, Elisabetta Bugianesi, Università di Torino
Antonio Liguori, Luca Miele, Policlinico Gemelli Roma
Federica Tavaglione, Umberto Vespasiani-Gentilucci, Campus Biomedico Roma
Grazia Pennisi, Salvatore Petta, Università di Palermo
Francesco Paolo Russo, Università di Padova
Alessandro Federico, Napoli
Giorgio Soardo, Udine
Helen Reeves, Newcastle-upon-Tyne

Corresponding Author:

Luca Valenti, MD
Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy,
Precision Medicine Lab, Biological Resource Center Unit, Department of Transfusion Medicine, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico; via Francesco Sforza 35, 20122, Milan, Italy
E-mail: luca.valenti@unimi.it

Niccolò Bolli, MD PhD
Department of Oncology and Hemato-Oncology, University of Milan, 20122 Milan, Italy
Hematology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, 20122 Milan, Italy
E-mail: niccolo.bolli@unimi.it

Funding:
Italian Ministry of Health (Ministero della Salute), Ricerca Finalizzata RF-2016-02364358 ("Impact of whole exome sequencing on the clinical management of patients with advanced nonalcoholic fatty liver and cryptogenic liver disease") (LV);
Italian Ministry of Health (Ministero della Salute), Rete Cardiologica "CV-PREVITAL" (DP, LV);
Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Ricerca corrente (LV, DP, NB);
"Liver BIBLE" (PR-0391) (LV);
Innovative Medicines Initiative 2 joint undertaking of European Union's Horizon 2020 research and innovation programme and EFPIA European Union (EU) Programme Horizon 2020 (under grant agreement No. 777377) for the project LITMUS (LV);
The European Union, programme "Photonics" under grant agreement "101016726" (LV);
Gilead_IN-IT-989-5790 (LV).
The European Union, Horizon-Europe "Genial" under grant agreement "101096312" (LV).

This work was supported by the European Research Council under the European Union's Horizon 2020 research and innovation program (817997, to N.B.); Associazione Italiana Ricerca sul Cancro (IG25739 to N.B)
Ricerca Finalizzata 2021 RF-2021-12373889, Italian Ministry of Health, Ricerca Finalizzata PNRR 2022 "RATIONAL: Risk strAtificaTIon Of Nonalcoholic fAtty Liver" PNR-MAD-2022-12375656.
European Community's Seventh Framework Programme (FP7/2001– 2013) under grant agreement HEALTH-F2-2009-241762 for the project FLIP (HR)
Cancer Research UK centre grant C9380/A18084; programme grant C18342/A23390 and Accelerator award C9380/A26813 (HR)

Conflicts of interest:
Roberta D'Ambrosio consults and is on the speakers' bureau for AbbVie and Gilead. Helen L. Reeves advises Bayer, Boston Scientific, and Eisai. She received grants from AstraZeneca. Vincenzo La Mura consults for AlfaSigma, Biomarin, CSL Behring, and Pfizer. He is on the speakers' bureau for Gore. He received grants from Sanofi and Takeda. Daniele Prati consults, advises, is on the speaker's bureau, and received grants from Macropharma. He consults, is on the speaker's bureau, and received grants from Diamed, Diatech, Grifols, Immucor, Ortho Clinical Diagnostics, and Terumo. He is on the speakers' bureau for Diasorin. Niccolò Bolli consults and is on the speakers' bureau for Janssen. He consults for Pfizer. He is on the speakers' bureau for from Amgen and Jazz. Luca Valenti consults, is on the speakers' bureau, and received grants from Gilead. He consults for AstraZeneca, Boehringer Ingelheim, Diatech, Intercept, Ionis, Novo Nordisk, Pfizer, and Resalis. He is on the speakers' bureau for AbbVie, AlfaSigma, and MSD. The remaining authors have no conflicts to report.

Authors' contributions:
Study design: LV, NB; funding and oversight: LV, NB, DP; writing of first manuscript draft: AM, LV, NB, SP; coordination of clinical study: SP, LV; clinical data collection: SP, EB LM, UVP, RDA, AF, PD, GS, MVCM, HLR, FP; data analysis: AM, SP, AM, FM, AR, LR, ML.
All authors read and approved the final manuscript version.

Graphical Abstract
GA1

ABSTRACT

Background & Aims: Metabolic dysfunction associated steatotic liver disease (MASLD) is a global epidemic and is the most rapidly rising cause of hepatocellular carcinoma (HCC). Clonal hematopoiesis of indeterminate potential (CHIP) contributes to neoplastic and cardiometabolic disorders and is considered a harbinger of tissue inflammation. CHIP was recently associated with increased risk of liver disease. The aim of this study was to examine whether CHIP is associated with HCC development in patients with SLD.

Methods: We considered individuals with MASLD-HCC (n=208) and controls with (n=414) and without (n=259) advanced fibrosis who underwent whole exome sequencing. CHIP was diagnosed when ≥2 variant callers identified a known myeloid mutation with VAF ≥2%.

Results: CHIP was observed in 116 participants (13.1%), most frequently in *DNMT3A*, *TET2*, *TP53* and *ASXL1*, and correlated with age ($p<0.0001$) and advanced liver fibrosis ($p=0.001$). Higher AST levels predicted non-*DNMT3A*-CHIP, in particular with variant allele frequency (VAF)≥10% (OR 1.14, 1.03-1.28 and OR 1.30, 1.12-1.49, respectively, $p<0.05$). After adjustment for sex, diabetes and a polygenic risk score of inherited MASLD predisposition CHIP was associated with cirrhosis (2.00, 1.30-3.15, $p=0.02$), and with HCC even after further adjustment for cirrhosis (OR 1.81, 1.11-2.00, 1.30-3.15, $p=0.002$). Despite the strong collinearity among aging and development of CHIP and HCC, non-*DNTM3A*-CHIP and *TET2* lesions remained associated with HCC after full correction for clinical/genetics covariates and age (OR 2.45, 1.35-4.53; OR 4.8, 1.60-17.0, $p=0.02$).

Conclusions: We observed an independent association between CHIP, particularly related to non-*DNTM3A* and *TET2* genetic lesions, and MASLD-HCC.

Keywords: CHIP; cirrhosis; genetics; MAFLD; NAFLD; TET2.

