UCLA
UCLA Electronic Theses and Dissertations

Title

From Bedside to Bench-side: the Clinical, Epidemiological and Molecular Basis for Nonalcoholic Steatohepatitis and Hepatocellular Carcinoma

Permalink

https://escholarship.org/uc/item/6zk519z2

Author

Benhammou, Jihane N

Publication Date

2019

Supplemental Material

https://escholarship.org/uc/item/6zk519z2#supplemental

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

From Bedside to Bench-side: the Clinical, Epidemiological and Molecular Basis for

Nonalcoholic Steatohepatitis and Hepatocellular Carcinoma

A dissertation submitted in partial satisfaction of the

requirements for the degree of Doctor of Philosophy

in Molecular, Cellular and Integrative Physiology

by

Jihane Benhammou

2019

ABSTRACT OF THE DISSERTATION

From Bedside to Bench: the Clinical, Epidemiological and Molecular Basis for Nonalcoholic

Steatohepatitis and Hepatocellular Carcinoma

by

Jihane Benhammou

Doctor of Philosophy in Molecular, Cellular and Integrative Physiology

University of California, Los Angeles, 2019

Professor Joseph R Pisegna, Chair

Non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH) affect 75-100 million U.S. citizens and carries an increased risk for liver, cardiovascular and cancer related morbidity and mortality. Similarly, chronic hepatitis C virus (HCV) infection is a major cause of liver disease and hepatocellular carcinoma (HCC) worldwide. Understanding the clinical, epidemiology and biological causes of NAFLD, with or without HCV, is of utmost importance given the lack of targeted therapies and the large economic burden it places on healthcare. Accordingly, we aimed to identify the clinical and epidemiological factors that affect HCV treatment in the setting of NAFLD and understand the risk of HCC in the NAFLD patients, compared to viral etiologies of HCC. To do so, we utilized the Electronic Medical Records of the Veterans Affairs Health Care System (VA HCS), the largest single-payer health system in the U.S., and UCLA Medical Center, one of the largest tertiary-care liver transplantation centers in the country. To further understand the molecular basis of NAFLD and NASH, we studied liver RNA-sequencing from a cohort of bariatric surgery patients with detailed liver histopathological data.

We found that NAFLD does not affect HCV cure and that resolution of HCV leads to improvements in insulin resistance. We also observe that NAFLD HCC can occur in a non-cirrhosis background in 18% of cases and that older Hispanic patients with larger BMIs were more likely to have cirrhosis when diagnosed with NAFLD HCC. Through analysis of our transcriptomics human liver RNA-sequencing, we identify a lipid responsive non-coding gene, *OLMALINC*, as a novel enhancer RNA (eRNA) in the *cis* regulation of stearoyl Co-A desature, a key triglyceride gene that has been a therapeutic target in NASH human clinical trials. In this work, we present the clinical and epidemiological phenotypes of NAFLD and identify important associations between insulin resistance, dyslipidemia, and BMI in HCC. Our functional genomics data in statin-users help us identify the first eRNA in lipid metabolism described to date. Bridging the understanding of clinical phenotypes that translate to human-relevant molecular studies is key to elucidating the mechanisms of NAFLD, NASH and HCC.

The dissertation of Jihane Benhammou is approved.

Päivi Elisabeth Pajukanta

Thomas Aguilar Vallim

Stephen J Pandol

Gregory A Brent

Joseph R Pisegna, Committee Chair

University of California, Los Angeles

2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors. Follow this format for each person. DO NOT EXCEED FIVE PAGES.

NAME: Jihane Nejma Benhammou

eRA COMMONS USER NAME (credential, e.g., agency login): Jihane11

POSITION TITLE: UCLA Clinical Instructor and VA Health Care System staff physician

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)*

INSTITUTION AND LOCATION	DEGREE (*if applicable*)	MM/YY	FIELD OF STUDY
UT, TX	B.S.	06/05	Human Biology
UCSF San Francisco, CA	M.D.	06/11	Medicine
UCLA Los Angeles, CA	Residency	6/14	Internal Medicine
UCLA Los Angeles, CA	Fellowship	06/18	Gastroenterology
UCLA Los Angeles, CA	CTSI Certification	06/18	Clinical Research

Honors/Awards

2002-2005	National Institutes of Health Undergraduate Scholarship Program
2010	UCSF Dean's Prize in Research Finalist
2011, 2012	Commendation for Excellence in Medical Student Teaching, UCLA
2012	Commendation for Excellent in Patient Communication, UCLA
2013	AASLD Emerging Liver Scholar
2016-2017	Selected UCLA Gastroenterology Chief Fellow
2017	AGA Young Investigator Recognition at DDW
2017	Selected Compassionate Award for Gastroenterology Fellowship
2018	UCLA Center for Ulcerative Research (CURE) Abstract of Distinction
2018	American Society of Human Genetics (ASGH) Ribbon of Choice Abstract
2018	UCLA STAR Research Innovator Award
2018	UCLA DOM Third Place Research Award for Junior Faculty
2018	Selected "Future Leader in NASH" at the NASH-TAG Conference

Publications:

1. McCullen C, **Benhammou J**, Majdalani N, Gottesman S. Mechanism of positive regulation of DsrA and RprA sRNAs: pairing both increases translation and protects *rpoS* mRNA from degradation. J Bacteriol 2010 Nov; 192 (21): 5559-71. PMC2953674

2. Gottesman S, McCullen C, Guillier M, Vanderpool C, Majdalani N, **Benhammou J**, Thompson K, Fitzgerld P, Sowa N, FitzGerald D. Small RNA Regulators and the Bacterial Response to Stress. Cold Spring Harb Symp Quat Biol. 2006 Dec; 71: 1-11. PMC3592358

3. Benhammou J, Vocke C, Santani A, Schmidt L, Baba M, Wu X, Korolevich S, Seyama K,

Nathanson K, Stolle C, Linehan M. Identification of deletions and duplications in Birt-Hogg-Dubé Syndrome. *Gene Chromos Can* 2011 June; 50 (6): 466-77. PMC3075348

4. Eisenhofer G, Vocke C, Elkahloun A, Lenders J, Timmers H, **Benhammou J**, Linehan M, Pacak K. Genetic screening for von Hipper-Lindau gene mutations in non-syndromic pheochromocytoma: low prevalence and false-positives or misdiagnosis indicates a need for caution. *Horm Metab Res* 2012 May; 44 (5):343-8. PMC3501345

5. Benhammou J, Boris R, Pacak K, Pinto P, Linehan M, Bratslavsky G. Functional and oncological outcomes of of partial adrenalectomy for pheochromocytoma in VHL patients after at least a 5-year follow-up. *J Urol* 2010 Nov; 184 (5): 1855-9. PMC3164541

6. Boris R, **Benhammou J**, Merino M, Pinto P, Linehan M, Bratslavsky G. The impact of germline BHD mutation on histological concordance and clinical treatment of patients with bilateral renal masses and known unilateral oncocytoma. *J Urol* 2011 June; 185 (6): 205-5. PMC3164767

7.O'Mahony F, Wroblewki K, O'Byrne S, Jiang H, Clerkin K, **Benhammou J**, Blaner W, Beaven S. Liver X receptors balance lipid stores in hepatic stellate cells via Rab18, a retinoid responsive lipid droplet protein. *Hepatology* 2014 Dec; 62 (2): 615-26. PMC4458237

8. Ko A, **Benhammou JN**, Kaminska D, Nikkola E, Pisegna JR, Hui ST, Lusis AJ, Cantor RM, Sinsheimer JS, Mohlke KL, Laakso M, Pihlajamaki J and Pajukanta P. Obesity-induced reduction of Death Associated Protein Kinase 2 predisposes to non-alcoholic fatty liver disease. Manuscript under revisions.

9. **Benhammou JN**, Ko A, Alvarez M, Kaikkonen MU, Rankin C, Garske KM, Padua D, Bhagat Y, Kaminska D, Kärjä V, Pihlajamäki J, Pisegna JR and Pajukanta P. Novel lipid lincRNA *OLMALINC* regulates the liver steatosis gene, SCD, as an enhancer RNA. Manuscript under revisions.

10. **Benhammou JN**, Abby ES, Namagardi G, Manansala K, Tong MJ. Clinical characteristics and outcomes of nonalcoholic fatty liver disease associated hepatocellular carcinoma. Manuscript in progress.

11.Garske KM, Pan DZ, Miao Z, Bhaghat YV, Comenho C, Robles CR, **Benhammou JN**, Alvarez M, Ko A, Ye JC, Piesgna JR, Mohlke KL, Sinsheimer JS, Laakso M, Pajukanta P (in press). Saturated fat changes promoter accessibility in human adipocytes: Towards illuminating molecular basis of gene-environment interactions. *Nature Metabolism*.

12. Rankin CR, Treger J, Faure-Kumar E, **Benhammou J**, Anisman-Posner D, Pothoulakis C, Padua DM. A method to over-express long non-coding RNAs using gene-activating CRISPR. *J Vis Exp* 2019 1 (145).

13. Jacobs JP, Dong TS, Agopian V, Lagishetty V, Sundaram V, Noureddin M, Ayoub W, Durazo F, **Benhammou J**, Enyati P, Elashoff D, Goodman MT, Pisegna JP, Hussain S. Microbiome and bile acid profiles in duodenal aspirates from cirrhotics: the micribiome microbial markers and liver disease study. *Hepatol Res* 2019 48 (13): 1108-1117. PMC6334634

14. Dong T, Aby E, **Benhammou JN**, Kawamoto J, Han S, May FP, Pisegna JR. Metabolic syndrome does not affect sustained virological response of direct-acting antivirals while hepatitis C clearance improves hemoglobin A1c. *World J Hepatol* 2018 10 (9): 612-621. PMC6177571

15. **Benhammou JN**, Dong TS, May FP, Kawamoto J, Dixit R, Jackson S, Dixit V, Bhattacharya D, Han S and Pisegna JR. Race Affects SVR12 in a large and ethnically diverse hepatitis C infected patient population following treatment with direct acting antivirals: analysis of a single center department of veteran's affairs cohort. *Pharmacol Research and Perspects*, 2018 22 (6): e00379. PMC5821896

16. Aby E, Dong T, Kawamoto J, Pisegna JR and **Benhammou JN**. Chronic kidney disease progression predicts sustained virologic response to direct acting antivirals for hepatitis C treatment among United States Veterans. *World J Hepatol*, 2017 9 (36): 1352-1360. PMC5756725

Chapter 1

Non-alcoholic fatty liver disease and the metabolic syndrome in the era of chronic hepatitis C

within a VA patient cohort

Introduction:

Chronic hepatitis C (HCV) is one of the most common causes of chronic liver disease world-wide and is the most common blood-born infection in the United States [1]. Long-term infection can lead to complications including cirrhosis, hepatocellular carcinoma (HCC), and death [2]. Eradication of HCV has become a world-wide focus given that sustained virological response (SVR) has been associated with reversal of hepatic fibrosis and decreased rates of HCC [3,4]. New direct-acting antivirals (DAAs) have revolutionized HCV therapy given their ease of administration, tolerability and reported SVR at 12 weeks (SVR12), considered to be a cure, with rates in the 90s, depending on the extent of liver fibrosis and genotype [5].

Individuals with chronic HCV are more likely to develop type 2 diabetes (T2DM) and patients with T2DM have at least a 2-fold greater risk of developing HCV infection than the general population [6,7]. Studies have also shown that chronic HCV infection is associated with a greater risk for the development of insulin resistance [8]. In a retrospective analysis of cirrhotic patients, those with HCV infection were 10 times more likely to have T2DM than those without HCV infection [8]. There is also evidence that patients with chronic HCV infection and increased insulin resistance have a higher prevalence of hepatic fibrosis, HCC, and other extrahepatic manifestations [9-12]. While there are unclear mechanisms for increased insulin resistance among those with HCV, factors such as the metabolic syndrome, have been implicated [13,14]

Non-alcoholic fatty liver disease (NAFLD), the liver manifestation of the metabolic syndrome, is on the rise and is predicted to continue to increase along its complications, including HCC [15]. Thus, in the era of the NAFLD epidemic and new DAA treatments for chronic HCV, the impact of NAFLD and the metabolic syndrome on HCV response rates and cure warrants further study,

especially given the association of chronic HCV to insulin resistance and its association with progression of liver disease. The Veterans Affairs (VA) is the largest single-system U.S. health care provider and is one of the most diverse and longitudinally followed patient population. Within the VA, the incidence of chronic HCV is 2-3 times higher than the general public [16]. Additionally, patients that receive care in the VA also have a higher prevalence of obesity and T2DM compared to the general population [17,18]. Thus, the VA provides the ideal population to evaluate the relationships between chronic HCV, T2DM and NAFLD.

Materials and Methods:

This study was approved by the Institutional Review Board and Research and Development Committee at the VA Greater Los Angeles Healthcare System.

Data source

This was an observational retrospective study of all HCV-infected patients treated at the VA Greater Los Angeles Healthcare System within the Corporate Data Warehouse for a diagnosis of chronic hepatitis C by the *International Classification of Diseases*, ICD-9 or ICD-10, coding. Data were extracted from January 1, 2014 to December 31, 2016 and included: baseline demographic and clinical characteristics, medication, laboratory results, outpatient visits, and previous diseases/diagnoses.

Study population

The study population consisted of consecutive HCV-infected patients who received DAA therapy at VA Greater Los Angeles Healthcare System. Patients without SVR data available 12 or more weeks following antiviral therapy were excluded from analysis. All genotypes 1-6 were included. Choice of DAA regimen was at the discretion of the provider. On-treatment and posttreatment monitoring followed an established protocol that included serum SVR evaluation every 2-4 weeks.

Sustained virological response

The primary outcome of our study was SVR12, which was defined as an undetectable HCV RNA (<15 IU/mL) 12 weeks or beyond the conclusion of treatment [19].

Baseline characteristics

Baseline demographic variables obtained at the initiation of therapy included: age, self-reported

race and ethnicity; HCV genotype; nonalcoholic fatty liver (NAFLD) fibrosis score (<−1.455

being unlikely to have advanced NAFLD fibrosis vs >0.676 being predictive of advanced NAFLD-

associated fibrosis) [20]; the Fibrosis-4 (Fib4; which will be referred to as advanced fibrosis from

here forward), as a marker of advanced liver disease using the formula (age x aspartate

aminotransferase)/(platelets X alanine aminotransferase$^{1/2}$) [21]; body mass index (BMI)

($\geq$30 kg/m^2 and <30 kg/m^2); HIV status; hemoglobin A1c (HbA1c); and HCV treatment status

(naïve or experienced). For race and ethnicity, we used a single variable that combines concepts

of race and ethnicity into five mutually exclusive categories for race/ethnicity: non-Hispanic

White, non-Hispanic Black (African-Americans), Hispanics, Asians, and Unknown/Other. The

presence of diabetes and its complications were determined by ICD-9 (250.00-250.92) and ICD-

10 codes, as were diagnoses of hypertension, dyslipidemia, HIV, and AIDS with their

complications. Psychiatric disorders, both organic and nonorganic (associated with substance

abuse or not) were included in our analysis as was substance abuse and homelessness.

Assessment of T2DM improvement

To determine how HCV cure affects T2DM, we determined the HbA1c before and after DAA

treatment. Serial BMI and HbA1c values were obtained from the year before and the year after

HCV treatment and summarized as one-year pre-treatment and one-year posttreatment averages,

respectively. A significant change of HbA1c was defined as a difference of 0.5 or greater,

consistent with prior similar studies [22,23]. Through chart review, we also documented whether a

patient had an increase or decrease in oral hypoglycemic dose and/or injectable insulin dose from

one year before to one year after treatment with DAA. A change of greater than 10% from baseline was considered a significant change in medication, similar to prior studies [24]. Daily insulin was calculated as a total amount of basal and/or meal-time insulin over a 24-h period as documented in the patient's medication list. We also documented if there was a change in the overall number of diabetes medications from one year before to one year after treatment of DAAs.

Medication adherence

Patient adherence was assessed by calculating the medication procession ratio (MPR) since this has been a validated method to determine adherence [25]. MPR was defined by the total number of pills supplied over the total number of pills expected to be dispensed by the pharmacy department based on length of treatment regimen and genotype. For purposes of simplicity, "MPR" will be defined and described as "adherence" from here onward.

Statistical analysis

Demographic data which included sex, race/ethnicity, DAA regimen, body mass index (BMI), advanced fibrosis, and genotype were summarized with frequencies and chi-square tests for comparisons. We conducted multivariable logistic regression analysis to model predictors of SVR12. A priori covariates selected for the model were age, race/ethnicity, genotype, treatment regimen, treatment length, being treatment naïve, HIV status, advanced fibrosis, NAFLD fibrosis score, and adherence as defined by MPR. Age and adherence were continuous variables in the analysis. To analyze the adherence data further, we stratified adherence into the following groups: ≥90%, 80%-89%, 60%-79%, and <60%. In addition to using adherence as a continuous variable we also performed regression models with MRP with the defined groups as a categorical covariate.