INTRODUCTION

Steatotic liver disease (SLD), most often associated with insulin resistance (nonalcoholic or now redefined metabolic dysfunction associated steatotic liver disease, NAFLD/MASLD), is now the leading cause of liver disease worldwide [1]. MASLD encompasses a range of disease severity, from simple steatosis to steatohepatitis, which can lead to liver fibrosis and HCC [2]. Advanced fibrosis is the main predisposing condition for the development of hepatocellular carcinoma (HCC) and decompensated cirrhosis, which are the main liver-related events [3]. HCC is the fifth most common cancer and third most common cause of cancer-related death globally. While progress has been made in treating viral hepatitis, the increasing prevalence of obesity and diabetes is driving a surge in HCC incidence, with MASLD being the most rapidly increasing cause of HCC, particularly in women [4, 5]. Insulin resistance and type 2 diabetes are major factors in the development of advanced fibrosis and HCC [6]. MASLD is a highly heritable condition, with common and rare genetic risk variants predisposing hepatocellular fat accumulation and lipotoxicity increasing the risk of steatohepatitis, fibrogenesis and HCC [7, 8], acting in synergy with insulin resistance [8, 9]. Inflammation triggered by lipotoxicity, as well as altered intestinal flora, and gut permeability, play a key role in MASLD progression [10, 11]. The development of acquired genetic mutations, selected by a lipotoxic environment, can also contribute to the evolution of hepatic parenchymal damage and multistage carcinogenesis [12].

Clonal hematopoiesis of indeterminate potential (CHIP) is defined as the presence of somatic mutations in hematopoietic stem cells with a variant allele frequency (VAF)≥2% located in genes affected in hematologic malignancies. CHIP prevalence increases with age and is associated with an increased risk of hematologic cancers, cardiovascular and other aging-related disorders [14]. Recently, CHIP has also been linked to increased susceptibility to SLD and severe liver disease [13]. The association of CHIP with liver inflammation and fibrosis was independent of steatosis but related to acquired mutations in some CHIP-related genes (especially *TET2*) in myeloid cells homing to the liver [13]. However, it is not yet known whether CHIP can facilitate the progression of liver disease to HCC independently of liver fibrosis.

The aim of this study was therefore to examine whether CHIP is associated with HCC independently of cirrhosis and other major clinical determinants in a cross-sectional case-control multi-center cohort of patients with MASLD and controls.

PATIENTS AND METHODS

Study cohorts

The EPIDEMIC-NAFLD (now MASLD) ("Exome sequencing for the identification of genetic mutations promoting hepatocellular carcinoma development in nonalcoholic fatty liver disease") is a cross-sectional multicenter case-control study cohort aimed at the identification of genetic variants predisposing to the development of HCC in individuals with MASLD. It enrolled European patients with MASLD-HCC (n=208) and two groups of controls: patients with MASLD and advanced fibrosis without HCC (n=414), individuals with MASLD without advanced fibrosis (n=107) as well as locally and ethnically matched healthy individuals (n=152). The diagnosis of NAFLD was based on the demonstration of fatty liver by imaging at the time of study inclusion or a previous positive clinical history in patients with advanced disease [15], daily alcohol intake <30/20 g/day in males/females, and absence of concurrent liver diseases and other hepatotoxic factors (including chronic viral or autoimmune hepatitis; genetic liver diseases including hereditary hemochromatosis, Wilson's disease, AAT deficiency, use of steatogenic/hepatotoxic drugs. All patients fulfilled the metabolic criteria for MASLD. Advanced liver fibrosis and HCC were diagnosed according to the EASL criteria [16, 17]. Age, sex, presence of type 2 diabetes (T2D), advanced liver

fibrosis, AST and ALT levels at the time of study enrolment were available in all patients. In addition, the main common germline risk variants for HCC in MASLD and a polygenic risk score (PRS) summarizing their effect were available for all [18]. Part of this cohort has previously been described contributing to the identification of new genetic risk loci for HCC [8, 19].
Blood samples to evaluate the presence of CHIP were collected at the time of study inclusion. The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki. The EPIDEMIC and SERENA study were approved by the Ethical Committee of the Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milan and participating centers (EPIDEMIC-TERT study Ethical approval n. 1882_2013; Perspective-SERENA multicenter Study approval n. 485_2017, Fondazione IRCCS Ca' Granda Ethical Committee). Informed consent was obtained from each participant. The FLIP study was approved by the Newcastle Hepatopancreatobiliary Research Tissue Bank (REC reference: 10/H0906/41), with patient samples and data shared from the Newcastle Academic Health Partners Bioresource.
The clinical features of the study cohorts are presented in Table 1.

Next generation sequencing and CHIP analysis
DNA sequencing was performed as previously described [8]. Briefly, DNA libraries were enriched for exome sequencing by the SureSelect Human All Exon v5/7 kits (Agilent, Cernusco sul Naviglio, Milan, Italy). Sequencing was subsequently performed on the HiSeq 4000/NextSeq2000 platforms (Illumina). Raw reads quality control was performed using FastQC software (Brabaham bioinformatics, Cambridge, UK). Reads mapping on the human GRCh37 genome was performed using the MEM algorithm of Burrows Wheeler Aligner (BWA) version 0.7.10 [20]. Reads with low quality alignments and duplicate reads were filtered out using Samtools to generate high quality bam files [21]. Mapping quality control was performed using Picard-tools (http://broadinstitute.github.io/picard) and Bedtools [22]. Sequencing mean depth was 73x, and no samples exhibit a mean depth lower than 50x. Somatic mutations were identified accepting a minimum variant coverage of 20, a minimum alternative allele count of 3 and a VAF between 0.02 and 0.46 in concordance with established variant filtering algorithms, the only exception was the *JAK2* V617F mutation, for which no VAF limitation was set [23][24]. Variants with a population mean allelic frequency higher than 1 in a 1000 were treated as polymorphisms. Variants were then curated to include only variants known to be somatic and associated to either malignancy or CHIP with the use of a semi-automatic pipeline [24][25][26]. Genes evaluated for CHIP attribution included: *NADK, GNB1, CBL, KMT2D, RHEBL1, PPM1D, JMJD6, MFSD11, DNMT3A, SF3B1, ASXL1, NOL4L, GNAS, RUNX1, AF015262.1, RPL34P3, U2AF1, TET2, CUX1, SH2B2, BCOR, BCORL1, MCAM, TP53, METTL23, SRSF2, NPM1P46, LINC01426, EZH2P1, ELF4, ATP1B2, ZNF316, ANAPC1, JAK2*. Three variant callers were employed for variant identification: Mutect, Vardict and Freebayes [27,28,29]; only variants recognized by at least 2 variant callers were then considered for further analysis.

Study design
The principal study aim was to assess the impact of CHIP and CHIP not driven by *DNMT3A* variants on the main outcome, that is the risk of developing HCC in the cross-sectional EPIDEMIC cohort. The choice of selecting non-DNMT3A is based on recent evidence suggesting that only selected DNMT3A mutations confer an inflammatory phenotype [27]. We considered unadjusted analyses, adjusted for main confounders available in the cohort (sex, presence of T2D and advanced fibrosis, genetic predisposition due to carriage of common germline mutations as summarized by PRS-5, Polygenic Risk Score 5) excluding

age (in the hypothesis that CHIP may be involved in mediating the effect of age on HCC risk), and with further adjustment by age. As secondary aims, to investigate the possible mechanism underlying the epidemiological association, we tested the impact of CHIP on liver damage: AST and ALT, correlating with inflammation and fibrosis and liver fat, respectively, and the presence of advanced fibrosis, and whether CHIP defined by VAF≥10% was more robustly associated with the study outcomes.
Finally, to explore whether genetic lesions in the most frequently mutated genes may specifically be associated with HCC, we also tested the impact of mutations at specific genes defining CHIP on the main outcome.

Statistical analysis
For descriptive statistics, categorical variables are shown as number and proportion, while continuous variables are shown as mean and standard deviation (SD) or median and interquartile range (IQR), as appropriate. The association of CHIP with AST and ALT levels was assessed by multivariable generalized linear models, whereas the association with cirrhosis and HCC through multivariable logistic regression models. As one of the main study aims was to examine the impact of CHIP on HCC, we included as covariates in multivariable models, sex, presence of T2D, advanced fibrosis and PRS-5. Analyses were then further adjusted for age.
Statistical analysis was carried out using the JMP 16.0 Pro Statistical Analysis Software (SAS Institute, Cary, NC), and R statistical analysis software version 4.1 (http://www.R-project.org). *P* values <0.05 (two tailed) were considered significant. P values were corrected with a *false discovery rate* approach, where appropriate.

RESULTS

Study cohort
The cross-sectional NAFLD-EPIDEMIC study cohort comprised 881 individuals: 208 patients had MASLD-HCC, 414 MASLD and advanced liver fibrosis without HCC, and 259 controls without advanced fibrosis (Table 1).
There was a progressive increase in age (p<0.001), AST levels (p<0.001), and prevalence of cirrhosis (p<0.001) from non-advanced fibrosis to advanced fibrosis and HCC. Patients with advanced fibrosis and HCC had higher BMI (p<0.001), prevalence of T2D (p<0.001), ALT levels (p<0.001), and prevalence of *PNPLA3* p.I148M variant homozygosity (p<0.001), and higher PRS-5 score (p=0.007) than those without advanced fibrosis. Males were overrepresented in all three cohort subgroups (63.3%, 54.6% and 79.8% respectively).

Prevalence of CHIP
CHIP-defining genetic lesions were identified in 116 out of 881 participants (13.1%). CHIP was found in 51 (24.5%) of patients with HCC, and among those without HCC in 54 (13.0%) of those with advanced fibrosis, and in 11 (4.2%) of those without advanced fibrosis (Table 1; p<0.001). As expected, the prevalence of CHIP was age-dependent (OR 2.02 per year, 95% CI 1.05-3.87; p<0.0001) and increased consistently across diagnosis groups, with the spike in CHIP incidence being observed after the age of 60 (Figure 1).
The oncoplot showing the frequency of specific CHIP-defining genetic lesions is shown in Figure 2. Across the cohort, 116 (13.1%) patients showed at least one CHIP-defining mutation. 43 (37%) carried at least an additional CHIP lesion. The most frequently mutated gene was *DNMT3A*, followed by *TET2*, *TP53* and *ASXL1* (Table 1). The spectrum of mutations was as expected, with 7.8% (4/51) of *DNMT3A* mutations occurring at codon R882, 40% (8/20) of *TET2* mutations being truncating, and 90% (9/10) of *ASXL1* mutations were in exon 13, whereas JAK2 lesions included the V617F lesion only. Overall, the median

VAF was 5.6% (range min-max 2-86%), with 31% (36/116) of patients showing at least one variant with VAF ≥10%. CHIP with an allele burden ≥10% was observed predominantly in the advanced fibrosis (20/36) and in the HCC groups (14/36). *TET2* and *TP53* variants were enriched in patients with HCC (p<0.001, p=0.022 respectively), while prevalence of *DNMT3A* was not significantly different between the advanced fibrosis and the HCC group. The *JAK2* V617F mutation appeared in three patients, one with HCC carrying a lesion with a high VAF (>80%) and in two cases of patients with non-fibrotic liver disease with a VAF of 33 and 7%, respectively.

Impact of CHIP on disease severity

We first sought to assess the impact of CHIP on liver disease severity, by analyzing association of mutations with biomarkers of liver damage. Although at univariate analysis in the overall cohort no significant association was found between AST or ALT levels and the presence of CHIP (p>0.05), after correcting for the main clinical confounders, namely, age, sex, T2D and BMI, CHIP cases with VAF≥10% were associated with higher ALT (OR 1.1, CI 95% 1.002-1.2) and AST levels (OR 1.1, CI 95% 1.01-1.2) (Fig. 3A). CHIP regardless of allele frequency, in the absence of *DNMT3A* mutations was associated with higher AST levels (OR 1.1, CI 95% 1.02-1.2) but not with ALT suggesting a sequenced relation with liver enzyme elevation.

Even higher levels of both AST and ALT were observed in patients with CHIP defined by mutations other than *DNMT3A* and VAF≥10% when corrected for the aforementioned variables (OR 1.20, CI 95% 1.10-1.30 and OR 1.21, CI 95% 1.10-1.35, respectively, Fig.3).

The association of CHIP on the risk of cirrhosis is presented in Table 2. At univariate logistic regression analysis, CHIP was associated with cirrhosis (OR 2.64, 95% CI 1.73-4.13, p<0.001). At multivariate analysis, the association of CHIP with cirrhosis remained independent of sex, T2D, and inherited genetic predisposition to SLD as captured by the PRS-5 (OR 1.70, 95% CI 1.10-2.73, p=0.02). The impact of CHIP appeared more consistent in men (OR 1.85, 95% CI 1.06-3.37, p=0.035) than in women (OR 1.41, 95% CI 0.66-3.27, p=0.38). However, the association between CHIP and cirrhosis was attenuated and lost statistical significance after correction for age (OR 1.26, 95% CI 0.78-2.10, p=0.35). Neither CHIP with VAF≥10% nor specific genetic lesions were associated with cirrhosis independently of age (not shown).