6

Prior to regression analyses, we tested for multiple collinearity, and no covariates were collinear as defined as a variance inflation factor of less than 10. Separate logistic multivariable regressions models were used to model SVR predictors in the African-American subgroup by assessing treatment length of 8- versus 12-week treatment, removing what is now considered suboptimal therapies (sofosbuvir/simeprevir ± ribavirin and sofosbuvir/ribavirin). We also performed regression models by race/ethnicity, genotype, and adherence on SVR12.

Due to the rare event of DAA failure, to examine the relationship between SVR12 and HbA1c, we performed univariable and multivariable penalized maximum likelihood logistic regression analyses similar to prior studies [26,27].

<u>Results</u>:

Baseline characteristics and treatment regimens

Of the 1204 patients meeting the inclusion and exclusion criteria, 1068 patients were included for analysis based on having complete demographic and follow-up data. Baseline characteristics of the cohort are presented in **Table 1**. Males comprised 97% of the study cohort, consistent with gender demographics within the VA system. The mean age was 61.8 (SE ± 0.2). White people and African-Americans were equally represented at 37.8% of the population. Genotype 1 (1a and 1b) was the most common genotype in 83.9% of patients (N = 896), followed by genotype 2 (7.9%, N = 84) and genotype 3 (6.9%, N = 74). None of the patients had genotype 5. Of all patients, 35.4% were considered to have advanced liver disease as defined by a Fib4 > 3.25. A minority of patients were HIV positive (3%, N = 35). The majority of patients were treatment naïve at the time of DAA initiation at 79.5% (N = 849).

DAA treatment regimen allocations for all patients and the corresponding SVR12 for each genotype are summarized in **Table 2**. The most common regimen was sofosbuvir/ledipasvir ± ribavirin, which occurred for 47.8% of the population. Since the study cohort included patients started on antiviral therapy from January 1, 2014, our data also include older antiviral regimens such as sofosbuvir/simeprevir (17.5%, N = 187), sofosbuvir/simeprevir ± ribavirin (0.7%, N = 7), and sofosbuvir/ribavirin (9.8%, N = 105). Our data also include a subgroup of African-Americans patients who only received 8 weeks of therapy instead of 12 (N = 159) based on pretreatment viral level. These patients were treated before 2015 and prior to recommendations to use 12-week regimens in African-Americans [28].

Predictors of SVR12

Predictors of SVR12 from adjusted regression models are summarized in **Table 3**. There were two clinically significant negative predictors of SVR12: African-American race/ethnicity (aOR = 0.43; 95% CI = 0.27-0.69) and advanced liver disease (Fib4 score >3.25) (aOR = 0.4; 95% CI = 0.26-0.68). Covariates that did not affect SVR12 included age, genotype, HIV status, advanced NAFLD fibrosis score, BMI ≥ 30 kg/m^2, features of the metabolic syndrome (hypertension, dyslipidemia, T2DM) and whether the patient was treatment naïve or experienced.

SVR12 by race/ethnicity

SVR12 differences by race/ethnicity were also observed. African-Americans reached SVR12 85% of the time, while White people had SVR rates of 89% and Hispanics of 83%. When older therapies (sofosbuvir/simeprevir, sofosbuvir/simeprevir ± ribavirin, sofosbuvir/ribavirin) were excluded from the analysis, African-Americans reached SVR12 87.8% of the time, while White people and Hispanics achieved SVR12 rates of 92.4% and 88.7%, respectively. Hispanics had an adjusted OR of 0.75 (95% CI = 0.43-1.31); Asians of 0.61 (95% CI = 0.67-5.43); and other/unknown race/ethnicity of 0.67 (95% CI = 0.13-3.37).

Changes in HbA1c and BMI

Overall, average HbA1c was significantly lower after DAA therapy: 7.44% vs 6.71%, P = 0.01. For the subgroup of patients that achieved SVR12, the average HbA1c before treatment was significantly higher than the average after treatment (7.35% vs 6.55%, P < 0.01). When SVR12

was not achieved, however, HbA1c was not significantly different before and after treatment: 8.60% vs 8.61%, P = 0.99 (**Table 4**).

SVR12 is not affected by the components of metabolic syndrome

After adjusting for age, sex, race/ethnicity, cirrhosis, treatment experience, HCV genotype, treatment regimen, HIV status, and treatment duration, the individual components of metabolic syndrome (obesity, HTN, HLD, and T2DM) and the presence of metabolic syndrome itself did not predict SVR12 (**Table 5**).

Forty-six patients were on insulin before treatment and 43 patients were on insulin after treatment. Of those patients who were on insulin, the average daily insulin requirement before treatment was 55.1 IU (5.7) and 49.7 IU (6.2) after treatment (P = 0.50). For patients on insulin who achieved SVR12, the average daily insulin requirement before treatment was 55.0 IU (5.85) and the average daily insulin requirement after treatment was 48.2 IU (6.30) (P = 0.42). Insulin requirement also did not change significantly for patients who did not achieve SVR12 [55.5 IU (20.4) vs 58.1 IU (21.8), P = 0.93]. No patients analyzed were on any non-insulin injectable diabetes medications. There was no difference between the number of diabetes medications per patient before or after DAA therapy (1.23 vs 1.26, P = 0.43). The study included 44 patients (41.5%) defined as overweight (BMI ≥ 25), 30 (28.3%) that were defined as obese (BMI ≥ 30), and 12 (10.4%) with severe obesity (BMI ≥ 40). The average BMI for all patients before treatment was 30.1 kg/m2 (0.53), and the average BMI for all patients after treatment was 30.2 kg/m2 (0.54) (P = 0.92). For patients who achieved SVR12, the average BMI before and after treatment were not statistically different: 30.3 kg/m2 (0.56) vs 30.3 kg/m2 (0.57), P = 0.96. Similarly, the average BMI was not different before and after treatment for the patients that did not achieve SVR12: 28.8 kg/m2 (1.4) vs 29.2 kg/m2 (1.5), P = 0.92.

SVR12 in African-Americans by treatment duration subgroup

Given previously published data suggesting that African-American patients require 12 weeks of therapy regardless of baseline viral load, we further investigated SVR12 rates for African-Americans when stratifying by duration of treatment (8 weeks and 12 weeks) [5,28]. Only 159 African-American patients were treated for 8 weeks of therapy given the recent change in clinical practice. The adjusted odds ratio for SVR12 for all genotypes in the 8-week group was 0.34 (95% CI = 0.09-1.29) compared to 0.4 (95% CI = 0.25-0.63) in the 12-week treatment group (N = 1043).

There were 746 African-American patients who were treated with DAA regimens other than sofosbuvir/ledipasvir ± ribavirin and sofosbuvir/ribavirin. In this subgroup, the adjusted OR for SVR12 among African-Americans was 0.45 (95% CI = 0.25-0.81), consistent with a significantly lower SVR12 for those only using optimal therapy when compared to those on obsolete therapies. However, when adherence, as defined by MPR, was included in the model, the adjusted OR was 0.47 (95% CI = 0.21-1.07), mitigating the effect. When adjusting for homelessness, substance abuse, and mental health disorders (N = 306), only homelessness affected SVR12 rates among African-Americans (aOR = 0.39; 95% CI = 0.19-0.81).

SVR12 in genotype 1 patients

Given the heterogeneity of our patient population in genotype and therapies, as well as the observation that genotype 3 is more difficult to eradicate [29], we performed subanalyses on patients with only genotype 1 disease (1a and 1b) (N = 872). Similar to the larger cohort, African-American race/ethnicity was a significant predictor for non-SVR12 with an adjusted OR of 0.48 (95%

11

CI = 0.29-0.80). When older therapies were excluded (N = 69), being African-American race/ethnicity remained a significant predictor with an adjusted OR of 0.47 (95% CI = 0.23-0.82). When adherence was included in the model, African-American race/ethnicity was not a significant predictor (aOR = 0.60; 95% CI = 0.31-1.17). When addressing SVR12 only in the African-American subgroup treated for genotype 1 (N = 358), the only predictor of SVR12 failure was advanced liver disease (aOR = 0.35; 95% CI = 0.12-0.97).

Discussion:

This study demonstrates that in a large ethnically diverse community-based VA practice, SVR12 rates for chronic hepatitis C are influenced by race/ethnicity and advanced liver disease, which corroborates previously published data [28,30,31]. Having the metabolic syndrome or features of it (T2DM, hypertension or dyslipidemia) did not affect SVR12 rates. We also observed that T2DM (as measured by HbA1c or changes in medication use) improves after DAA therapy, independently of BMI. In our cohort, the lower SVR rates observed in African-Americans relative to Whites is persistent despite at least 12 weeks of therapy even when only using current "optimal" therapies. One important consideration in our analysis was the effect of adherence by measures of MPR in an ethnically diverse population treated with direct acting antivirals, which has not been evaluated in detail previously. We find that adherence explains some of these differences. These data suggest potential underlying biological differences between Whites and African-Americans in medication response.

While adherence and medication tolerability are not a concern in well-resourced large clinical trials, they can be more difficult to measure in real-world effectiveness data for chronic hepatitis C treatment. In 2007, Backus and colleagues assessed SVR12 rates in a VA cohort during the interferon era and demonstrated adherence, as defined by MPR, to be a predictor of SVR success. Patients with adherence of 90% or greater reached SVR 88% of the time. There was a clear threshold where any adherence less than 80% negatively impacted SVR, resulting in SVR12 rates of only 8% [30]. More recently in 2017, Louie and colleagues assessed real-world effectiveness SVR12 with the use of DAAs in the Kaiser Permanente Southern California health care system. Although African-Americans only represented 8% of their population (N=17), adherence was also

a predictor of success with an adjusted OR of 2.28 if it reached 80% or higher [32]. Our study demonstrated similar findings. We find that patients with adherence rates of <90% did not reach SVR12. Adherence appeared to attenuate the association between race/ethnicity and SVR12, thus explaining some of the differences observed, although biological causes have not been addressed.

One potential biological explanation for these observations is SVR12 differences in drug metabolism, driven by the patients' genetic background. Drug metabolism differences have been described to exist between African-Americans and Whites. Although some of these differences have been attributed to environmental factors such as diet and concomitant medications, intrinsic host factors such as genetic variability and gene polymorphisms in drug metabolizers such as CYP2D6 and CYP2C19 have also been described [33,34]. Understanding genetic polymorphisms in African-Americans and elucidating their mechanism of action, namely in the IL28B gene also offered great advances in understanding the underlying genetic differences between African-Americans and Whites in the interferon era [35,36]. Other genetic differences such as the variant in HAVCR1 gene variant (rs6880859) were also subsequently identified [37]. Although the polymorphisms and genetic differences listed above have not been shown to affect SVR12 with DAAs, such undiscovered genetic polymorphisms could explain these findings. The effects of genetic variants on SVR in the DAA era have not been addressed in any systematically other than to assess patterns of resistance [38]. The previous findings that extending therapy in African-Americans from 8 to 12 weeks to reach similar SVR12 to Whites, also points to potential biological differences between differences races and ethnicities. To address this, Large Genome Wide Association Studies (GWAS) or comparing genotypes in African-Americans and Whites who reached and did not reach SVR are needed.

The other important finding from our study is the improvement in HbA1c in patients cured of chronic HCV in the DAA era as measured by improvements in HbA1c or T2DM medication use, which has been previously described [39]. By analyzing each individual patient and examining both changes in the number of T2DM medication and the dosage, we find that the change of HbA1c was not due to an increase in oral hypoglycemic or insulin treatment. The findings imply that the change in HbA1c was most likely due to a change in host insulin resistance due to HCV clearance. This is in line with a recent study showing that HCV clearance with DAA reverses insulin resistance [40]. The data presented here are also consistent with data during the pegylated-interferon era where SVR clearance was associated with decreased insulin resistance and improved beta-cell function [13,41,42].

In conclusion, we present a large, ethnically and medically heterogeneous population within the VA System and their SVRs rates in the DAA era. Our data demonstrate the importance of racial/ethnic differences in SVR12 as well as regression of insulin resistance with HCV cure. Although DAAs have revolutionized chronic HCV care, it remains to be seen what the long term outcomes are since DAAs have only become available in 2014. What improvements in HCV-related insulin resistance mean in the context of having underlying NAFLD and NASH, where the prevalence of T2DM is high, has yet to be determined due to lack of long-term follow up of cured patients and the difficulty in identifying NAFLD and NASH patients using Electronic Medical Records within the VA and in other health care systems [43,44]. These questions are especially relevant in the setting of recent studies demonstrating a rise in HCC, which is closely associated with insulin resistance/T2DM and NAFLD [11,45,46]. Further understanding of these relationships will need to be the focus of new studies given the controversies associated with HCC screening and the large economic burden this places on our primary care and sub-specialty clinics.

Table 1: Demographics of study population.

Demographic	N=1068
Age in years, mean	61.8
SVR, %	87.0
Race/Ethnicity, % (N)	
Whites	37.5 (400)
African American	37.5 (401)
Hispanic	15.1 (161)
Asian	0.7 (7)
Other/Unknown	21 (9.3)
Genotype 1 (a & b), % (N)	83.9 (896)
Treatment naïve, % (N)	79.5 (849)
BMI > 30, % (N)	33.5 (358)
T2DM	14.4 (154)
Dyslipidemia	41.9 (448)
Hypertension	56.3 (601)
Advanced fibrosis, % (N)	35.4 (462)
HIV positive, % (N)	3.3 (35)
Homelessness, % (N)	22 (236)
Substance abuse, % (N)	27.9 (298)
History of psychiatric disorder, % (N)	58.5 (625)
Medical Procession Ration, % (N)	
<60%	5.3% (57)
60-79%	0.5% (5)
80-89%	0.2% (2)
≥90%	94%(1004)

Table 2: Treatment regimen allocation and SVR12 rates for all patients and genotypes.

Treatment Regimen by Genotype	SVR12, % (N)	
Genotype 1	*1a*	*1b*
sofosbuvir/ledipasvir	95.0 (314)	90.0 (116)
sofosbuvir/ledipasvir/ribavirin	87.2 (89)	92 (27)
sofosbuvir/ribavirin	100.0 (1)	85.5 (1)
sofosbuvir/simeprevir	78.0 (138)	100.0 (57)
sofosbuvir/simeprevir/ribavirin	100.0 (8)	100.0 (1)
paritaprevir/ritonavir/ombitasvir + dasabuvir	100.0 (1)	91.5 (47)
paritaprevir/ritonavir/ombitasvir + dasabuvir/ribavirin	83.6 (152)	91.5 (50)
grazoprevir/elbasvir	100.0 (1)	100.0 (8)
Genotype 2		
sofosbuvir/ribavirin	77.8 (97)	
sofosbuvir/ledipasvir/ribavirin	100.0 (1)	
Genotype 3		
sofosbuvir/ledipasvir	100.0 (1)	
sofosbuvir/ledipasvir/ribavirin	82.1 (43)	
sofosbuvir/ribavirin	61.1 (20)	
sofosbuvir/daclatasvir	83.3 (7)	
sofosbuvir/daclatasvir/ribavirin	81.8 (13)	
Genotype 4		
sofosbuvir/ledipasvir	100.0 (8)	
sofosbuvir/ribavirin	100.0 (3)	
paritaprevir/ritonavir/ombitasvir + dasabuvir/ribavirin	100.0 (1)	
paritaprevir/ritonavir/ombitasvir/ribavirin	100.0 (2)	
Genotype 6		
sofosbuvir/ledipasvir	100.0 (1)	

Table 3: Odds ratios for SVR12 for all patients.

Patient characteristic	Unadjusted		Adjusted	
	OR	95% CI	OR	95% CI
African American (ref. Whites)	*0.77*	*0.54-1.11*	*0.43*	*0.27-0.69*
Age	1.02	1.00-1.05	1.04	1.00-1.07
Genotype 2 (ref. genotype 1)	0.5	0.29-0.88	2.47	0.25-24.5
Treatment Naïve (ref. treatment experienced)	1.44	0.95-2.17	1.46	0.90-2.37
Advanced fibrosis (ref. Fib4<3.25)	*0.4*	*0.28-0.58*	*0.40*	*0.26-0.68*
BMI $\geq$ 30 (ref. BMI<30 kg/m^2)	0.9	0.62-1.31	1.18	0.79-1.78
HIV positive (ref. HIV negative)	1.44	0.43-4.8	1.77	0.48-6.54
MPR 80-89% (ref. $\geq$90%)	0.15	0.01-2.37	*	*
MPR 60-79% (ref. $\geq$90%)	0.04	0.004-0.32	<0.1	0
MPR <60% (ref. $\geq$90%)	0.02	0.01-0.05	0.01	13-166

(*) omitted in the analysis given only 2 patients.