Impact of CHIP on HCC risk

The impact of CHIP on the risk of HCC is presented in Table 3. Logistic regression highlighted a significant association between CHIP and HCC (OR 3.04, 95% CI 2.01-4.55, p<0.001). Remarkably, the association remained significant after correction for sex, T2D, and PRS-5 (OR 2.21, 95% CI 1.42-3.41, p<0.001), and even after further correction for the presence of cirrhosis (OR 2.01, 95% CI 1.30-3.15, p=0.002), suggesting it is independent of the stage of liver disease. In this model, the association of CHIP with HCC was significant in men (OR 2.81, 95% CI 1.65-4.83, p<0.001), but not in women (OR 0.71, 95% CI 0.23-1.83, p=0.5).

Despite a relevant proportion of patients with HCC were found to carry CHIP, the association of HCC and CHIP was not statistically significant when correcting for age (p>0.05), due to the older median age in patients with HCC. However, CHIP associated with HCC even in the context of age correction when *DNMT3A* lesions were excluded from the analysis (OR 2.45, CI 95% 1.34-4.53, p=0.02).

Finally, we examined the impact of specific CHIP genetic lesions on the risk of HCC. We highlighted an association between *TET2* and *TP53* mutations with HCC (Table 4). Strikingly, 70% of *TET2* mutations were observed in HCC cases. On the contrary, *DNMT3A*-

driven CHIP, despite being the most represented event, was equally distributed with respect to HCC prevalence. The association between HCC and *TET2* remained statistically significant at multivariable logistic regression when correcting for covariates including age, sex, T2D, presence of cirrhosis and PRS-5 score (OR 4.80, 95% CI 1.24-18.1.60-17.0, p=0.02; Table 4 and Figure 4). Notably, 25% of *TET2* mutated HCC cases had VAF≥ 10% with a trend for higher AST levels with respect to the low allelic burden counterpart (median AST 84 vs 40 IU/l), reinforcing the concept that a higher fraction of CHIP circulating myeloid elements associates with a higher inflammatory burden.

DISCUSSION

Here we report the prevalence of CHIP in a multicenter cross-sectional cohort of patients with severe MASLD and controls without advanced liver fibrosis and describe the association of CHIP with liver damage and HCC development, with the latter being the main study focus. Indeed, recent evidence suggests that the occurrence of somatic mutations in the liver and hematopoietic cells accompany and contribute to the progression of MASLD to steatohepatitis and fibrosis [12, 13]. Importantly, clonal mutations defining CHIP, and particularly *TET2* mutations, have been shown to promote liver disease progression by inducing a proinflammatory phenotype in myeloid cells homing to the liver [13]. However, despite inflammation being involved in hepatic carcinogenesis, no data were yet available on the impact of CHIP on HCC onset.

First, we observed that 13% of the cohort population showed one or more genetic mutations defining CHIP, this being in line with the expected proportion of affected genes and types of mutations [30], with a prevalence distribution sharply rising at the age of 60 across all study groups.

Secondly, CHIP-bearing patients, especially those with a higher clonal burden (≥10%), showed levels of AST higher than age and comorbidity-matched peers. The contribution of larger CHIP clones on liver inflammatory markers has been previously described in larger cohorts [13]. Whether this finding is the result of higher circulating CHIP clone fraction contributing more consistently to organ inflammation is not currently known. In the present cohort, elevated AST and less strictly ALT levels correlated with CHIP, especially when *DNMT3A* driven cases were not considered. This observation is in line with previous evidence that *DNMT3A* mutations had a lower impact on the risk of progressive MASLD than those in *TET2* and *ASXL1* [13] and might subtend a different inflammatory potential of circulating CHIP clones according to their driving genetic lesion. In addition, the inflammatory potential of *DNMT3A* mutations seems to be lesion specific. In our case, only one patient carried the R882H variant, that has been associated with the higher inflammatory burden by recent studies [31]. Hence, we do not exclude a priori a role for *DNMT3A* in liver pathology, but if a role is present, it could be mutation-specific. Evaluation of *DNMT3A* mutations in larger cohorts is required to define their role in liver disease. Individuals positive for CHIP were at almost 2-fold higher risk of being affected by cirrhosis independently of the main clinical covariates, such as sex and presence of T2D, and genetic predisposition towards progressive SLD. However, due to the strong link between CHIP and aging, the association was lost after correcting for age at enrolment. Therefore, we could not conclude whether CHIP may partially mediate the effect of aging on progressive liver disease or was an epiphenomenon.

The main study finding, however, was the evidence of an association between CHIP and the occurrence of HCC in the NAFLD-EPIDEMIC cohort. Importantly, as a whole, CHIP was associated with a 2-fold increase in the risk of HCC when corrected for other contributing

factors, including the presence of cirrhosis. Notably, CHIP was able to discriminate between patients who developed HCC and those who did not better than genetic predisposition intended as inheritance of common risk variants predisposing to severe SLD, which are associated with progression to cirrhosis. However, the association between CHIP and HCC was not independent of age when considering DNMT3A lesions, but it did so when DNMT3A lesions were excluded. This finding could be explained by the strong collinearity between development of CHIP and HCC with increasing age, but also by a mutation specific effect, with most *DNMT3A* variants described in our series being neutral.

Notably, a strikingly clear signal has emerged showing to our knowledge for the first time, a specific enrichment of *TET2* and *TP53* lesions in HCC samples. The association of *TET2* genetic lesions (and of non-*DNMT3A*-CHIP overall) with HCC remained significant even after adjusting for all covariates, including clinical and genetic factors as well as age at enrolment. The risk of developing HCC was found to be approximately 4.8 times higher in individuals with *TET2*-CHIP. This observation supports a diverse contribution to inflammation and organ damage from circulating myeloid cells bearing different CHIP lesions. Indeed, it has been suggested that different CHIP clones engage selected immune phenotypes, with *DNMT3A* driven clonal hematopoiesis being associated with increased IFN activity and *TET2*-driven hematopoiesis with an NLRP3 dependent IL-1B production and elevated IL-6 (Interleukin 6) and IL-8 (Interleukin 8) levels. Supporting these findings, wild-type *TET2* gene seems required for suppression of IL-6 production [31, 32, 33]. While IL-6 activity has been reported to be higher also in *DNMT3A* driven CHIP, this is the case for R882H mutations only, further corroborating our hypothesis [31].

Overall, these data suggest that specific genetic lesions defining CHIP, and definitively those in *TET2*, may play a causal role in hepatic carcinogenesis by inducing liver inflammation and be at least partly involved in mediating the impact of aging on the risk of HCC development. Current limitations include the lack of an independent validation cohort, inclusive of different ethnicities, with a prospective approach to assess the impact of CHIP incidence on HCC risk. The contribution of CHIP to HCC may be confounded by the strong age-dependency of CHIP, with age being a determinant of HCC risk as well. In our study, non-DNMT3A CHIP was associated with HCC even after adjusting for age in a multivariate model. However, only larger prospective cohort studies are likely to determine whether the impact of CHIP to HCC risk is substantial enough to become a clinically useful parameter. Not the least, there is a need for mechanistic demonstration of the impact of specific CHIP-defining genetic lesions on hepatic carcinogenesis.

In conclusion, in this cross-sectional multicenter case-control cohort of European patients with MASLD we observed a high prevalence of CHIP, which was associated with a more severe liver damage and a 2-fold higher risk of HCC. HCC association with CHIP was independent of clinical and genetic cofactors, and of age when considering non-DNMT3A lesions. In addition, when analyzing the specific drivers, a significant four-fold impact of *TET2* driven CHIP on HCC was observed, independent of age. However, given the relatively small size of our cohort and its retrospective nature, additional studies will be crucial in validating our findings. External validation with a prospective cohort is required to fully establish CHIP as a risk factor for liver disease progression and untangle the interaction between aging and clonal hematopoiesis. Furthermore, experimental and in vivo studies are needed to determine a pathogenic contribution of tumor infiltrating leukocytes in the inflammatory milieu of HCC.

ABBREVIATIONS

Clonal Hematopoiesis of Indeterminate Potential (CHIP), Hepatocellular Carcinoma (HCC), Polygenic Risk Score 5 (PRS-5), Aspartate Aminotransferase (AST), Alanine

Aminotransferase (ALT), Next Generation Sequencing (NGS), Non-Alcoholic Fatty Liver Disease (NAFLD), Fatty Liver Disease (FLD), Interferon (IFN), Interleukin 6 (IL6), Interleukin 8 (IL8), Variant Allele Frequency (VAF), Type 2 Diabetes (T2D), Body Mass Index (BMI), Odds Ratio (OR), Confidence Interval (CI), metabolic dysfunction associated fatty liver disease (MAFLD).

ACKNOWLEDGMENTS
We thank Rossana Carpani for administrative support, Oriana Lo Maglio for graphical support.

REFERENCES

[1] Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 2016;64:73-84.

[2] European Association for the Study of the L, European Association for the Study of D, European Association for the Study of O. EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. J Hepatol 2016;64:1388-1402.

[3] Boursier J, Tsochatzis EA. Case-finding strategies in non-alcoholic fatty liver disease. JHEP Rep 2021;3:100219.

[4] Huang DQ, Singal AG, Kono Y, Tan DJH, El-Serag HB, Loomba R. Changing global epidemiology of liver cancer from 2010 to 2019: NASH is the fastest growing cause of liver cancer. Cell Metab 2022;34:969-977 e962.

[5] Tan DJH, Setiawan VW, Ng CH, Lim WH, Muthiah MD, Tan EX, et al. Global burden of liver cancer in males and females: Changing etiological basis and the growing contribution of NASH. Hepatology 2023;77:1150-1163.

[6] Valenti L, Bugianesi E, Pajvani U, Targher G. Nonalcoholic fatty liver disease: cause or consequence of type 2 diabetes? Liver Int 2016;36:1563-1579.

[7] Trepo E, Valenti L. Update on NAFLD genetics: From new variants to the clinic. J Hepatol 2020;72:1196-1209.

[8] Baselli GA, Jamialahmadi O, Pelusi S, Ciociola E, Malvestiti F, Saracino M, et al. Rare ATG7 genetic variants predispose patients to severe fatty liver disease. J Hepatol 2022;77:596-606.

[9] Stender S, Kozlitina J, Nordestgaard BG, Tybjaerg-Hansen A, Hobbs HH, Cohen JC. Adiposity amplifies the genetic risk of fatty liver disease conferred by multiple loci. Nat Genet 2017;49:842-847.

[10] Tilg H, Moschen AR. Evolution of inflammation in nonalcoholic fatty liver disease: the multiple parallel hits hypothesis. Hepatology 2010;52:1836-1846.

[11] Tacke F, Puengel T, Loomba R, Friedman SL. An integrated view of anti-inflammatory and antifibrotic targets for the treatment of NASH. J Hepatol 2023.

[12] Ng SWK, Rouhani FJ, Brunner SF, Brzozowska N, Aitken SJ, Yang M, et al. Convergent somatic mutations in metabolism genes in chronic liver disease. Nature 2021;598:473-478.

[13] Wong WJ, Emdin C, Bick AG, Zekavat SM, Niroula A, Pirruccello JP, et al. Clonal haematopoiesis and risk of chronic liver disease. Nature 2023;616:747-754.

[14] Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. Science 2019;366.

[15] Valenti L, Romeo S, Pajvani U. A genetic hypothesis for burnt-out steatohepatitis. Liver Int 2021;41:2816-2818.

[16] EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2012;56:908-943.

[17] European Association for the Study of the Liver. Electronic address eee, Clinical Practice Guideline P, Chair, representative EGB, Panel m. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis - 2021 update. J Hepatol 2021;75:659-689.

[18] Bianco C, Jamialahmadi O, Pelusi S, Baselli G, Dongiovanni P, Zanoni I, et al. Non-invasive stratification of hepatocellular carcinoma risk in non-alcoholic fatty liver using polygenic risk scores. J Hepatol 2021;74:775-782.

[19] Pelusi S, Baselli G, Pietrelli A, Dongiovanni P, Donati B, McCain MV, et al. Rare Pathogenic Variants Predispose to Hepatocellular Carcinoma in Nonalcoholic Fatty Liver Disease. Sci Rep 2019;9:3682.
[20] Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 2010;26:589-595.
[21] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-2079.
[22] Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841-842.
[23] Vlasschaert C., Mack T., Heimlich J. B., Niroula A., Uddin M. et al. A practical approach to curate clonal hematopoiesis of indeterminate potential in human genetic data sets. Blood 2023; 141 (18): 2214–2223.
[24] Bolli N., Manes N., McKerrell T., Chi J., Park N., et al. Characterization of gene mutations and copy number changes in acute myeloid leukemia using a rapid target enrichment protocol. Haematologica 2015;100(2):214-222
[25] McKerrell T., Moreno T., Ponstingl H., Bolli N., M. L. Dias J. et al.; Development and validation of a comprehensive genomic diagnostic tool for myeloid malignancies. Blood 2016; 128 (1): e1–e9.
[26] Da Vià M., Lionetti M., Marella A., Matera A., Travaglino E. et al.; MGUS and clonal hematopoiesis show unrelated clinical and biological trajectories in an older population cohort. Blood Adv 2022; 6 (21): 5702–5706.
[27] Silver A. J, Brown D.J., Vlasschaert C., Bhat P., Puddu F., et al. DNMT3A R882H Exhibits Greater Inflammatory Potential Than R882C in Primary Hematopoietic Stem and Progenitor Cell Knock-in Model and Population Data, Blood 2023/10.1182/blood-2023-187244.
[28]Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D et al.; Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013 Mar;31(3):213-9. doi: 10.1038/nbt.2514.
[29] Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, et al.; VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res. 2016 Jun 20;44(11):e108. doi: 10.1093/nar/gkw227.
[30] Garrison, Erik, and Gabor Marth. "Haplotype-based variant detection from short-read sequencing." arXiv preprint arXiv:1207.3907 (2012).
[31] Niroula A, Sekar A, Murakami MA, Trinder M, Agrawal M, Distinction of lymphoid and myeloid clonal hematopoiesis. Nat Med. 2021 Nov;27(11):1921-1927. doi: 10.1038/s41591-021-01521-4.
[32] [Belizaire, R., Wong, W.J., Robinette, M.L. et al. Clonal haematopoiesis and dysregulation of the immune system. Nat Rev Immunol (2023). https://doi.org/10.1038/s41577-023-00843-3
[33] Zhang Q, Zhao K, Shen Q, Han Y, Gu Y, Li X, Zhao D, Liu Y, Wang C, Zhang X, Su X, Liu J, Ge W, Levine RL, Li N, Cao X. Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. Nature. 2015 Sep 17;525(7569):389-393. doi: 10.1038/nature15252.