Table 4: Mean HbA1c before and after HCV treatment by genotype and SVR12 status

Genotype	Before DAA (± SE)	After DAA (± SE)	*P*-value
Mean HbA1c of patients who achieved SVR12			
1a	*7.5 (0.19)*	*6.68 (0.14)*	*0.001*
1b	*7.3 (0.22)*	*6.59 (0.21)*	*0.03*
2	7.09 (0.35)	6.29 (0.21)	0.06
3	7.12 (0.37)	6.10 (0.39)	0.08
4	5.5 (0.13)	5.30 (NA)	NA
Overall	*7.35 (0.13)*	*6.55 (0.11)*	*<0.01*
Mean HbA1c of patients who did not achieved SVR12			
1a	8.98 (1.14)	8.95 (1.41)	0.98
1b	6.4 (NA)	6.50 (NA)	NA
2	No observations		
3	8.5 (NA)	8.70 (NA)	NA
4	No observations		
Overall	8.6 (0.89)	8.61 (1.08)	0.99

Table 5: Odds ratios for SVR12 for features of the metabolic syndrome

Patient characteristic	Unadjusted		Adjusted	
	OR	95% CI	OR	95% CI
Obese	0.90	0.62-1.31	1.25	0.82-1.90
Hypertension	1.06	0.74-1.52	0.81	0.51-1.28
Dyslipidemia	1.42	0.97-2.10	1.31	0.87-1.97
T2DM	0.72	0.48-1.10	0.82	0.55-1.09
Metabolic syndrome	1.04	0.68-1.60	1.81	0.75-4.37

Chapter 2

Clinical characteristics and outcomes of nonalcoholic fatty liver disease associated with

hepatocellular carcinoma

Introduction:

The metabolic syndrome, as defined by the clustering of biochemical and clinical features, has increased to epidemic proportions. Non-alcoholic fatty liver disease (NAFLD), the liver manifestation of the metabolic syndrome, has increased in parallel and is the most common cause of liver disease in the United States [47]. NAFLD can progress to nonalcoholic steatohepatitis (NASH), fibrosis, cirrhosis as well as hepatocellular carcinoma (HCC) [48]. NAFLD and its complications are predicted to continue to increase over the next decade, which is likely a reflection of the progression of disease and the aging patient population [15]. This has and will continue to have a large economic burden on society, especially given the lack of optimal current therapies for NASH.

Partly fueled by the NAFLD and NASH epidemic, HCC incidence has also shown a continuous increase over the years, which has placed economic strains on health care [49]. NAFLD patients have been shown to have a 7-fold increase in HCC incidence which tends to occur in older Hispanic patients with cirrhosis [45,46]. Although the majority of NAFLD and NASH-related HCC cases occur in a cirrhosis background, non-cirrhosis HCC cases have been described to occur in up to 50% of cases [50,51]. Features of the metabolic syndrome and more specifically type 2 diabetes mellitus (T2D) are highly associated with HCC development [11]. In a large prospective study with a 26-year follow-up, the length of T2D and the number of features of the metabolic syndrome were associated with HCC development in patients with and without cirrhosis, suggesting a mechanistic role between chronic inflammation, insulin resistance and liver carcinogenesis [52].

Early HCC detection has been shown to improve mortality [53-55]. However, current society guidelines only recommend HCC screening in patients with cirrhosis or those with high risk

features of chronic hepatitis B (HBV) infection. This poses a clinical dilemma given the number of patients with NAFLD, the increase in HCC incidence and the potential for cancer in non-cirrhosis population not currently targeted for screening [56]. Thus, identifying clinical high-risk factors and understanding the tumor growth rate in NAFLD-associated HCC may provide valuable insight into how to identify and stratify an "at-risk" patient population with NAFLD for screening.

Accordingly, we aimed to study the clinical features of NAFLD-associated HCC in cirrhosis and non-cirrhosis patients, study NAFLD-HCC outcomes compared to viral etiologies of HCC, and measure tumor growth as compared to viral etiologies of HCC.

Materials and Methods:

The study was approved by the Institutional Review Board of the University of California, Los Angeles (IRB#17-000015).

Data source

This is a retrospective case control study comparing NAFLD (including non-alcoholic steatohepatitis or NASH), chronic HBV and C (HCV) HCC cases. Our data source for the NAFLD-HCC cases were evaluated between 1/1/2000 to 12/31/2016 and comprised of the UCLA Jonsson Comprehensive Care Center (JCCC) cancer registry as well as review of liver surgical, hepatology and oncology clinic patient visits identified in the UCLA Electronic Medical Records (EMR) using *International Classification of Diseases* (ICD-9 and ICD-10 codes) for HCC with a diagnosis of NAFLD, NASH, the metabolic syndrome, or features of the metabolic syndrome as defined by diabetes (including its complications diabetic nephropathy, diabetic polyneuropathy, diabetic retinopathy), hypertension and dyslipidemia (see appendix). Given the under-reported cases of NAFLD and NASH using EMR ICDs [43,44], we identified additional cases using natural language processing of all pathology, operative, diagnostic and interventional radiology (using *Current Procedural Terminology*, CPT codes) reports with the following key terms: "NASH", "NAFLD", "steatohepatitis", "ballooning" or "NAFLD activity score". HBV and HCV cases were used as a comparative group and were identified from the Liver Cancer Center in Pasadena, CA. Hepatitis B and C patients received their care at the local clinic and/or at a tertiary transplant center including at UCLA Medical Center [53].

NAFLD patient population

Men and women ≥ 18 years were included. Patients with mixed HCC and cholangiocarcinoma on pathology report were excluded as were patients who reported excessive alcohol consumption as defined by the AASLD guidelines (>21 standard drinks on average per week for men and >14 standard drinks on average in women). All patients with a diagnosis of HBV (positive surface antigen), HCV (positive HCV RNA or history of SVR12), primary biliary cholangitis, primary sclerosis cholangitis, alpha-1 anti-trypsin, auto-immune hepatitis, Wilson disease and hemochromatosis were also excluded.

Baseline laboratory and clinical data

All patients who met inclusion criteria had laboratory data, including body mass index (BMI), evaluated closest to the time of HCC diagnosis. If patients had laboratory data after 6 months from the time of HCC diagnosis, they were included in our final analysis without their laboratory data. Hypertension, dyslipidemia, T2D and glucose intolerance were defined by diagnosis (as defined by ICD codes or from review of cardiology notes) or being on a medication associated with that diagnosis. HCC cases were defined as anyone with evidence of Li-RADS-5 lesions on CT or MRI or evidence of HCC on liver biopsy or on evaluation of the explanted liver (including autopsy in the event of death).

HCC characteristics and tumor growth

Tumor number and size were collected for all patients from MRI or CT scans with and without contrast. Pathology data (from biopsy, explant or resections) were reviewed when available. Studies done outside UCLA (interpreted locally or outside the institution) were included if local imaging data were not available. Abdominal ultrasound data were excluded. Tumor growth was determined based on the number of patients having two consecutive 2D images by CT or MRI

(prior to any therapy) and determining the size and time difference in cm between the two studies. When available, tumor growth in patients with multiple tumors was measured in those meeting Li-RADS-5 criteria [57]. Tumors measuring <2cm on initial imaging study were included if they were confirmed HCC on subsequent studies by imaging or pathology data. HCC cases were classified using the Milan criteria (single lesion 5 cm, maximum of three lesions with none >3 cm) and by the University of California at San Francisco (UCSF) criteria (single lesion 6.5 cm, maximum of three lesions with none >4.5 cm, or a total tumor burden of 8 cm). Metastasis was determined based on abdominal CT or MRIs as well as CT chest and bone scans. Recurrence was defined if on a subsequent scan a new lesion was identified after evidence of resolution on prior imaging study. We excluded any lesions that did not meet Li-RADS-5 criteria or if a study was done outside of UCLA without contrast agent.

Statistical analysis

The p values for between group comparisons of continuous variables that did not follow the normal distribution were calculated using the non-parametric Kruskal-Wallis method. The p values for comparing continuous variables such as age that followed the normal distribution were computed using a one-way analysis of variance model.

The p values for comparing binary data across groups were computed using Fisher's exact test.

Recurrence free survival curves were calculated using the Kaplan-Meier method and the p values for their comparison were performed using the log rank test.

A Cox proportional hazard model was used to compare recurrence free survival curves adjusted for covariates. The Hazard (event rate) ratio (HR) and its 95% confidence bounds under this model are reported. A linear regression model was used to compute age and gender adjusted means for (log) tumor growth. To measure tumor growth, we used the log tumor growth since log tumor

growth is normally distributed. Logistic regression was used to model cirrhosis risk as a function

of gender, age, T2D/glucose intolerance, dyslipidemia and BMI.

Results:

Validation of NAFLD HCC cases

3,358 HCC cases were identified in the JCCC cancer registry of which only 22 had a diagnosis of NASH. One patient who had both HCV and NASH was excluded. 4,809 cases were identified from the EMR abstraction using ICD codes and natural language processing. To validate our algorithm, key words from natural language processing (NASH, NAFLD, steatohepatitis, ballooning and metabolic syndrome) were counted the number of times they appeared in an individual's chart. Cases were then ranked based on those numbers, and individual patient chart reviews were conducted to validate these definitions. All highly ranked "NASH" patients were evaluated, followed by all highly ranked "NAFLD", "steatohepatitis", "ballooning", "NAFLD activity score" and "metabolic syndrome" cases. A total of 10-15 charts were reviewed in patients who had a key term appear once or twice and no confirmed cases were identified. 10-15 cases were reviewed in patients who only had one key term appear without identifying any NAFLD or NASH cases, and an additional 20 charts were reviewed without any key terms, which did not confirm any cases. Of the 430 charts reviewed, 127 met inclusion criteria. One patient was removed from the final analysis due to being the only one having received stereotactic body radiation therapy; another patient was removed due to an unclear sequence of events related to the diagnosis of HCC (outside scans only available for review) and only a small focus on explant that was not consistent with the imaging data (<0.5 cm without field defect from prior therapies). Therefore, 125 cases were included in our final NAFLD cohort. Of note, one patient had possible autoimmune hepatitis versus NASH, two patients had positive HCV antibody without positive RNA or a history of HCV treatment, and one patient was homozygous for C282Y mutation with elevated ferritin (>1000), but did not show evidence of iron overload on pathology. Due to the retrospective nature of the

study, patients self-identified their race and ethnicity. Given the large missing data for race, only ethnicity was included in the final analysis.

Demographics of NAFLD HCC cases

The mean age of the NAFLD-associated HCC cases was 64.8 years with a mean BMI of 31.2 kg/m^2 ($\pm$ 10.5 kg/m^2). The majority had hypertension (n=85, 68%) and dyslipidemia (n=44, 35%). Seventy percent (n=87) had T2D (n=82) or glucose intolerance (n=4). The majority had T2D or glucose intolerance for $\geq$10 years (n=32), with 16 patients having had the disease for 2-10 years and only 2 for 0-2 years. Of the patients with T2D, 32% (n=27) were on insulin therapy. The median A1c was 6.1 (IQR 5.4-6.95); however, the majority were on therapy by the time of A1c analysis. The majority of patients identified themselves as Hispanics (n=52, 42%).

Clinical characteristics of NAFLD, HBV and HCV associated HCC cases

Four patients who were co-infected with HBV and HCV were removed from the final analysis. Therefore, 168 cases were included in the HBV cohort and 158 in the HCV cohort. Demographics of all three groups are presented in **Table 1**. Unlike HBV and HCV cases who had mostly men (n=135, 80%, and n=97, 61%, respectively), NAFLD cases were equally distributed between men and women (n=59, 47%). NAFLD cirrhosis patients were more likely to have decompensated liver disease with 51% of the cohort with a Child-Turcotte-Pugh score of B and C (n=45), when compared to HBV who only comprised 29.4% (n=50) and HCV 23% (n=36) of the cohorts. This is consistent with more patients in the NAFLD group having hepatic encephalopathy compared to HBV and HCV (25% versus 11% and 6%, respectively; p<0.0001) and ascites/volume overload (36% versus 2% and 15%, respectively; p<0.0001). NAFLD cases also were more likely to have T2D or glucose intolerance (70%, p<0.0001). The HCC screening rate between all three groups

was not significantly different (p=0.0503). Patients with HBV were more likely to have a family history of HCC (25%) when compared to NAFLD (8%) or HCV (5%).

Tumor characteristics of NAFLD, HBV and HCV cases

NAFLD and HCV HCC cases were more likely to be within Milan and UCSF criteria for liver transplantation than the HBV group (**Table 2**). This was further confirmed when assessing the median size of the first tumor which was similar in the NAFLD and HCV groups (2.9 and 2.1 cm, respectively) but markedly larger in the HBV group at 4 cm (p=0.0003) (**Table 2**). To measure tumor growth, 230 cases had serial imaging studies available for review for all three groups. Tumor growth, as measured by % increase per month (given its linear increase), was not different between the three groups after adjusting for age and gender. The adjusted mean for the HBV tumor growth in males was 8.18% when compared to 7.79% in females. In the HCV group, the adjusted mean for males was 7.6% while it was 7.51% in females. In NAFLD, similar trends were seen with the adjusted mean in males being 6.69% and 6.29% in females, although none were statistically significantly different. We also found that NAFLD HCC patients were less likely to have a positive AFP (AFP ≥ 10) when compared to the other two groups (p<0.0001) (**Table 2**).

Clinical and tumor characteristics of cirrhosis versus non-cirrhosis HCC cases

To characterize which clinical features were more likely to predict cirrhosis and advanced fibrosis in the NAFLD HCC cohort, we classified the group into "cirrhosis/advanced fibrosis" or "non-cirrhosis". We defined cirrhosis/advanced fibrosis as anyone with clinical evidence of cirrhosis (platelets <150K, evidence of portal hypertension or as diagnosed by a hepatologist) or pathology review (all F3, F-4 and F4 on trichrome stain by METAVIR scoring system) [58]. The non-cirrhosis group was defined as anyone who had F0, F1-2, F2-3 disease on the pathology review or as

diagnosed by a hepatologist. A total of 86.5% of our cohort had pathology available for review. Eighteen patients (14%) had no evidence of clinical cirrhosis; 2 (2%) had F0-1; 4 (3%) had F1-2; one patient had F2-3 (1%); 6 (5%) had F3; 5 (3%) had F3-4 and 71 (89%) had clinical evidence of cirrhosis of F4 disease on liver biopsy.

Based on these definitions, 102 (82%) had cirrhosis/advanced fibrosis and 23 (18%) had non-cirrhosis liver disease. The demographics comparing the two groups are presented in **Table 3**. In an unadjusted bivariate analysis comparing clinical factors associated with cirrhosis, we identified that being of non-Hispanic ethnicity (OR=0.07, p=0.0001), having dyslipidemia (OR=0.321, p=0.0268) and having a lower BMI (p=0.00065) were associated with non-cirrhosis NAFLD HCC. In an adjusted multivariate analysis, patients who identified themselves as Hispanic (adj. OR=12.34, 95% CI 2.59-58.82) and who had a higher BMI (adj. OR=1.19, 95% CI 1.066-1.330) were more likely to have cirrhosis. T2D and glucose intolerance diagnoses showed an increased trend towards the cirrhosis group; however, the differences were not statistically significant (adj. OR=1.46, 95% CI 0.464-4.614). Similarly, age and gender were not predictive of cirrhosis/advanced fibrosis (adj. OR=0.970, 95% CI 0.859-1.095), while dyslipidemia demonstrated a trend towards decreased cirrhosis (adj. OR=0.538, 95% CI 0.1643-1.762).

Survival outcomes are not different between the NAFLD, HBV and HCV groups

We evaluated the overall and recurrence-free survival among the three groups and found that 217 patients died and 82 developed recurrences. **Figure 1A** represents the unadjusted overall survival for all three groups with a median follow up of 22.4 months. Our data demonstrate a better overall survival in the NAFLD group when compared to HBV and HCV. These findings persist after removing the 23 non-cirrhosis patients in the NAFLD group (**Figure 1B**). Unadjusted recurrence free survival, which had a median follow up of 19 months, was also better in the NAFLD group

when compared to HBV and HCV with or without the NAFLD non-cirrhosis patients (**Figure 1C-D**). The adjusted multivariate models for the overall and recurrence free survival are presented in **Table 4**. In the multivariate model adjusting for gender, ethnicity and most definitive treatment, we found that HBV patients had the lowest survival when compared to HCV (adj. HR=0.70, 95% CI 0.514-0.952) and NAFLD (adj. HR=0.236, 95% CI 0.103-0.539). NAFLD patients had improved survivals when compared to HCV (adj. 0.337, 95% CI 0.156-0.725). As expected, the type of definitive therapy influenced the survival rates of the groups, with the orthotopic liver transplantation (OLT) giving the most decrease in death or recurrence rates (adj. HR=0.115, 95% CI 0.062-0.214). In the adjusted recurrence-free model, we also observe that HBV patients had decreased survival rates compared to NAFLD patients (adj. HR=0.487, 95% CI 0.242-0.977). No differences were seen in the recurrence-free survival rates between the HCV and HBV groups (adj. HR=0.792, 95% CI 0.590-1.063) or the NAFLD and HCV group (adj. HR=0.615, 95% CI 0.325-1.161).

Discussion:

We present the largest NAFLD HCC cohort with the longest follow-up to date with detailed clinicopathological data. We demonstrate important clinical differences between NAFLD and viral etiologies of HCC, including that HBV-associated HCC patients present at a younger age and have larger tumors at the time of presentation, which lends them to be outside of OLT criteria. Although NAFLD patients tend to have more decompensated liver disease at the time of HCC presentation, the overall survival rates are better when compared to HBV and HCV, as is the recurrence-free survival in the NAFLD group when compared to HBV patients. Hester et al. recently analyzed the outcomes of a group of 97 NASH HCC patients. When compared to HBV, HCV and alcoholic-associated liver (ALD) disease, NASH HCC patients had worse overall survivals than ALD patients but similar survival rates as HCV or HBV cases (median follow up time of only 16 months) [59]. Wakai et al. evaluated post-surgical outcomes in 17 NAFLD-associated HCC cases and demonstrated that although the overall survival was not different between NAFLD, HBV and HCV patients, the recurrence-free survival was improved in the NAFLD cohort at a median follow-up time of 87 months, similarly to our findings [60,61]. Some of the differences observed between the studies can likely be explained by different patient demographics (especially given regional differences in liver transplantation allocations), sample sizes, and longer follow-up times.

HCC in the non-cirrhosis liver has been reported to occur in NAFLD [51,52]. Since distinguishing NAFLD, NASH and different stages of fibrosis remains a diagnosis based on pathology, assessing liver histology in NAFLD-associated cases of HCC is critical but is often lacking in larger studies. Our detailed pathological analysis enabled us to distinguish between cirrhosis/advanced fibrosis cases when compared to non-cirrhosis. We report that ~18% of our cohort did not have any cirrhosis, although our definition was conservative due to including all bridging fibrosis cases (F3

and F3-F4 by METAVIR) in the advanced fibrosis/cirrhosis group. We found that patients who self-identified as Hispanic and had a larger BMI are more likely to develop HCC in a cirrhosis background. We also noted that the cirrhosis HCC group was less likely to have dyslipidemia although this was no longer significant after adding BMI in our model. We interpret these data as potential statin effects (our definition of dyslipidemia included patients on lipid-lowering agents including statins) since their use has been shown to decrease fibrosis progression and HCC [52,62]. Few studies have attempted to differentiate statin effects between those with and without cirrhosis. In a recent case-control study comparing cirrhosis and non-cirrhosis cases based on histology, dyslipidemia (as defined by a high LDL cholesterol or triglycerides) was independently associated with HCC development in the non-cirrhosis group (adjusted OR=1.74, p<0.05) [63]. Although the cirrhosis group had a larger BMI (29.2 kg/m^2) when compared to the non-cirrhosis (26.1 kg/m^2) group, those differences were not significant (p=0.05), which is possibly explained by only having 28 NAFLD patients in the cohort of 545 individuals (5%). Statin use has also been shown to be associated with a decreased HCC mortality, although again in most studies NAFLD cases only comprise a small group of the patient population [64,65]. In a large retrospective study assessing the occurrence of HCC in 18,080 non-cirrhosis patients with NAFLD using Taiwan's National Health Insurance Research, Lee and colleagues found that statin use significantly decrease HCC incidence (HR=0.29, 95% CI 0.12-0.68) [66]. BMI differences were not addressed in their models, although Asian NAFLD patients often develop lean or non-obese NAFLD where BMI criteria are not as relevant, thus representing a different patient population from the ones normally seen within the U.S [67,68]. Teasing out the effects of dyslipidemia and statin use in different ethnic backgrounds will be important for future studies given our findings. Several clinical trials are ongoing assessing the effects of statin therapy on liver fibrosis progression and HCC recurrence, which will provide

valuable insight into the chemoprotective effects of statins (NCT03219372, NCT03024684, NCT03275376, NCT03654053, NCT02968810).

In addition to understanding the clinical factors that differentiate cirrhosis and non-cirrhosis cases, understanding tumor growth may help risk stratify the non-cirrhosis group. HCC screening practices have been developed partially based on image-based tumor growth measurements, which have mostly been studied in the context of viral etiologies of HCC [69]. Given the recent birth of NAFLD as a major etiology of cirrhosis and HCC in the U.S., few studies have evaluated this topic in the NAFLD patient population. To evaluate tumor growth, we evaluated the % change of the tumor per month, which is stable and steadily increases over time, given that most tumors will grow exponentially when small and will level with time [70]. We found that the median growth rate of HCC in all three groups was not different after adjusting for age and gender, suggesting that screening intervals should not differ between viral etiologies of HCC and NAFLD. We were not able to compare the tumor growth of cirrhosis and non-cirrhosis in the NAFLD group due to the small sample size of the non-cirrhosis group of only 6. However, this should be the subject of further studies to better understand whether and how often to screen this patient population, especially given that NAFLD patients are less likely to be AFP-producers (a HCC serum biomarker), as shown by us and others [45].

While our study highlights important differences between the cirrhosis and non-cirrhosis NAFLD HCC patient population as well as viral etiologies of HCC, there are limitations. UCLA is a large tertiary-care liver transplantation center in the U.S.; therefore, the majority of our patients were referred from outside institutions, thus creating a bias towards OLT evaluation and treatment. This may explain the large proportion of decompensated NAFLD patients who received OLT (versus

regional differences in transplantation allocation where region 5, which includes California, tends to have a sicker patient population [71]). However, this also allowed for a more diverse patient population due to the large referral pattern seen in Los Angeles. Being in a transplant center also provided for a detailed review of the pathology, which is often not available in large cohorts. To the best of our knowledge, there are no previous free language processing approaches that would have identified cases within the EMR. This approach also allowed us to minimize selection bias of only studying patients seen by hepatologists and therefore have OLT or other curative treatments offered. Inter-observer differences between radiographic assessments of HCC also introduced differences in tumor growth measurements given that most initial imaging studies were done outside of UCLA. We attempted to normalize these by only including CT or MRI studies that were re-interpreted at UCLA using the validated Li-RADS score. Another limitation is the small sample size of the non-cirrhosis cases of HCC (23), restricting further analyses, such as tumor growth, and teasing out the effects of dyslipidemia and statin treatment. This is especially relevant because recent obese mouse models have demonstrated that NASH HCC can occur through independent mechanisms of NASH [72].

In conclusion, we present a large, diverse NAFLD HCC patient population with detailed clinical and pathological data, allowing for important differences to be identified between various stages of fibrosis. Identifying population-specific biomarkers, which will likely require a combination of clinical risk factors, laboratory data and tumor growth data, will be important in this group of patients. These studies also support the use of longitudinal biomarker studies to identify potentially useful diagnostic targets. The association between BMI and dyslipidemia remains of crucial clinical significance due to the non-cholesterol and pleiotropic effects of statins on HCC and liver fibrosis. This provides an avenue for statin use as a chemoprotective agent not only in NAFLD

cirrhosis patients but specifically in the sub-group of non-cirrhosis patients who are not currently being targeted for screening. Although the ongoing prospective randomized clinical trials with statins will help elucidate the benefits of statins in cirrhosis and HCC, new prospective studies are needed to assess the benefits of statins in the non-cirrhosis group, which comprises a large portion of the NAFLD.

Table 1- Demographics and clinical characteristics of NAFLD, HBV and HCV cases

	NAFLD	HBV	HCV	*P* value
Males n, %	59 (47)	135 (80)	97 (61)	<0.001
Mean age at HCC dx ± SD	64.8 ± 8.5	57.7 ± 12.7	65.9 ± 10.3	<0.0001
Hispanic ethnicity, n (%)	52 (42)	2 (1)	29 (18)	<0.0001
T2D and glucose intolerance, n (%)	87 (70)	21 (13)	28 (17)	<0.0001
Decompensation, n (%)				
HE	31 (25)	19 (11)	10 (6)	<0.0001
Ascites/volume overload	45 (36)	3 (2)	25 (15)	<0.0001
Child-Pugh Score				
A	41 (41)**	120 (71)	121 (77)	
B	34 (34)	41 (24)	30 (19)	<0.0001
C	10 (10)	9 (5)	6 (3.7)	
Missing data	8	0	2	
Median INR (IQR)	1.2 (11-1.3)	1.1 (1-1.2)	1.1 (1.1-1.3)	0.0001
Median AST (IQR)	45 (33-60)	65 (35-113)	84 (49-128)	<0.0001
Median ALT (IQR)	32 (21-45)	54 (32-84)	64 (38-114)	<0.0001
Median bilirubin (IQR)	1.2 (0.7-2.3)	0.9 (0.7-1.6)	1.1 (0.8-1.6)	0.0246
Screened for HCC, n (%)	56 (44)	79 (47)	95 (58)	0.0503
Family history HCC, n (%)	10 (8)	42 (25)	8 (5)	<0.0001

** excludes patients without cirrhosis (see table 3)

Table 2- Presenting tumor characteristics between NAFLD, HBV and HCV

	NAFLD	**HBV**	**HCV**	*P* value
Within Milan	85 (68)	78 (46)	112 (69)	<0.0001
Within UCSF	100 (80)	92 (55)	131 (80)	<0.0001
Median first tumor size (cm) (IQR)	2.9 (2-4.5)	4 (2.4-7.6)	3 (2.1-4.6)	0.0003
Median tumor numbers (IQR)	1 (1-2)	1 (1-2)	1 (1-2)	0.232
Median tumor growth (% per month) (IQR)	4.5% (1.3-13.7)	7.3% (1.8-14)	5.7% (1.6-11.4)	0.6636
AFP-producers	35 (35)	111 (66)	124 (76)	<0.0001
Most definitive therapy, n (%):				
OLT	50 (40)	19 (11)	30 (19)	
Resection	14 (11)	37 (22)	14 (9)	
RFA	26 (21)	20 (12)	24 (15)	<0.001
TACE/Y-90	13 (10)	27 (16)	26 (16)	
PEI	0 (0)	2 (1)	6 (4)	
Chemotherapy	5 (4)	14 (8)	4 (3)	
Supportive	17 (14)	51 (30)	55 (35)	

IQR=interquartile range; AFP= alpha-fetoprotein; RFA=radiofrequency ablation; TACE=trans-arterial chemoembolization; OLT=orthotopic liver transplantation

Table 3- Clinical and tumor characteristics between the cirrhosis/advanced fibrosis versus non-cirrhosis group within the NAFLD cohort

	Cirrhosis/advanced fibrosis (n=102)	Non-cirrhosis (n=23)	*P* value
Males n, %	57 (56)	9 (39)	0.170
Mean age at HCC dx ± SD	64.3 ± 7.4	67.1± 12	0.286
Hispanic ethnicity	50 (49)	2 (9)	0.0001
Median BMI (IQR)	31.7 (28-34)	25.5 (22-30)	0.0007
T2D/GI, n (%)	72 (71)	14 (61)	0.478
Median A1c (IQR)	5.9 (5.4-6.9)	6.1 (5.4-6.5)	0.909
Hypertension, n (%)	67 (67)	17 (74)	0.623
Dyslipidemia, n (%)	30 (29)	13 (56)	0.0268
AFP-producers, n (%)			
Yes	29 (29)	6 (26)	>0.999
No	63 (62)	14 (58)	
Missing or no AFP	8 (8)	3 (13)	
Screened, n (%)	60 (58)	1 (4)	<0.0001
FHx of HCC, n (%)	9 (9)	4 (17)	0.266
FHx LD, n (%)	27 (26)	4 (17)	0.187

SD=standard deviation; IQR= interquartile range; GI=glucose intolerance; T2D=type 2 diabetes; FHx=family history; AFP=alpha-fetoprotein; LD=liver disease.

Figure 1-Overal and recurrence free survival of NAFLD, HBV and HCV cases. A. Overall survival of the three groups; B. Overall survival in all three groups without non-cirrhosis NAFLD group; C. Recurrence free survival for all three groups; D. Recurrence free survival in all three groups without the non-cirrhosis NAFLD group.
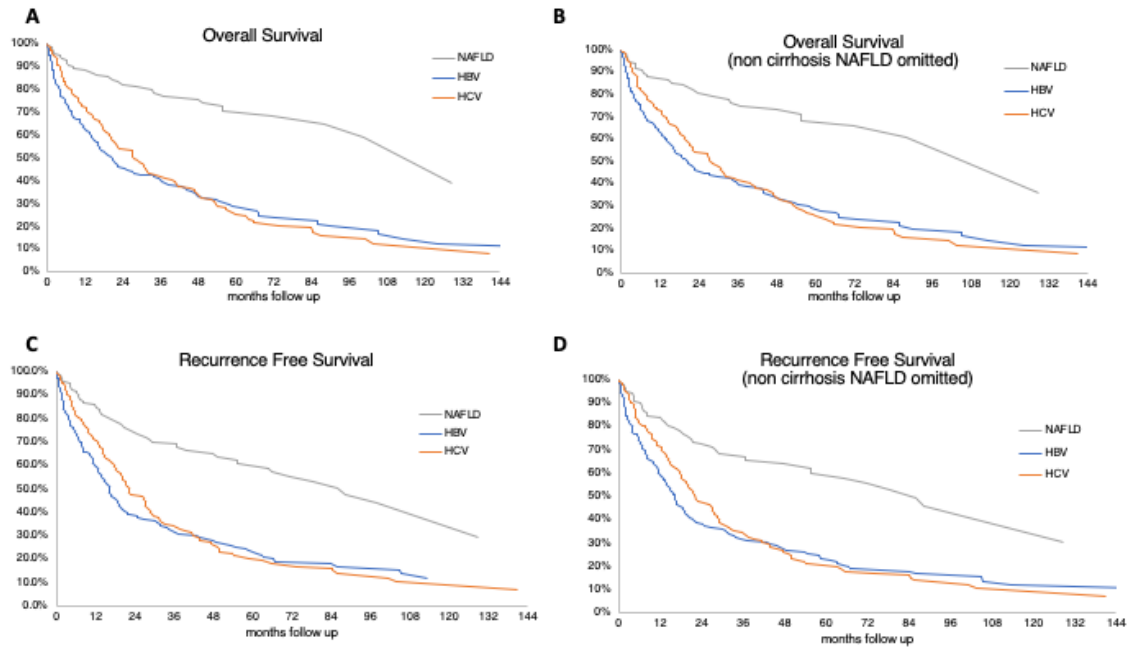
Table 4- Cox multivariable analysis of patients and treatment variables associated with overall

survival and recurrence free survival.