TABLES

Table 1. Clinical features of the 881 individuals included in the NAFLD-EPIDEMIC (MASLD) cross-sectional study cohort.

	HCC	No HCC		
		Advanced Fibrosis	No Advanced Fibrosis	P*
n	208	414	259	
Age, years	70.0 (64.5, 76.0)	63.0 (56.0, 69.0)	51.0 (41.7, 60.2)	<0.001
Sex, male	166 (79.8)	226 (54.6)	164 (63.3)	<0.001
T2D, yes	111 (53.4)	150 (36.4)	4 (1.5)	<0.001
BMI, Kg/m^2	29.1 (26.1, 33.0)	29.8 (26.9, 34.0)	23.9 (22.0, 25.61)	<0.001
AST, IU/l	44.50 (33.0, 58.3)	37.0 (27.0, 49.0)	22.0 (19.0, 29.0)	<0.001
ALT, IU/l	37.0 (27.0, 52.0)	38.0 (28.0, 55.0)	19.0 (15.0, 27.0)	<0.001
Cirrhosis, yes	148 (82.7)	176 (66.9)	0	<0.001
PNPLA3, p.148M/M	58 (28.0)	135 (32.7)	13 (12.0)	<0.001
PRS-5, score	0.44 (0.22-0.66)	0.40 (0.25, 0.60)	0.33 (0.13-0.49)	0.007
CHIP, yes	51 (24.5)	54 (13.0)	11 (4.2)	<0.001
CHIP w VAF≥10%	14 (6.7)	20 (4.8)	2 (0.8)	ns
DNMT3A, yes	16 (7.7)	32 (7.7)	3 (1.2)	0.1
TET2, yes	14 (6.7)	4 (1.0)	2 (0.8)	0.001
TP53, yes	8 (3.8)	5 (1.2)	2 (0.8)	0.06
ASXL1, yes	6 (2.9)	4 (1.0)	0	ns
JAK2, yes	1 (0.5)	0	2 (0.7)	ns

Data are shown as N (%), or median (IQR), when appropriate. HCC: hepatocellular carcinoma; BMI: body mass index; T2D: type 2 diabetes; HCC: hepatocellular carcinoma; LSM: liver stiffness measurement by Fibroscan (available in 28% of cases). *P values were calculated among pairs through Kruskal-Wallis test for continuous variables (non-normality assumed) and Fisher test for categorical variables and corrected for multiple testing through *false discovery rate*. *ns:* not statistically significant (p>0.05).

Table 2. Independent determinants of cirrhosis (n=474) in 881 individuals in the NAFLD-EPIDEMIC (MASLD) study cohort.

	Unadjusted		Model 1		Model 2	
	OR, 95% CI	p value	OR, 95% CI	p value	OR, 95% CI	p value
Age, years	1.10, 1.08-1.11	<0.0001	-	-	1.06, 1.04-1.08	<0.0001
Sex, M	1.03, 0.78-1.35	0.81	0.92, 0.66-1.28	0.63	1.00, 0.71-1.42	0.96
T2D, yes	3.71, 2.71-5.15	<0.0001	1.98, 1.41-2.80	<0.0001	1.45, 1.01-2.10	<0.041
PRS-5, unit	4.2, 2.25-8.00	<0.001	4.27, 2.25-8.24	<0.001	3.75, 1.90-7.54	0.0001
CHIP, yes	2.64, 1.72-4.14	<0.001	1.70, 1.07-2.73	0.02	1.26, 0.77-2.10	0.35

At logistic regression analysis adjusted for the covariates reported in the models. Unadjusted; Model 1: adjusted for sex, type 2 diabetes (T2D), polygenic risk score of fatty liver disease (PRS-5); Model 2: further adjusted for age at enrolment. CHIP: clonal hematopoiesis of indeterminate potential.

Table 3. Independent determinants of hepatocellular carcinoma (HCC, n=208) in 881 individuals in the NAFLD-EPIDEMIC (MASLD) study cohort.

	Unadjusted		Model 1		Model 2		Model 3		Model 4	
	OR, 95% CI	p value	OR, 95% CI	p value	OR, 95% CI	p value	OR, 95% CI	p value	OR, 95% CI	P value
Age, years	1.11, 1.10-1.13	<0.0001	-	-	-	-	1.10, 1.07-1.12	<0.0001	1.10, 1.07,1.12	<0.001
Cirrhosis, yes	6.35, 4.31-9.60	<0.0001	-	-	3.50, 2.30-5.43	<0.0001	3.04, 1.85-5.03	<0.0003	2.32, 1.46,3.73	<0.001
Sex, M	2.87, 2.00-4.20	<0.0001	2.90, 2.00-4.34	<0.0001	3.10, 2.10-4.70	<0.0001	3.80, 2.50-5.90	<0.0001	3.82, 2.51-5.92	<0.001
T2D, yes	3.84, 2.80-5.33	<0.001	2.62, 1.86-3.70	<0.0001	2.33, 1.64-3.33	<0.0001	1.97, 1.35-2.88	0.0004	1.98, 1.36-2.89	<0.001
PRS-5, unit	1.32, 0.70-2.50	0.38	1.44, 0.74-2.82	0.27	1.05, 0.52-2.12	0.88	0.90, 0.43-1.90	0.80		0.31
PRS-5, high risk (≥0.495)	0.76, 0.55-1.06	0.10							1.22, 0.83-1.80	
CHIP, yes	3.04, 2.02-4.56	<0.0001	2.08, 1.26-3.26	0.0004	2.00, 1.30-3.15	0.002	1.41, 0.86-2.30	0.17	1.46, 0.89-2.40	0.13

At logistic regression analysis adjusted for the covariates reported in the models. Unadjusted; Model 1: adjusted for sex, type 2 diabetes (T2D), polygenic risk score of fatty liver disease (PRS-5); Model 2: further adjusted for cirrhosis; Model 3: further adjusted for age at enrolment. CHIP: clonal hematopoiesis of indeterminate potential. Model 4 evaluated PRS-5 as a nominal variable, with high risk PRS being ≥0.495.