Variable	HR	95% CI	*P* value
Overall Survival			
Male gender	0.140	0.893-1.543	0.250
Etiologies:			
HCV vs HBV	0.711	0.523-0.966	0.0293
NAFLD vs HBV	0.235	0.105-0.525	0.0004
NAFLD vs HCV	0.331	0.145-0.694	0.0034
Ethnicity			
African American	Ref	-	-
White	0.888	0.312-2.523	0.8230
Asian	0.825	0.298-2.285	0.7108
Hispanic	1.331	0.449-3.948	0.6066
Not Hispanic	1.089	0.287-4.131	0.9001
Most definitive treatment:			
Chemotherapy	Ref	-	-
OLT	0.115	0.062-0.214	<0.0001
PEI	0.366	0.142-0.938	0.0364
Resection	0.177	0.096-0.328	<0.0001
RFA	0.175	0.093-0.329	<0.0001
Supportive care	1.325	0.784-2.240	0.2936
TACE	0.539	0.309-0.939	0.0290
Recurrence Free Survival			
Male gender	1.197	0.920-1.558	0.1799
Etiologies:			
HCV vs HBV	0.792	0.590-1.063	0.1199
NAFLD vs HBV	0.487	0.242-0.977	0.0429
NAFLD vs HCV	0.325	0.325-1.161	0.1336
Ethnicity:			
African American	Ref	-	-
White	0.966	0.340-2.745	0.9489
Asian	0.947	0.343-2.619	0.9169
Hispanic	1.280	0.433-3.780	0.6553
Not Hispanic	1.146	0.323-4.061	0.8327
Most definitive treatment:			
Chemotherapy	Ref	-	-
OLT	0.114	0.062-0.211	<0.0001
PEI	0.432	0.168-1.111	0.0816
Resection	0.306	0.168-0.556	0.0001
RFA	0.339	0.190-0.605	0.0002
Supportive care	1.446	0.852-2.455	0.1718
TACE	0.694	0.401-1.203	0.1935

Appendix:

JCCC and UCLA EMR

1. Males and females
2. Age ≥18 years at the initial diagnosis of HCC and NAFLD and/or NASH
3. Dates 01/01/2000-12/31/2016
4. Diagnosed with Hepatocellular Carcinoma (HCC)
   a. Either listed on JCCC Cancer Registry list (~3k patients)
   b. OR has diagnosis of Hepatocellular carcinoma (HCC) by EMR as defined by:
      1. ICD-9 or ICD-10: (155.2, 155.0) [C22.8, C22.0, C22.7]

AND DIAGNOSED with one of the following comorbidities:

5. Non-alcoholic fatty liver disease (NAFLD) patients
   a. NAFLD by (ICD-9) or [ICD-10]: (571.8, 571.9, 573.8, 573.9) [K76.9, K76.0, K76.89, K76.81]
6. OR Non-alcoholic steatohepatitis (NASH) patients
   a. NASH by (ICD-9) and [ICD-10]: (571.8, 571.9, 573.8, 573.9) [K75.81, K75.8, K75.9]
   b. NASH by pathology natural language processing using the following keywords: (NAFLD Activity Score (NAS) and include fibrosis score F0-F4):
   c. "steatohepatitis", "ballooning", "NAFLD activity score", "NASH" or "NAFLD"
7. OR metabolic syndrome (Please flag metabolic syndrome patients)
   a. Metabolic syndrome defined by ICD-9 or ICD-10: (277.7) [E88.8]
   b. OR Metabolic syndrome as defined by: diabetes [(ICD-9: 250.00) [ICD-10: E11]: Diabetic nephropathy N08.3, Diabetic polyneuropathy G63.2, Diabetic retinopathy H36.0], hypertension [portal hypertension (572.3 or K76.6) or Hypertension (401.9)] and hyperlipidemia (272.4) or [E78.4, E78.5]

ADDITIONALLY,

Evaluated all EMR patient from oncology liver clinic and all patients who had hepatic resection from 2010-2016 (separate lists).

Abstracted medications, A1c, liver tests, AFP level, all radiology data, all pathology data.

All EMR data was ranked based on following keywords: "steatohepatitis", "ballooning", "NAFLD activity score", "NASH", "NAFLD" or "metabolic syndrome"

Chapter 3

Novel lipid lincRNA *OLMALINC* regulates the liver steatosis gene, *SCD*, as an enhancer RNA

Introduction:

The metabolic syndrome (MetS), as defined by the clustering of phenotypic, biochemical and clinical factors, has reached epidemic proportions in the United States [73]. Non-alcoholic fatty liver disease (NAFLD), the liver manifestation of the MetS, has also increased in parallel with other determinants of the MetS [74]. NAFLD ranges from simple steatosis to inflammatory non-alcoholic steatohepatitis (NASH), which can lead to fibrosis, cirrhosis, and hepatocellular carcinoma (HCC) [48]. The pathophysiology and interplay of the MetS and NAFLD are complex, multi-factorial, and include both genetic and environmental contributions.

Intrahepatic lipid accumulation, steatosis, is the hallmark of NAFLD [75,76]. Although the pathogenic pathways that cause progression from steatosis to steatohepatitis and fibrosis remain elusive, human and murine models have demonstrated that lipid dysregulation plays an important role in the NAFLD pathogenesis [76-79]. Blood lipidomics data in NAFLD patients [77,80,81] and murine knock-out models [82] have also shown the importance of the monounsaturated fatty acid rate-limiting enzyme, stearoyl-CoA desaturase (SCD) in the MetS, steatosis, and NAFLD [77,80,81]. Targeting SCD in murine NASH models has shown promising results [83], which has recently led to human clinical trials with early phase data demonstrating reversal of hepatic steatosis using Aramchol, an SCD activity inhibitor [84].

As advances in deep and high-throughput sequencing have emerged, novel players have been identified in lipid biology, including the identification of a unique group of non-coding genes called long non-coding RNAs (lncRNAs) [85]. LncRNAs are >200 nucleotides long, show tissue and cell-type specificity, and can differentially regulate signaling pathways [86]. Understanding their biology has provided insight into new ways in which known key metabolic genes and proteins are

regulated beyond previously described mechanisms, such as acting as scaffolds to complex proteins and enhancer RNAs (eRNAs), and modifying chromatin states [85,87]. This has included the role of lncRNAs in the regulation of cholesterol and lipid pathways [88]. However, to the best of our knowledge, no eRNA lincRNAs have been discovered to regulate lipid metabolism as of yet.

In the present study, we identified the long intervening non-coding RNA (lincRNA), *O*ligodendrocyte *M*aturation-*A*ssociated *L*ong *I*ntergenic *N*on-Coding RNA (*OLMALINC*), in a statin- and triglyceride (TG) -associated liver co-expression network utilizing liver RNA-sequencing (RNA-seq) from 259 Finnish bariatric surgery patients from the Kuopio OBesity Surgery (KOBS) cohort with refined clinical phenotypic and liver histology data. We demonstrate that *OLMALINC* liver expression is highly correlated with the key lipid and TG pathway genes in the liver RNA-seq data, including *SCD*. We further functionally show that *OLMALINC* regulates this central TG metabolism gene, *SCD* as a regional eRNA. Taken together, these novel data indicate that *SCD* is regulated by the adjacent lincRNA, *OLMALINC*, which likely contributes to its central function in TG metabolism and liver steatosis.

Materials and Methods:

Study cohorts

The Kuopio OBesity Surgery (KOBS) cohort was recruited at the University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland [89]. All participants provided informed consent and the study was approved by the local ethics committee. The liver RNA-seq cohort comprises 259 Finnish KOBS participants who underwent bariatric surgery, during which liver biopsies were obtained. Clinical measurements were performed as described previously [89]. We also analyzed liver RNA-seq data on 96 GTEx samples [90].

Histological assessment of the liver biopsy and meta-liver trait, D1

The NASH Clinical Research Network (CRN) criteria were used to evaluate the liver histological data [91]. The following attributes were used: steatosis grade (0-3), lobular inflammation (0-2), ballooning (0-2), and fibrosis stage (0-4). The diagnosis for NASH was also determined by the pathologist following the standard guidelines [92,93]. To determine the NAFLD status with liver RNA-seq data, we performed a non-linear principal component analysis (PCA) using the homals R package [94] on the four CNR liver histological phenotypes and used the first principal component (PC1) as the aggregated meta-liver trait (D1) for NAFLD (Figure 1A). We note that the D1 is negatively correlated with the histological parameters, i.e. a higher D1 represents a healthier liver (Figure 1A).

Liver RNA-seq and expression quantification

RNA samples were isolated a using the miRNeasy (Qiagen) kit and sequencing libraries were prepared using Ribo-Zero gold (Illumina) kit to remove ribosomal RNAs. External RNA Controls Consortium (ERCC) spike-ins (ThermoFisher Scientific) were added as controls. We quantified

the transcript abundance as read counts and transcript per million (TPM) using Kallisto [95], based on GENCODE version 25 liftover to hg19 gene annotation. Gene-level quantification is estimated as the sums of read counts and TPM of all transcripts of a gene. To remove lowly expressed genes, a gene had to have $\geq$10 reads in 80% of samples, resulting in 15,670 genes in the final analysis.

Hidden covariate estimation for RNA-seq

We performed a supervised surrogate variable analysis (sSVA) [96] on TPMs and used the 92 ERCC spike-in transcripts as invariable controls to estimate hidden confounders in the liver RNA-seq data. The following covariates were included in the sSVA analysis: uniquely aligned reads %, mitochondrial reads %, 3' bias, BMI, sex, and age. Overall, 25 latent factors were estimated and we included all sSVA factors and known covariates in down-stream analyses. GTEx data do not contain ERCC spike-ins, so we did not carry out sSVA analysis, but adjusted for the same covariates as in KOBS.

Statistical analysis for WGCNA, gene correlations, and expression-trait associations

Statistical analyses were performed in R. We transformed raw TPM to log2(TPM+1) and then performed empirical Bayes-moderated linear regression implemented in the WGCNA package [97] (function *empiricalBayesLM*) to correct for covariates while retaining the variation due to the trait of interest. We calculated pairwise gene correlation using biweight correlation allowing a maximum of 5% outliers, and subsequently built a signed network using the soft threshold power of 12. The eigen-gene of each module was calculated and used for trait association tests. To test the module preservation in GTEx, we re-processed the RNA-seq raw reads using our pipeline, the same QC and genes expressed in both KOBS and GTEx. A module with a preservation summary Z-statistics >10 was considered as strongly preserved [98]. Pair-wise gene expression correlation

between *OLMALINC* and all other genes were calculated using biweight correlation and the adjusted TPMs. We used linear and logistic regression in all trait association tests where the adjusted gene expression level and trait were treated as dependent and independent variables, respectively. The quantitative traits were adjusted for age and sex as well as inverse normal transformed to avoid outlier effects.

Cell culture

We maintained HepG2 (ATCC) and Fa2N4 (Xeno Tech) cells in a monolayer culture at $37^{\circ}C$ with 5% CO^2. The base medium was EMEM for HepG2 (Corning) or base media for Fa2N4 (Xeno Tech) containing 100 U/ml penicillin and 100 μg/ml streptomycin sulfate (GE Healthcare Sciences). We tested the cells for mycoplasma contamination using SoutherBiotech Mycoplasma Detection Kit.

Reagents and transfections

For ASO treatment, 0.5 million cells were grown to ~70% confluency in 6-well plates in triplicates (in 10% FBS containing 1g/L of glucose with penicillin/ampicillin). Cells were treated with Opti-MEM (Gibco), Lipofectamine RNAiMax (Invitrogen 13778100) and the ASO (IDT) at a final concentration of 50-100 nanomoles. The control ASO was designed to have similar modifications to the *OLMALINC* ASO. Cells were transfected at a final concentration to 30 pmoles for siRNAs. ASO and siRNA sequences are provided in Supplemental Tables 3S and 4S. For plasmid transfections, we used Lipofectamine 3000 (Invitrogen) with 2 ug of DNA. For the time point experiments, cells were incubated overnight in 0.25% BSA (Simga) followed by treatment in corresponding conditions outlined in the figures [99]. We obtained Lipoprotein deficient medium (LPDS) from Kalen Biomedical LLC; Simvastatin sodium salt from Calbiochem dissolved in

DMSO; and GW 3965 and mavelonic acid were purchased from Sigma Aldrich. Oleic acid was purchased from Sigma Aldrich. For cellular localization experiments, we used the PARIS Kit (Invitrogen). GFP control and *OLMALINC* cDNA plasmids were obtained from GeneCopoeia.

RNA purification, cDNA synthesis and real time quantitative PCR (RT-qPCR)

We harvested cells in TRIzol (Invitrogen) and extracted their RNA using Direct-Zol (Zymo Research) according to the manufacturer's protocol. We synthesized cDNA using the Maxima First Strand cDNA Synthesis Kit (Thermo Scientific). RT-qPCR was performed using SYBRGreen reaction mix (Applied Biosystems) and Studio 5 detection system (Applied Biosystems). *36B4* was used as an internal control to normalize the data. The primer list is provided in Supplemental Table 2S.

Conservation and synteny of OLMALINC

To study the conservation of the *OLMALINC* locus, we used the NCBI HomoloGene and the mouse and human ENSEMBL data. We evaluated the conservation of *OLMALINC* between human and mouse by aligning DNA segments sequentially between the mouse and the human using blast (GRCh37/hg19), utilizing the blastn function: word size 11, expected threshold 10, match score of 2, and mismatch score of -3. We also used the mouse ENCODE data (Mouse mm10) to identify RNA polymerase II and histone methylation markers.

Promoter Capture Hi-C

We performed the promoter Capture Hi-C in 2 biological replicates of 10 million HepG2 cells [100]. The libraries were sequenced on the Illumina HiSeq 4000 to obtain ~114 million paired-end reads. The reads were processed as described in [101], using HiCUP [102] v0.7.2 software and aligning to the

GRCh37/hg19 [102]. Significant interactions were identified using the CHiCAGO software [103] v1.1.1.

GRO-sequencing

GRO-seq libraries were prepared according to previously described protocols in HepG2 cells (10% FBS) [104,105]. Illumina HiSeq 2000 platform was used to sequence the libraries after size selection (180-350 bp). After quality control, the data was aligned using the GRCh37/hg19. The GRO-Seq data is accessible under GEO accession GSE92375.

Activating CRISPR dCas9 stable cell lines

To generate the activating CRISPR dead Cas9-VP64 (aCRISPR dCas9) stable cell lines, we used the pHAGE EF10apha dCas9-VP64 (Addgene #50918) plasmid [106]. Cells were transduced with polybrene (1μg/ml) for 2-3 days followed by selection with 4ug/ml of puromycin for 7 days. Single clone isolation was obtained following the serial dilutions. Clones expressing the dCas9 were confirmed by RT-qPCR of the dCas9 gene. We used two *OLMALINC* guide RNAs (gRNAs) targeting the promoter region of *OLMALINC* [107]. gRNAs were obtained from VectorBuilder (Shenandoah, TX).

CRISPR Cas9 of the OLMALINC enhancer/promoter region

Using IDT Alt-R CRISPR-Cas9 genome editing tools, gRNAs were designed to flank the enhancer/promoter region of *OLMALINC* which was identified using ENCODE, GRO-seq, and promoter Capture Hi-C. Four gRNAs were used to identify the most efficient gRNAs (supplemental table 3S). RNA protein complexes were prepared using Alt-R S.p. Cas9 Nuclease V3 (IDT) with the *OLMALINC* gRNAs. HepG2 cells were transfected with Opti-MEM (Thermo

Fisher Scientific) and Lipofectamine RNAiMax (Invitrogen) for 48 hours. Transfection efficiency was evaluated using light and fluorescent microscopy (Texas Red-X) using the BZ-X710 fluorescent microscope. The FACAriaII cytometer was used to quantify the efficiency of transfection using FACDiva Version 8.0.2. HepG2 genomic DNA was extracted using PureLink Genomic DNA extraction kit (Thermo Fisher Scientific). The PCR of the genomic DNA was conducted using primers flanking the gRNA cut sites to detect efficiency of all clones, as well as to amplify regions within the *OLMALINC* wild type. These were confirmed using RT-qPCR.

Western blots

Cells were washed and lysed in 1X Laemmli SDS sample buffer (Alfa Aesar). Lysates were seprated by SDS-PAGE (4-15% polyacrylamide) pre-cast gels (BioRad) overnight, transferred to a polyvinylidene difluoride membrane (Immobilon, Millipore Corp.), and blocked for 1 hours in 5% blocking solution (Biorad). The membrane was incubated in 1:1000 primary SCD antibody (ThermoFisher) overnight at 40C followed by washes in and 1:1000 secondary mouse antibody for 45 min. The membrane was washed after which immunoreactive proteins was detected using chemiluminescence (Biorad). Beta-actin and secondary mouse antibodies was kindly provided by Dr. Enrique Rozengurt's laboratory as a loading control.

Statistical methods of the cellular data

For the *in vitro* HepG2 and Fa2N4 experiments, numeric outcomes are summarized as means ± standard deviation (SD) or ± standard error of the mean (SEM). All relative expression values were measured using the $^{\Delta\Delta}$Ct. Experimental groups were compared using unpaired Student's *t* test (for two groups). Analyses were performed using GraphPad Prism version 7.0c. Statistical significance was defined as $P < 0.05$. Graphs were made in GraphPad Prism and assembled in Inkspace.

Results:

Identification of *OLMALINC* in the statin- and triglyceride-associated liver co-expression network:

To identify new genes involved in central liver functions, we performed a weighted gene co-expression network analysis (WGCNA) on the liver transcriptomes from 259 participants (40% statin users) in the KOBS bariatric surgery cohort and tested the association of co-expression modules with statin use, serum TGs, and other metabolic and liver histology phenotypes measured in this cohort. Thirteen of the 19 co-expression modules were significantly associated (FDR<0.05) with at least one of the clinical or histological traits (Figure 1A), including the light cyan module (75 genes) that was significantly associated with statin use (FDR=2.0×10^{-15}) and serum TGs (FDR=7.7×10^{-5}), among other traits (Figure 1A). We validated the module preservation in an independent human liver RNA-seq cohort, Genotype-Tissue Expression (GTEx), by investigating the GTEx subjects whose causes of death were not liver diseases (n=96). Most trait-associated liver modules, such as the statin- and TG-associated light cyan module, were either preserved (Z score>3) or highly preserved (Z score>10) in the GTEx livers (Supplementary Figure 1), respectively, suggesting that gene co-regulation related to main liver functions is robust and consistent across human cohorts. Notably, we observed that the 75 genes in the statin- and TG-associated light cyan network module (Figure 1A-B) comprise 19 known cholesterol pathway genes, 33 fatty acid and metabolic pathway genes, and several potentially novel statin response and TG genes, including the lincRNA, *OLMALINC*. In line with its statin and TG associations, this light cyan module was enriched for the steroid biosynthesis pathway, fatty acid metabolism, and other metabolic pathways (FDR<0.05) (Supplementary Figure S2) using the KEGG pathway database.

Since the lincRNA O*LMALINC*, identified in the lightcyan module, resides immediately downstream from the main TG metabolism gene, stearoyl-CoA desaturase (*SCD*), on human chromosome 10, and given that lincRNAs often regulate adjacent coding genes [108], we next individually tested the correlation of the *OLMALINC* liver expression with *SCD* and detected a significant correlation of ß=0.44 with FDR=4.57×10^{-11} (Supplemental Table 1S). We observed that *SCD* in turn resides in another WGCNA network, the midnight blue module, that is strongly associated with serum TGs (FDR=2.7×10^{-9}) and liver steatosis (FDR=5.9×10^{-7}) (Figure 1A).

Next, we followed up the *OLMALINC* and *SCD* co-expression findings and their mutual associations. We first tested if the liver expression of *OLMALINC* is individually associated with statin usage. When counting for multiple testing of the 75 genes in the lightcyan module using Bonferroni (which is a conservative approach because these co-expressed module genes are not entirely independent), *OLMALINC* was nominally associated with statin use (p=0.0035, Figure 1B). Thus, the statin users appear to have a higher *OLMALINC* liver expression than the non-users in the KOBS cohort, fully supported by our *in vitro* statin response results in HepG2 cells (see below; Figure 3). Similarly, *SCD* liver expression was also higher in the statin users of the KOBS cohort (p=0.0027), again in line with our *in vitro* HepG2 results (see below; Figure 3). We also detected a significant association between *OLMALINC* liver expression and fasting serum TGs in the KOBS cohort (ß=0.27, p=0.001, passing the Bonferroni correction for 7 traits, Supplemental Table 2S). In line with this observation, *SCD* liver expression was significantly associated with serum TGs (ß=0.48, p=0.13×10^{-7}) in the KOBS cohort as well. Finally, although *OLMALINC* was not associated with steatosis or other liver histology traits (Supplemental Table 2S), *SCD* liver expression was associated with liver steatosis (ß=0.35; p=0.0054) but not NASH (ß=0.27; p=0.107). Taken together these novel data suggest the possibility that *OLMALINC* regulates its

54

adjacent regional protein coding gene, *SCD*, which is likely the driver in liver steatosis among the two, while both genes are associated with serum TG levels and respond to the statin use. To further investigate this new hypothesis that *OLMALINC* regulates *SCD*, we performed functional genomics studies, as described below.

To assess *OLMALINC* gene expression in other human tissues, we analyzed the RNA-seq data from the GTEx project and found that *OLMALINC* is ubiquitously lowly expressed, as expected from a lincRNA. After the brain, the most abundant *OLMALINC* expression can be seen in the liver and other endocrine/hormone-regulated organs (Supplemental Figure S3).

Overview of our functional genomic approaches to study *OLMALINC* in lipid metabolism:
We aimed to study the function of *OLMALINC* by utilizing molecular genomics approaches (Supplemental Figure S4). Since the chromosomal location of *OLMALINC* is directly downstream of *SCD* (see below), we first demonstrate that *OLMALINC* is an enhancer of *SCD* transcription by forming a DNA-DNA looping interaction (Supplemental Figure S4A). This was confirmed by CRISPR-Cas9 genetic deletion of this region (Supplemental Figure S4B) and endogenous transcriptional over-expression using the activating CRISPR-dead Cas9 (aCRISPR-dCas9) gene editing system (Supplemental Figure S4d). To complement our CRISPR-Cas9 gene editing, we confirm that *OLMALINC* positively regulates *SCD* expression (Supplemental Figure S4C) by using an ASO which preferentially localizes to the nucleus. We further show that *OLMALINC* expression increases with *SCD* siRNA (Supplemental Figure S4E) but decreases with oleic acid treatment, a by-product of SCD enzyme activity.

***OLMALINC* is statin, sterol, and LXR responsive:**

Using data from the ENCODE project and chromatin immunoprecipitation sequencing (ChIP-seq) from HepG2 cells, we found two active transcription start sites (TSSs) characterized by a RNA polymerase II binding site, a 5' CAGE peak, and active histone modification markers characteristic of enhancer- and promoter-elements, in the *OLMALINC-SCD* region (Figure 2A). GRO-seq data in HepG2 cells, used to assess nascent RNA, not only confirms two active TSSs in the enhancer and promoter of *OLMALINC*, but also demonstrates bi-directional transcription, suggesting that *OLMALINC* could function as an enhancer to *SCD* (Figure 2A). Using the ENCODE project data, we identified SREBP1 and SREBP2 ChIP-seq sites at the *OLMALINC* TSSs (Figure 2B). We hypothesized that *OLMALINC* expression would be statin and sterol responsive, based on our correlative results from the liver RNA-seq data in the KOBS cohort. Using RT-qPCR, we demonstrate that *OLMALINC* expression increases with statin and sterol treatments in a time-dependent manner, demonstrating that it is both a sterol- and statin- responsive gene in HepG2 (Figure 3A). Although there is a similar trend in the non-cancerous Fa2N4 cell line none of the increases show a significant increase (Figure 3B). These data are consistent with the human liver RNA-sequencing results in the KOBS cohort, which demonstrate a positive correlation of *OLMALINC* with liver cholesterol gene expression and a membership in the statin module of our WGCNA analysis (Figure 1; Supplemental Table 1S). We also show that *OLMALINC* expression is LXR-responsive as cells treated with the synthetic liver LXRα and LXRβ agonist, GW3965, increase *OLMALINC* expression (Figure 3, C-D). We identified an LXR responsive element (LXRE-DR4)T(G/A)A(C/A)C(T/C)XXXXT(G/A)A(C/A)C(T/C) in the *OLMALINC* promoter (Supplemental Figure 5S). This is consistent with *OLMALINC* having a retinoid X receptor alpha (RXRα) ChIP-seq binding site, which forms a heterodimer with LXRα and LXRβ to activate transcription (Figure 2B), suggesting thus a direct role of LXR in regulating *OLMALINC* liver expression.

***OLMALINC* function:**

To study *OLMALINC* function, we analyzed its cellular localization, which did not demonstrate a significant difference between the cytoplasmic and nuclear extracts for exons 1-2 (RT-qPCR of exons 2-3 demonstrates a preferential cytoplasmic expression of the stable transcript) (Supplemental Figure S6B). All subsequent RT-qPCR data that we present were conducted by measuring exons 1-2 (shared between the identified isoforms). A ~50% knock-down of *OLMALINC* by an ASO (of exon 2) resulted in a decrease in *SCD* expression (Figure 4A; Supplemental Figure S4C). Conversely, when *SCD* is knocked down, we observed an increase in *OLMALINC* expression (Figure 4B-C). These data suggest that *OLMALINC* expression is responsive to *SCD* expression, its protein level, or the monounsaturated fatty acid (MUFA) byproducts. Given that *SCD* resides upstream of *OLMALINC* and previous observations that lincRNAs can regulate genes in *cis*, we hypothesized that *SCD* is regulated locally by *OLMALINC* in *cis*.

The *cis* effects of *OLMALINC* on *SCD* expression:

OLMALINC resides directly downstream of *SCD*, the microsomal enzyme that converts polyunsaturated fatty acids into monounsaturated fatty acids (MUFAs). *OLMALINC* liver expression is significantly correlated with *SCD* expression (ß=0.44; FDR=4.57E-11, Supplemental Table S1) and serum TGs (Supplemental Table S2), suggesting a role for *OLMALINC* in TG regulation. The chromosome 10 region of *OLMALINC* and *SCD* in humans has synteny with chromosome 19 of the mouse genome where *WNT8B*, *SCD1*, *SCD2*, *SCD3* and *SCD4* are localized in a ~330 kb region (Supplemental Figure S7). However, no orthologues of *OLMALINC* were

57

identified in the mouse. Consistent with these findings, no histone methylation markers or RNA polymerase II ChIP-seq sites were found in the mouse genome between *WNT8B* and *SCD1* to suggest a TSS (Supplemental Figure S8). Similar to other lincRNAs, *OLMALINC* only shows a high homology in primates [109].

Since lincRNAs often exert their function by affecting adjacent genes, we hypothesized that *OLMALINC* may regulate *SCD* expression in *cis* by acting as an enhancer. To further investigate this, we performed promoter capture Hi-C in liver HepG2 cells (in 10% FBS) and identified a DNA-DNA looping interaction between the promoter of *SCD* and the annotated promoter/enhancer of *OLMALINC* (Figure 5A, Supplemental Figure S4A). This interaction is cell-type specific given that no interaction was identified between *SCD* and *OLMALINC* in human adipocytes despite the high *SCD* adipocyte expression [100]. These promoter capture Hi-C interaction data suggest that *OLMALINC* acts via looping in *cis* to affect transcription of *SCD*. It is worth noting that since *OLMALINC* and *SCD* have a bi-directional promoter (Figure 2B), it is possible that the looping interaction is strand-specific; however, only the positive strand was interrogated when targeting the promoter for CRISPR Cas9 (see below).

To further investigate the *cis* local regulatory effects, we used an activating CRISPR dead Cas9-VP64 (aCRISPR dCas9) to over-express *OLMALINC* endogenously using previously validated gRNAs in a constitutively expressing dCas9 cell line [107,110]. By RT-qPCR, we demonstrate that a ~1.8-fold increase in *OLMALINC* expression resulted in a 2-fold increase in *SCD* expression (Figure 5B, Supplemental Figure S4D).

To further tease out the local transcriptional versus post-transcriptional effects of *OLMALINC* regulation, we investigated the effects of its transcript on *SCD* expression. *OLMALINC* is annotated to have several transcripts (data not shown). Expression of a stable transcript with 3 exons was confirmed by Sanger sequencing of the PCR products (Supplemental Figure S6A) and alignment analysis of the liver RNA-seq (data not shown). When the mature *OLMALINC* transcript is overexpressed using a cDNA construct (exons 1-3), we observed no downstream effects on *SCD* gene expression (Figure 5C). In conjunction with the endogenous over-expression data (aCRISPR dCas9), our results confirm that *SCD* regulation by *OLMALINC* occurs at the transcriptional level, likely through the *cis* effects.

To target the *cis* effects of *OLMALINC* on *SCD*, we used CRISPR-Cas9 gene editing to delete the ~3.5 kb region of *OLMALINC*, which encompasses the SREBP1/2 binding sites, TSSs, LXR element, and the capture Hi-C looping interactions (Figure 6A-C). Using a fluorescently-labeled tracrRNA, we determined that our transfection efficiency of the HepG2 cells was 84% (Supplemental Figure S9), thus showing success in targeting the majority of the cells. The cells demonstrate ~50% decrease in *OLMALINC* expression, which causes a decrease in *SCD* expression (Figure 6D, Supplemental Figure S4B). Whether the *SCD* expression effects are specific to disruption of DNA-DNA interactions between *SCD* and *OLMALINC* encompassing the promoter/enhancer region or are a by-product of large DNA deletions remains to be tested. *Wnt8B*, the gene downstream of *OLMALINC*, is not expressed in human liver, as confirmed by the GTEx cohort and our RT-qPCR data in HepG2 cells (data not shown), thus ruling out a *Wnt8B*-specific effect. Taken together, our detailed functional genomic manipulation of *OLMALINC* expression (over-expression at the transcriptional level using aCRIPR-dCas9, over-expression post-transcriptionally using the mature cDNA transcript, and knocking down *OLMALINC* RNA via

CRISPR-Cas9 and ASO) show that *OLMALINC* regulates *SCD* expression in *cis* as an enhancer, likely via looping interactions.

***OLMALINC* regulation:**

In conjunction with the ENCODE data, we demonstrated that *OLMALINC* is sterol-, statin- and LXR- responsive (Figures 2 and 3). Given the *cis* effect of *OLMALINC* on *SCD* and the known regulation of *SCD* by the SREBP1 and SREBP2 pathways [111], we sought to further understand *OLMALINC* regulation by these transcription factors. To accomplish this, we knocked down *SREBP1* and *SREBP2* using siRNAs to study those effects on *OLMALINC* expression. We observed that knock-down of *SREBP2* or *SREBP1* alone does not affect *OLMALINC* expression or SREBP1/2 dependent genes, likely from compensatory effects of the SREBPs (data not shown). However, when both SREBP1 and SREBP2 siRNAs are used in conjunction, their target genes, including *SCD*, are decreased, while *OLMALINC* expression does not decrease (Figure 7A). We therefore hypothesized that *OLMALINC* expression is regulated by SCD by-products, which are MUFAs. To test these, we treated HepG2 cells with the MUFA oleic acid at different time points and show that *OLMALINC* expression decreases with oleic acid treatment (Figure 7B), which is consistent with the observed increase in *OLMALINC* expression when knocking down *SCD* (Figure 4B). We observe that *OLMALINC* gene expression decreases early (18 hours) prior to seeing an effect on *SCD* gene expression, which occurs later at 24 and 48 hours of treatment (Figure 4B), when we also see a decrease in *SREBP1a* and *SREBP1c*. These data suggest that *OLMALINC* senses and mediates *SCD* gene expression locally early before SREBP1 transcription factor proteins can regulate *SCD* expression. This is in line with our finding that the *OLMALINC* expression is positively correlated only with serum TGs and not with the other phenotypes in the KOBS cohort (Supplemental Table 2S).

60

Discussion:

In the present study, we combined human liver transcriptomic and *in vitro* experimental data to identify and characterize the lincRNA, *OLMALINC* in lipid metabolism. We first detected *OLMALINC* in tight correlation with known lipid genes in human liver RNA-seq data, and then demonstrate that our human correlative expression data translate to important effects of *OLMALINC* on a key triglyceride gene, *SCD*. Our study also describes the first eRNA in lipid metabolism as our data show that *OLMALINC* regulates the *SCD* gene in *cis*. Specifically, we observed that *OLMALINC* regulates *SCD* at the transcriptional level in *cis* by forming a looping interaction with the *SCD* enhancer/promoter region at important DNA elements where transcription factors and enhancers can interact and activate gene transcription. Furthermore, as *SCD* encodes an enzyme involved in fatty acid biosynthesis, including the synthesis of the monounsaturated fatty acid oleic acid [79], it is noteworthy that in our context-specific lipid loading experiments, *OLMALINC* expression is responsive to the SCD byproduct oleic acid early, independently of SREBP1, prior to seeing changes in *SREBP1a/c* which occurs later. This suggests that *OLMALINC* may have evolved through an independent mechanism to sense and fine-tune *SCD* gene expression early given its proximity to the gene, perhaps to maintain the important monounsaturated fatty acid homeostasis.

Cellular cholesterol and lipid homeostasis are tightly regulated to maintain essential lipid-related processes in the human membrane [111]. Important feedback mechanisms are in place to preserve homeostasis at the transcriptional, post-transcriptional and protein level, partly through the SREBP transcription factors, which are the master regulators of cellular lipid and cholesterol processes, with SREBP1c preferentially activating the fatty acid synthesis pathway [111-113]. Recent studies have demonstrated the role of lincRNAs in regulating and helping regulate SREBPs in their

functions [85]. For instance, MALAT1, the nucleus specific lincRNA, inhibits degradation of SREBP1c protein by preventing its ubiquitination in the nucleus [114]. Similarly, the lncRNA H19 stabilizes SREBP1c both at the transcript and protein levels, depending if it exerts its function in the cytoplasm or nucleus, respectively [115]. In our current study, we demonstrate that *OLMALINC* acts as an enhancer for *SCD* and regulates *SCD* expression through sensing of its by-products early prior to SREBP1-dependent effects.

Patients with NASH and NAFLD have previously been shown to exhibit altered cholesterol and triglyceride metabolism [77,80]. Since the majority of the participants in the KOBS cohort have some form of NAFLD, it is possible that the statin-associated co-expression module we identified in the WGCNA analysis may also reflect the primary effect that NAFLD and NASH have on cholesterol metabolism. However, the correlative WGCNA data cannot alone separate these two possibilities. As SCD and SREBP2 have been shown to be dysregulated in NAFLD and NASH [76,77,80] future studies are warranted to elucidate the role of *OLMALINC* in cholesterol metabolism perturbed by NASH.