Table 4. Impact of genetic lesions at specific genes defining CHIP on the risk of hepatocellular carcinoma (HCC, n=179) in 530 individuals in the NAFLD-EPIDEMIC study cohort.

	Model 1		Model 2	
	OR, 95% CI	p value	OR, 95% CI	p value
CHIP	3.04, 2.01-4.60	<0.001	1.41, 0.86-2.31	0.24
CHIP without *DNMT3A*	4.47, 2.72-7.42	<0.001	2.45, 1.35-4.53	0.02
DNM3TA	1.51, 0.80-2.76	0.18	0.60, 0.27-1.22	0.24
TET2	8.02, 3.17-22.90	0.002	4.8, 1.60-17.0	0.02
TP53	3.81, 1.35-11.00	0.045	1.93, 0.50-7.92	0.37
ASXL1	5.00, 1.41-19.60	0.015	2.00, 0.45-10.30	0.37

At logistic regression analysis adjusted for the covariates reported in the models. Unadjusted; Model 1: adjusted for CHIP defining genetic lesions; Model 2: further adjusted for age at enrolment, sex, T2D, cirrhosis and PRS-5 score. A false discovery rate approach was employed to account for multiple testing.

Figure 1. Clonal Hematopoiesis of Indeterminate Potential (CHIP) prevalence as a function of age ($p<0.0001$). A steep expected increase is seen after 60 years of age and appears to be consistent across patient groups. The cumulative prevalence of CHIP in the overall study cohort settled around 13.1%. Figure detail: cumulative prevalence of DNMT3A and TET2 mutations showing divergent age dependency in the HCC cohort with TET2 plateauing after 75 years of age.

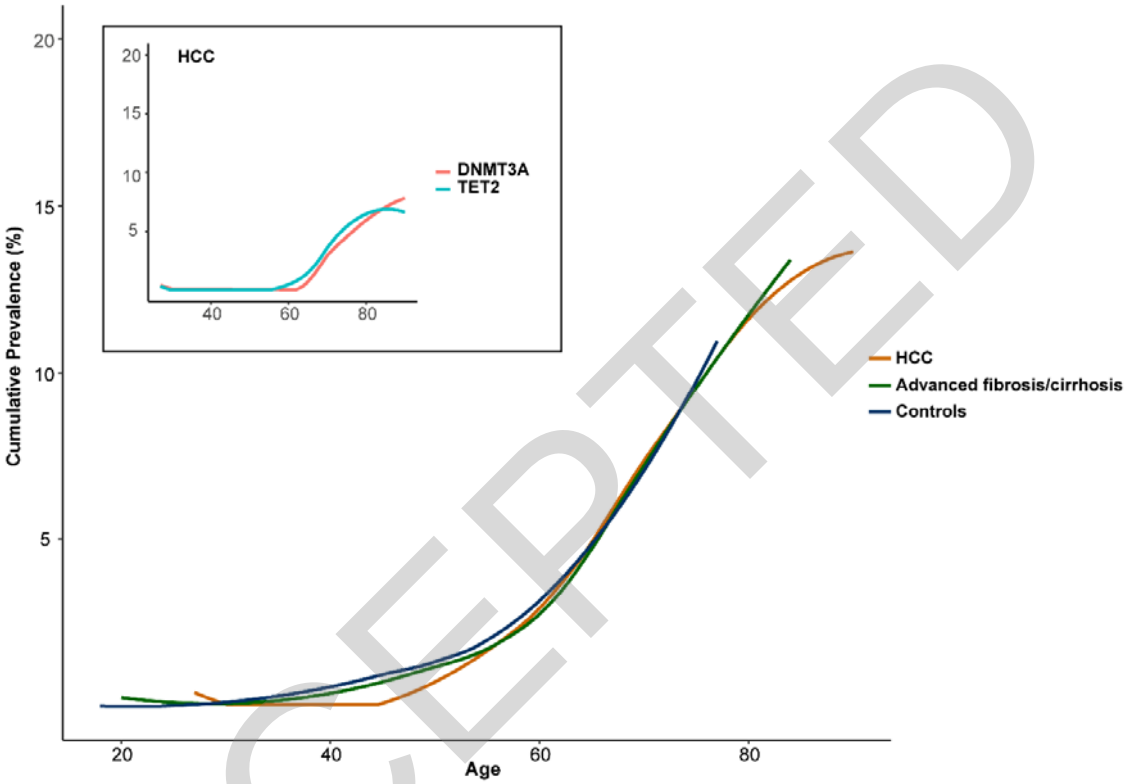

Figure 2. Oncoplot showing Clonal Hematopoiesis of Indeterminate Potential (CHIP) defining lesions across patient cohorts. Genes are ordered according to cohort frequency. Variants are color coded according to legend.
Median lesion Variant Allele Frequency (VAF) is shown in the left barplot. Bottom ribbons identify disease groups, presence or absence of cirrhosis and hepatocellular carcinoma (HCC).