Recent studies have demonstrated that lncRNAs affect nearby coding gene expression similarly to the effects of *OLMALINC* on *SCD* expression [116]. Through detailed transcriptional analyses, it has also been elucidated that the effects on the near-by genes by lncRNAs are not necessarily mediated through the transcript but rather by transcriptional regulation (through enhancers and promoters) and/or splicing machinery [87]. In addition to the important enhancer/promoter region via which *OLMALINC* affects *SCD*, we show that *OLMALINC* has a stable, spliced and polyadenylated transcript. Given that enhancers generally produce unstable transcripts without a poly-A tail or

splicing [117], *OLMALINC* likely has a secondary function on other targets independently of its *cis* effects on *SCD* expression, which remains to be elucidated.

Consistent with the importance of SCD in metabolic disorders, patients with NASH demonstrate an increased *SCD* expression in the liver [80]. Plasma oleate to stearate (18:1/18:0) and palmitoleate to palmitate (16:1/16:0) ratios, which are used as surrogates for systemic *SCD* activity, are also increased in patients with the MetS and NASH, supporting an increase in SCD activity [81]. These data are corroborated by recent clinical trials targeting SCD protein in NASH (n=58) and HIV (who also develop hepatic steatosis) (n=25) patients, which demonstrates reversal of hepatic steatosis with treatment [84,118]. In agreement with the human data, $SCD^{-/-}$ mouse models are protected from adiposity, have decreased *de novo* lipogenesis and increased fatty acid oxidation [82]. It has also been shown that repletion of oleate through dietary supplementation in global and liver-specific *SCD* knock-out murine models prevents hepatic ER stress and inflammation [119]. Given these findings, it would not be surprising for a lincRNA to have evolved to maintain MUFA homeostasis and provide another layer of early regional regulation to *SCD* gene expression epigenetically via chromosomal looping of this adjacent coding gene. Although far from therapeutic considerations, further understanding of *OLMALINC* function opens up unexplored avenues for gene modification and treatment considering its cell- and tissue-specificity.

The present study highlights a novel lincRNA, *OLMALINC*, that affects a key TG gene by affecting *SCD* expression in *cis* as a regional eRNA. *OLMALINC* joins a group of lipid lincRNAs that have been described and continue to emerge in lipid homeostasis and pathology [88]. In addition to their role in regulating important coding genes, they could be one of many factors that explain the cross-species differences in lipid metabolism. Further unraveling of their biology will provide insight

into new cellular mechanisms and may pave the way for better understanding of complex cardiometabolic disorders in humans.

a

|	Steatosis	Fibrosis	Lobular inflammation	Ballooning	D1	DM	Triglycerides	Cholesterol	LDL cholesterol	HDL cholesterol	Cholesterolmed	Fasting glucose	ALT
yellow:1390	0.0088 (0.051)	0.013 (0.03)	0.0067 (0.27)	0.007 (0.32)	−0.26 (0.077)	−4e−04 (0.48)	0.01 (0.024)	0.0033 (0.38)	0.002 (0.43)	5e−04 (0.47)	−0.0054 (0.37)	0.00042 (0.48)	0.002 (0.41)
turquoise:6710	0.0063 (0.15)	0.01 (0.079)	0.0097 (0.17)	0.0035 (0.42)	−0.24 (0.091)	0.018 (0.058)	0.0057 (0.18)	0.011 (0.074)	0.013 (0.074)	−0.0038 (0.3)	0.0065 (0.33)	−0.0043 (0.37)	−0.00078 (0.46)
tan:90	−0.0079 (0.077)	0.0015 (0.45)	0.0017 (0.45)	−0.014 (0.13)	0.1 (0.32)	−0.018 (0.053)	−0.0078 (0.08)	0.0034 (0.38)	0.011 (0.12)	−0.0019 (0.4)	−0.002 (0.45)	−0.0032 (0.4)	−0.0021 (0.41)
salmon:83	−0.017 (9.6e−05)	−0.014 (0.023)	−0.026 (3e−04)	−0.023 (0.017)	0.52 (0.00015)	−0.019 (0.041)	−0.0093 (0.041)	0.008 (0.19)	0.013 (0.068)	0.0041 (0.27)	−0.0078 (0.31)	−0.011 (0.14)	−0.0099 (0.053)
red:246	−0.0016 (0.42)	0.0079 (0.17)	−0.0035 (0.39)	−0.0029 (0.43)	−0.021 (0.46)	−0.0045 (0.39)	−0.0021 (0.39)	−0.005 (0.31)	−0.0063 (0.28)	0.0052 (0.21)	−0.0074 (0.31)	0.0055 (0.32)	0.0037 (0.32)
purple:112	−0.011 (0.012)	−0.0097 (0.1)	−0.023 (0.0012)	−0.017 (0.079)	0.41 (0.0037)	−0.025 (0.0049)	−0.0034 (0.31)	0.0082 (0.19)	0.011 (0.12)	0.0057 (0.18)	−0.017 (0.074)	−0.0086 (0.21)	−0.0085 (0.098)
pink:154	0.0023 (0.38)	0.0044 (0.32)	0.006 (0.31)	0.00074 (0.48)	−0.1 (0.32)	0.0055 (0.36)	0.0032 (0.32)	0.015 (0.025)	0.018 (0.013)	−0.0015 (0.43)	−0.0028 (0.27)	−0.007 (0.27)	−0.0012 (0.45)
midnightblue:76	0.021 (5.9e−07)	0.017 (0.0036)	0.021 (0.0042)	0.025 (0.01)	−0.61 (7.5e−06)	0.021 (0.025)	0.024 (2.7e−09)	0.014 (0.019)	0.008 (0.18)	0.0013 (0.43)	−0.00099 (0.47)	0.015 (0.025)	0.0012 (0.44)
magenta:132	6e−04 (0.46)	0.012 (0.044)	0.0036 (0.39)	−0.0013 (0.46)	−0.11 (0.31)	−0.00062 (0.48)	−0.00081 (0.48)	−0.0047 (0.29)	−0.0036 (0.36)	0.0013 (0.43)	−0.0046 (0.24)	−0.007 (0.25)	0.0012 (0.45)
lightgreen:32	0.0032 (0.32)	0.0062 (0.24)	0.011 (0.11)	0.0049 (0.39)	−0.18 (0.18)	−0.0045 (0.39)	3e−06 (0.5)	−0.0017 (0.44)	0.002 (0.43)	−0.0046 (0.24)	−0.0053 (0.37)	−0.0079 (0.25)	0.0053 (0.24)
lightcyan:75	0.0095 (0.036)	0.02 (0.00036)	0.015 (0.043)	0.012 (0.19)	−0.43 (0.0023)	0.05 (3.2e−09)	0.017 (7.7e−05)	0.0011 (0.45)	−0.0041 (0.31)	−0.0044 (0.25)	0.066 (2e−15)	0.0014 (0.45)	−0.00025 (0.48)
grey60:51	−0.00064 (0.46)	0.00087 (0.46)	−0.0085 (0.21)	−0.0056 (0.37)	0.057 (0.41)	0.0021 (0.45)	0.0036 (0.31)	0.0026 (0.41)	0.0026 (0.41)	0.0012 (0.44)	0.0018 (0.46)	0.0016 (0.45)	−0.0047 (0.27)
grey:1717	−0.0016 (0.42)	0.0037 (0.35)	−0.0051 (0.32)	−0.0075 (0.31)	0.027 (0.46)	0.0026 (0.44)	0.0017 (0.42)	0.009 (0.15)	0.012 (0.089)	−0.00038 (0.48)	−0.000033 (0.49)	−0.0038 (0.38)	−0.0032 (0.35)
greenyellow:92	−0.0046 (0.24)	−0.0062 (0.24)	−0.0066 (0.28)	−0.01 (0.24)	0.16 (0.22)	0.0016 (0.46)	−0.00063 (0.46)	0.016 (0.015)	0.022 (0.0016)	−0.0034 (0.32)	0.0029 (0.43)	−0.0074 (0.25)	−0.0067 (0.17)
green:374	0.017 (5e−05)	0.03 (8.4e−08)	0.04 (2.7e−09)	0.029 (0.0027)	−0.78 (2.7e−09)	0.03 (0.00073)	0.0062 (0.15)	0.0068 (0.21)	0.0087 (0.15)	−0.0066 (0.13)	0.011 (0.2)	−0.012 (0.32)	0.0035 (0.31)
cyan:78	−0.0073 (0.1)	−0.004 (0.34)	−0.015 (0.04)	−0.011 (0.21)	0.25 (0.079)	−0.005 (0.38)	−0.0046 (0.24)	0.0077 (0.21)	0.01 (0.14)	0.0052 (0.21)	−0.00059 (0.48)	−0.012 (0.14)	−0.0076 (0.14)
brown:1678	0.014 (0.0012)	0.029 (1.4e−07)	0.032 (3.5e−05)	0.012 (0.19)	−0.6 (8.4e−06)	0.025 (0.0071)	0.0039 (0.29)	0.0029 (0.4)	0.0052 (0.31)	−0.0068 (0.12)	0.012 (0.12)	−0.0042 (0.34)	0.0076 (0.098)
blue:2418	0.018 (2.1e−05)	0.021 (0.00023)	0.023 (0.0012)	0.019 (0.054)	−0.58 (1.7e−05)	0.021 (0.025)	0.016 (0.00033)	0.0098 (0.11)	0.0084 (0.18)	−0.0035 (0.31)	0.0034 (0.42)	0.0041 (0.37)	0.0047 (0.27)
black:162	0.0095 (0.036)	0.016 (0.0049)	0.028 (9.1e−05)	0.011 (0.21)	−0.45 (0.0012)	0.0012 (0.46)	−0.0014 (0.43)	−0.0028 (0.4)	0.0031 (0.4)	−0.0057 (0.18)	−0.00048 (0.48)	−0.0026 (0.42)	0.0059 (0.21)

−Log10(FDR): 10, 5

b

Scatter plot — x-axis: Module membership (0.5, 0.6, 0.7, 0.8, 0.9); y-axis: −log10(P) (0, 5, 10, 15, 20, 25).

Gene labels: TM7SF2, CYP51A1, TMEM97, FDPS, C14orf1, SQLE, IDI1, ACAT2, HMGCS1, SC5D, MSMO1, LSS, FDFT1, HMGCR, DHCR7, ACSS2, VPS37D, NPC1L1, MVD, CAPN12, RDH11, HSD17B7, MMAB, PCOLCE, FADS1, ALDOC, C2orf82, SREBF2, SLC29A2, PCYT2, STARD4, DERL3, FAM213A, ECHDC1, TIMM10B, DNHD1, LDLR, TMEM170B, AP001468.1, SLC26A6, FADS2, TDRKH, ACLY, BBOF1, ALKBH6, NEK8, POLD2, ACACA, MVK, FLVCR1, NPC1, PRSS8, MFSD14A, PDP2, ELOVL6, PGD, CYB5B, TRIB3, NAT6, PANK3, DHCR24, PCSK9, RP11−297N6.4, SHMT1, AP001065.15, KCNE1, ELOVL5, TMEM41B, SYDE2, SYNE4, MRPS27, OLMALINC, RP11−180D21.3, RP11−344P13.6, PNPLA3

Figure 1- A, Liver weighted gene co-expression network analyses (WGCNA) identify a statin-associated network module (i.e. the light cyan module). The association results between the liver WGCNA modules and statin use, serum triglycerides (TGs), and other metabolic and histological liver phenotypes in the Finnish KOBS cohort (n=259). D1 indicates the aggregated meta-liver trait for NAFLD (see Methods). Numbers in the cells and parenthesis indicate the effect sizes and FDRs, respectively. **B,** Genes in the light cyan module (n=75) are

strongly associated with statin medication and involved in cholesterol synthesis. The strength of

association with statin medication is highly correlated with the module membership of the light

cyan module. The red line indicates the threshold for the Bonferroni corrected p-value of 0.05.
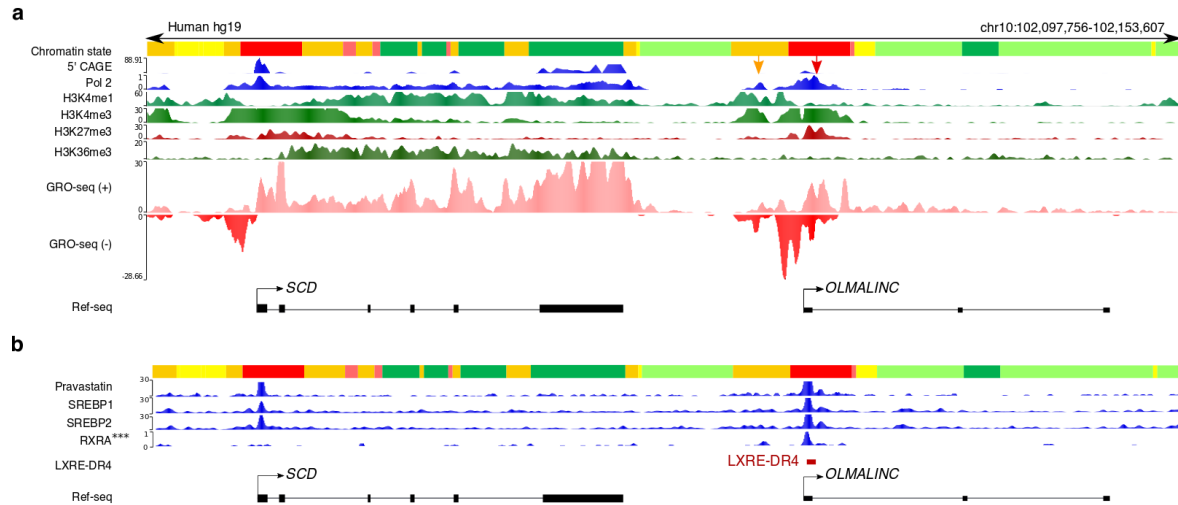
Figure 2- *OLMALINC* **resides downstream of** *SCD* **and demonstrates similar regulatory**

regions. **A**, The annotated *OLMALINC* promoter (red) and enhancer (orange) demonstrate histone

methylation marks, 5' CAGE, and polymerase II ChIP-seq binding sites using the ENCODE data.