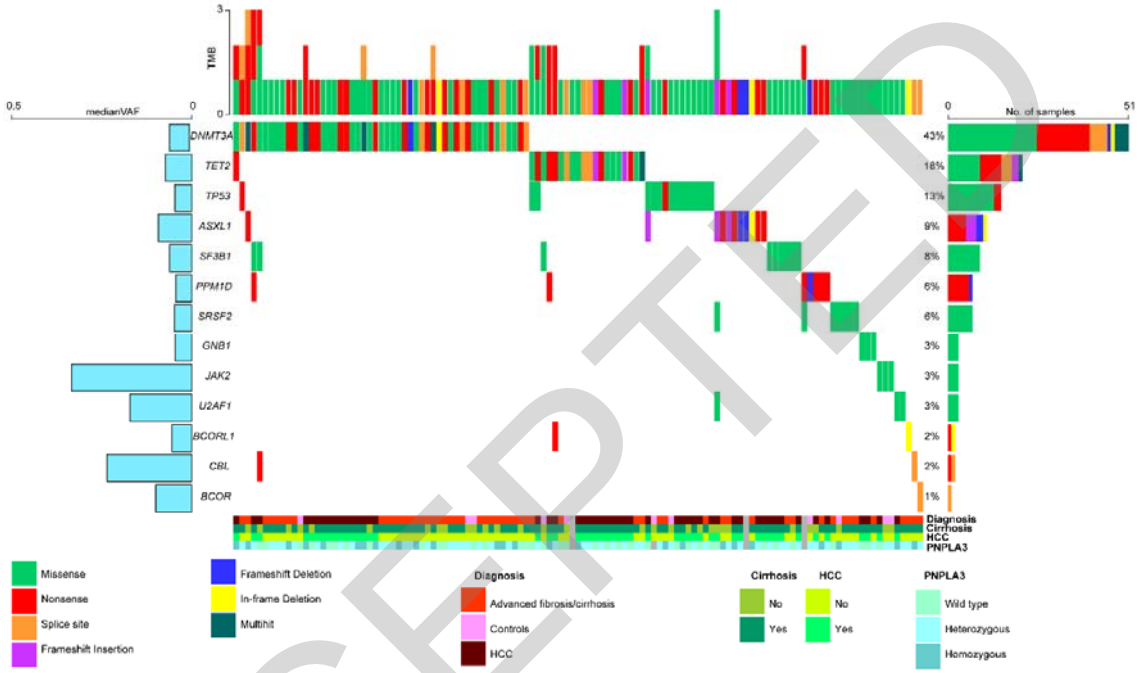

Figure 3: Violin plot displaying AST and ALT levels (log10(IU/L)), as biomarkers of liver fat and AST of liver inflammation, respectively, between samples with detectable Clonal Hematopoiesis (CHIP; in the absence of *DNMT3A* mutations), CHIP with an allele fraction ≥10%, and non-mutated samples. Additional stratification according to cirrhotic status is shown. Statistically significant differences are marked by an asterisk ($p<0.05$, at generalized linear models adjusted for Type 2 Diabetes (T2D), Body Mass Index (BMI), age and gender.

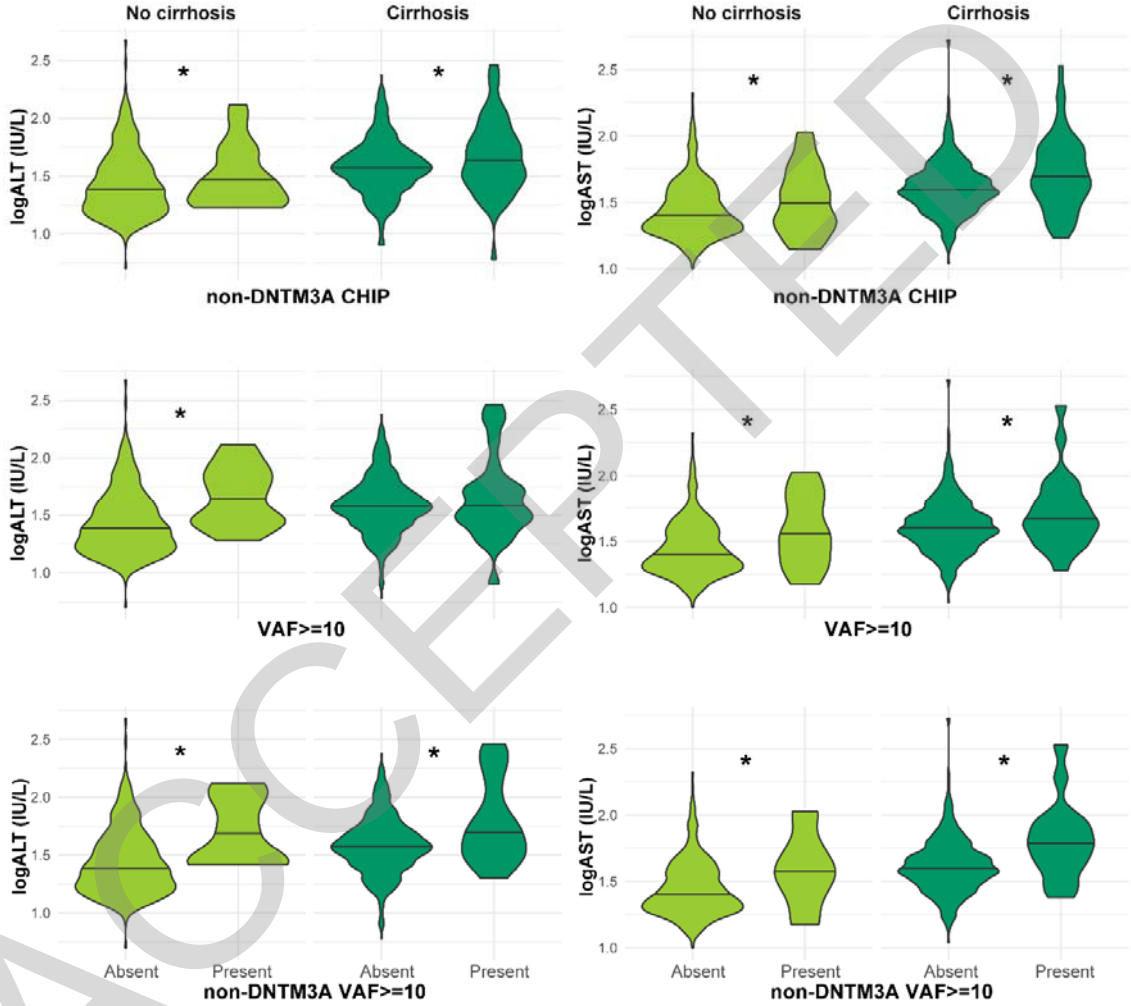

Figure 4. Forest plot displaying association between Hepatocellular Carcinoma (HCC), CHIP and CHIP subgroups Multivariable at analysis (MVA) including T2D, age at study enroll, cirrhosis, sex and PRS-5 score as covariates. Absolute counts of patients with concomitant HCC and CHIP and HCC without CHIP are shown in the first two columns. Odds ratio and their 95% confidence interval associating CHIP subtypes to HCC are shown in the third column. Rows represent different CHIP subtypes.

	CHIP+HCC (n)	CHIP without HCC (n)	OR [95% CI]
CHIP	51	65	1.41 [0.86-2.31]
CHIP without DNMT3A	38	32	2.45 [1.35-4.53]
DNMT3A	16	35	0.59 [0.27-1.22]
TET2	14	6	4.76 [1.60-17.0]
TP53	8	7	1.93 [0.50-7.92]
ASXL1	8	4	1.98 [0.45-10.0]