There are two transcription start sites (TSSs): The orange arrow denotes the enhancer-TSS, while

the red arrow highlights the promoter-TSS. Our GRO-seq data in HepG2 cells show active

transcription and nascent *OLMALINC* RNA expression bi-directionally. **B**, *OLMALINC* has

SREBP1/2, pravastatin (pravastatin-treated HepG2 cells with SREBP1/2 peaks), and RXRA (***)

binding sites where an LXR element (LXRE-DR4) is identified using sequence comparisons.
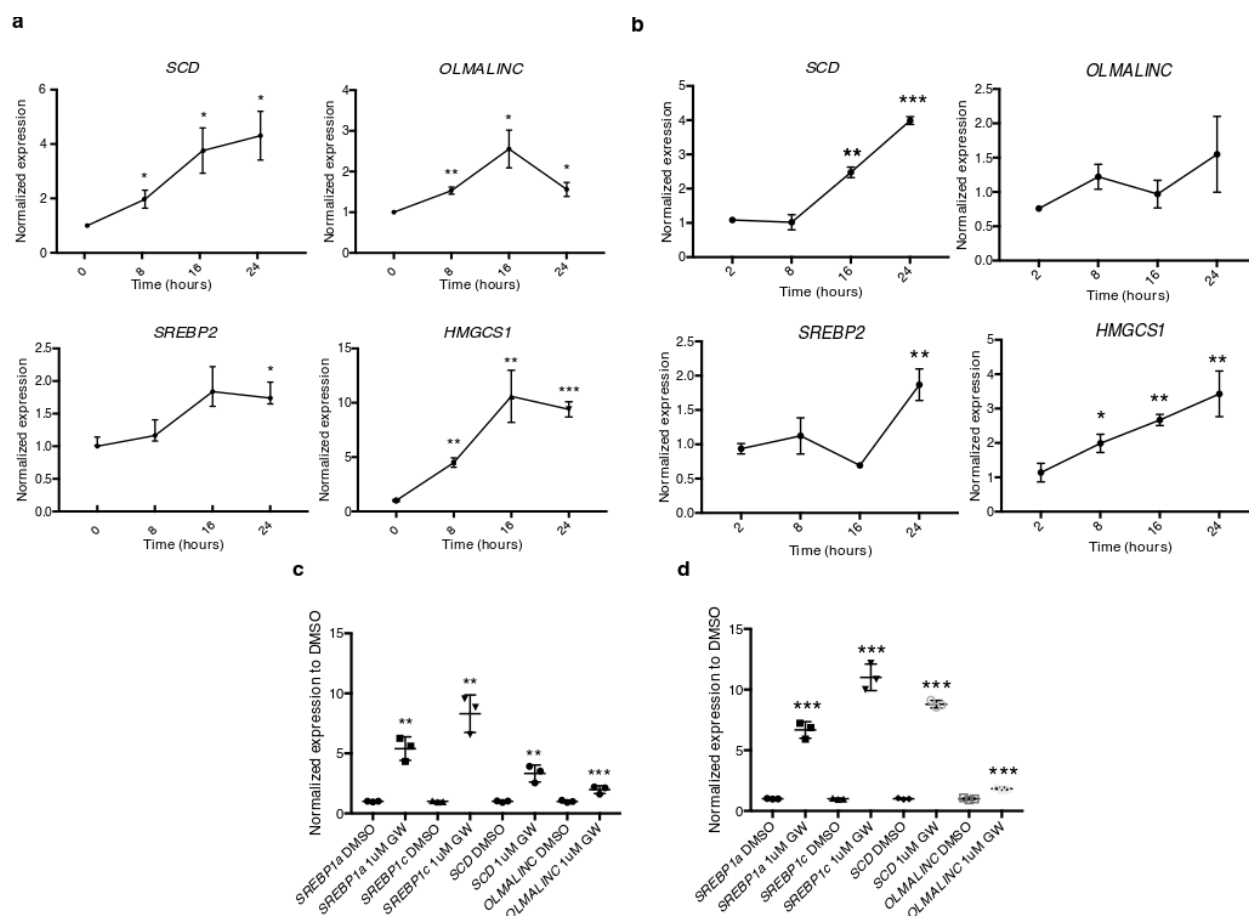
Figure 3- *OLMALINC* **expression is responsive to sterols, statins and LXR agonists in HepG2 (A, C) and Fa2N4 (B, D) cells**. **A-B**, *OLMALINC* and *SCD* increase expression by RT-qPCR in a time-dependent manner under sterol-depleted conditions supplemented with statin treatment (5% lipoprotein deficient media with 5 µM simvastatin and 50 µM mavelonic acid) when compared to sterol rich conditions (10% FBS) supplemented with DMSO vehicle control, similarly to *SREBP2* and its downstream gene *HMGCS1*. Each time point was normalized to its DMSO 10% FBS-treated time point. **C-D**, *OLMALINC* gene expression increases after 24-h treatment of GW3695 (an LXRα and LXRβ agonist) when compared to the DMSO vehicle control in 5% LPDS with 5 µM simvastatin and 50 µM mavelonic acid, as measured by RT-qPCR. Values are mean ± s.d. (n=3) for A and C or mean ± s.e.m. for B (n=3). *P<0.05; **P<0.01, and ***P<0.001 (unpaired Student's *t* test was used for two groups).
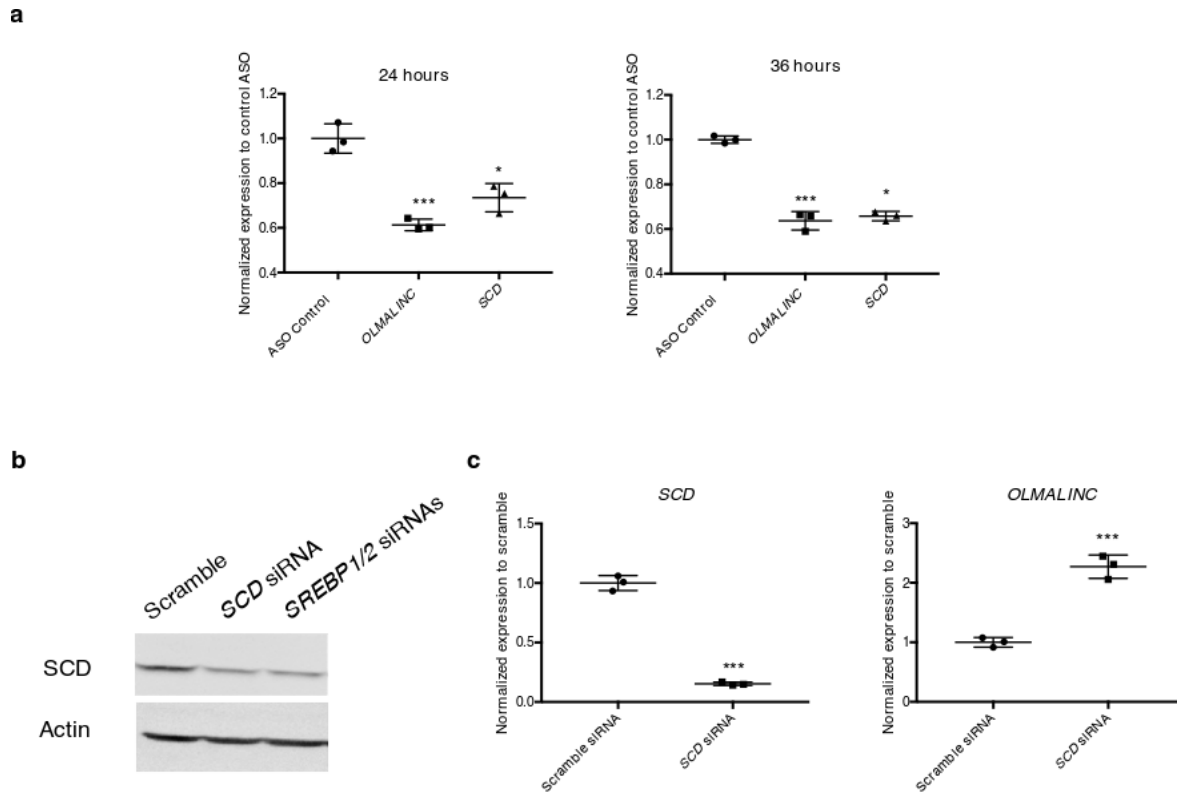
Figure 4- *OLMALINC* **ASO introduced to HepG2 cells cause a decrease in expression of** *OLMALINC* **and target genes**. **A**, *OLMALINC* and target gene expression, measured by RT-qPCR, decrease after 24-h and 36-h treatment with an ASO targeting exon 2 of the *OLMALINC* gene. **B**, Validation of SCD protein antibody (38 kDa) after treatment with scramble, *SCD* and *SREBP1* with *SREBP2* siRNAs after 96-h. **C**, *OLMALINC* gene expression increases after 48-h treatment with an *SCD* siRNA compared to the scramble control. Values are mean ± s.d. (n=3). *P<0.05; **P<0.01, and ***P<0.001 (unpaired Student's *t* test was used for two groups).
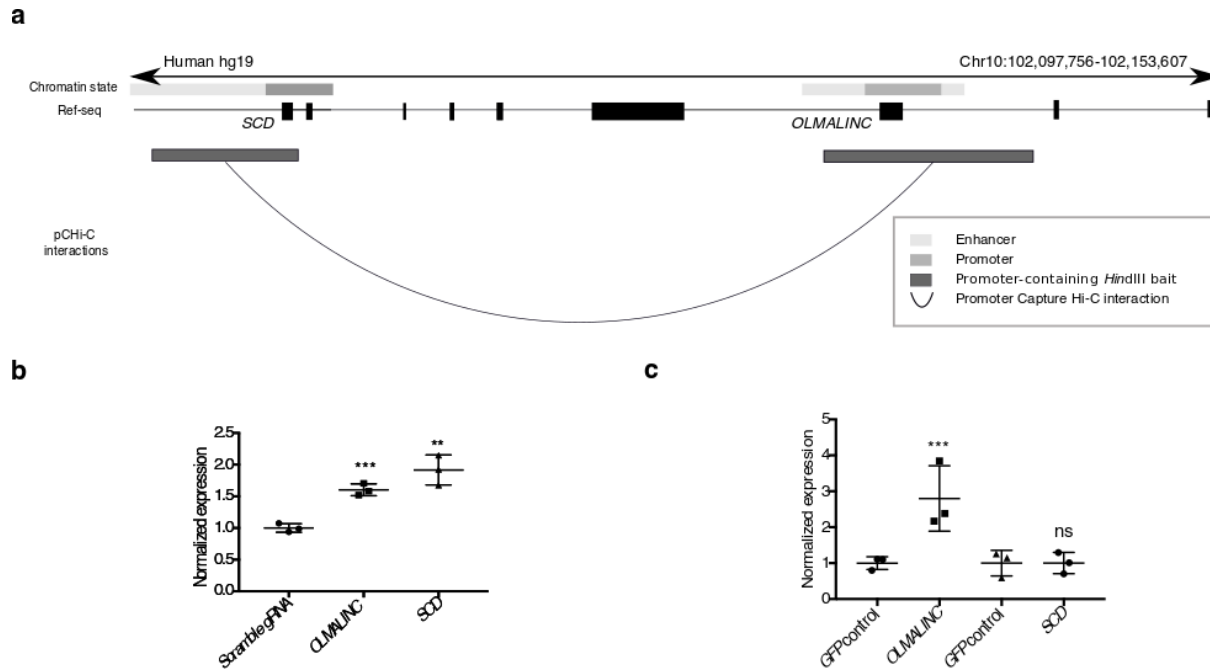
Figure 5- *OLMALINC* **regulates** *SCD* **gene expression in** *cis* **by forming DNA-DNA looping**

interactions. A, Promoter-Capture Hi-C data in HepG2 cells demonstrate DNA-DNA looping

interactions between the *OLMALINC* enhancer/promoter and the *SCD* promoter/enhancer regions.

B, Endogenous *OLMALINC* over-expression using aCRISPR-dCa9 gene editing increases

expression of *SCD*. **C**, Over-expression of the spliced *OLMALINC* stable transcript (exons 1-3)

for 48-h does not affect *SCD* gene expression. Expression data are normalized to a GFP negative

control. Values are mean ± s.d. (n=3). *P<0.05; **P<0.01, and ***P<0.001 (unpaired Student's *t*

test was used for two groups).
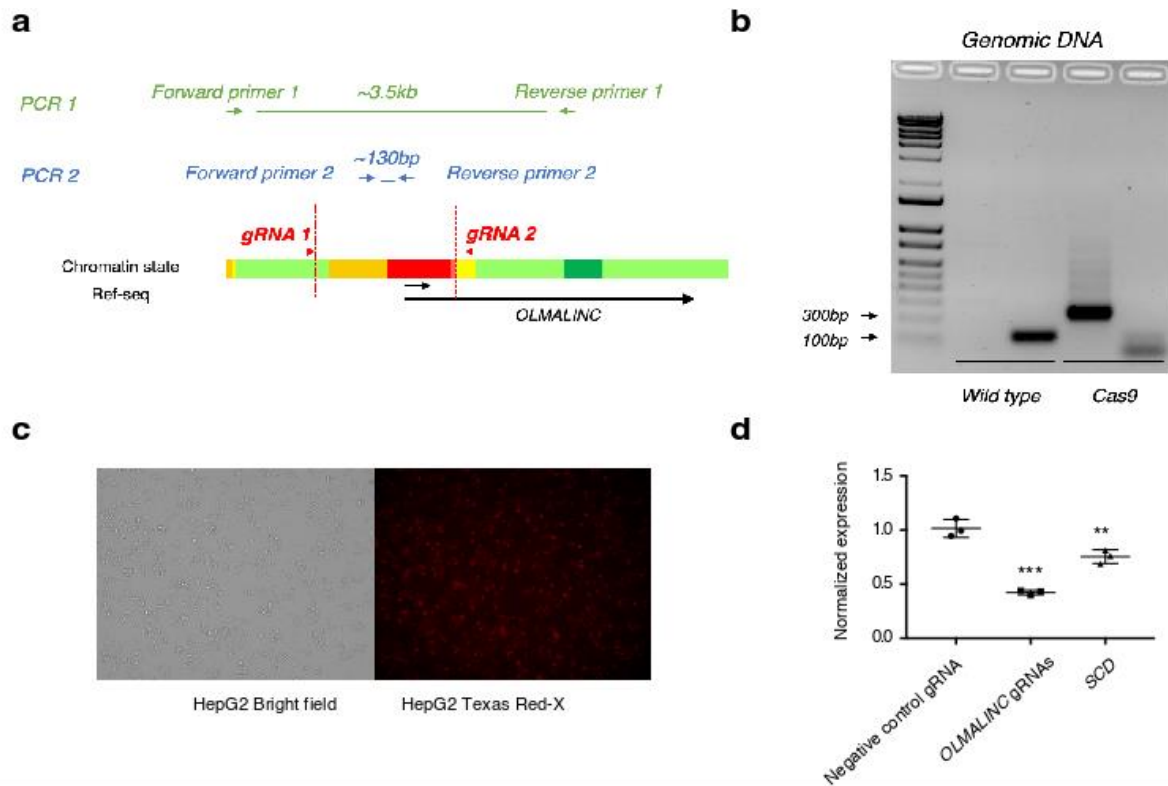
Figure 6- *OLMALINC* **enhancer/promoter deletion using CRISPR-Cas9 gene editing decreases *SCD* gene expression. A**, Schematic of primer designs for genomic PCR amplification of wild type versus CRISPR-Cas9-mediated *OLMALINC* promoter/enhancer deletion. Per ENCODE HepG2 chromatin state data, red highlights *OLMALINC* promoter while yellow highlights the enhancer. **B**, Gel electrophoresis of PCR products from amplification of the wild type and CRISPR-Cas9 *OLMALINC* enhancer/promoter deletions from the genomic DNA from HepG2 cells. **C**, Evaluation of transfection efficiency of HepG2 with fluorescently labeled tracRNA with ATTO-550 after 24-h, with left panel demonstrating bright field cells and right panel the corresponding labeled cells. **D**, *OLMALINC* and *SCD* gene expression by RT-qPCR after 48-h transfection with the Cas9 enzyme and *OLMALINC* gRNAs flanking the enhancer/promoter region. Values are mean ± s.d. (n=3). *P<0.05; **P<0.01, and ***P<0.001 (unpaired Student's *t* test was used for two groups).

71
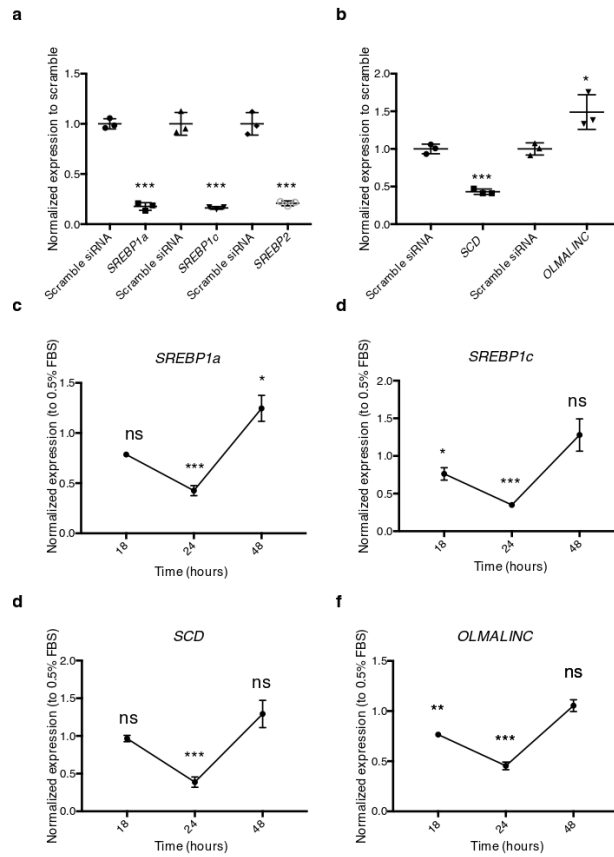
Figure 7- *OLMALINC* is regulated by monounsaturated fatty acids but not by SREBP1/2. A, *SREBP1a*, *SREBP1c,* and *SREBP2* gene expression after *SREBP1* and *SREBP2* siRNA co-transfection for 48 hours, relative to scramble siRNA control. **B**, *OLMALINC* expression does not decrease after a 48-h co-transfection with SREBP1 and SREBP2 siRNAs, while *SCD* decreases. **C-D**, *SREBP1a*, *SREBP1c*, and *SCD* expression decreases after lipid loading with monounsaturated fatty acids (200 µM oleic acid) 24-h of treatment only, following 8 hours of starvation in 0.5% FBS **F**, *OLMALINC* decreases its expression after lipid loading with monounsaturated fatty acids (200 µM oleic acid) after 18-h and 24-h of treatment, following 8 hours of starvation in 0.5% FBS. All expression time points are normalized to the corresponding gene expression in 0.5% FBS. Values are mean ± s.d. (n=3). *P<0.05; **P<0.01, and ***P<0.001 (unpaired Student's *t* test was used for two groups).
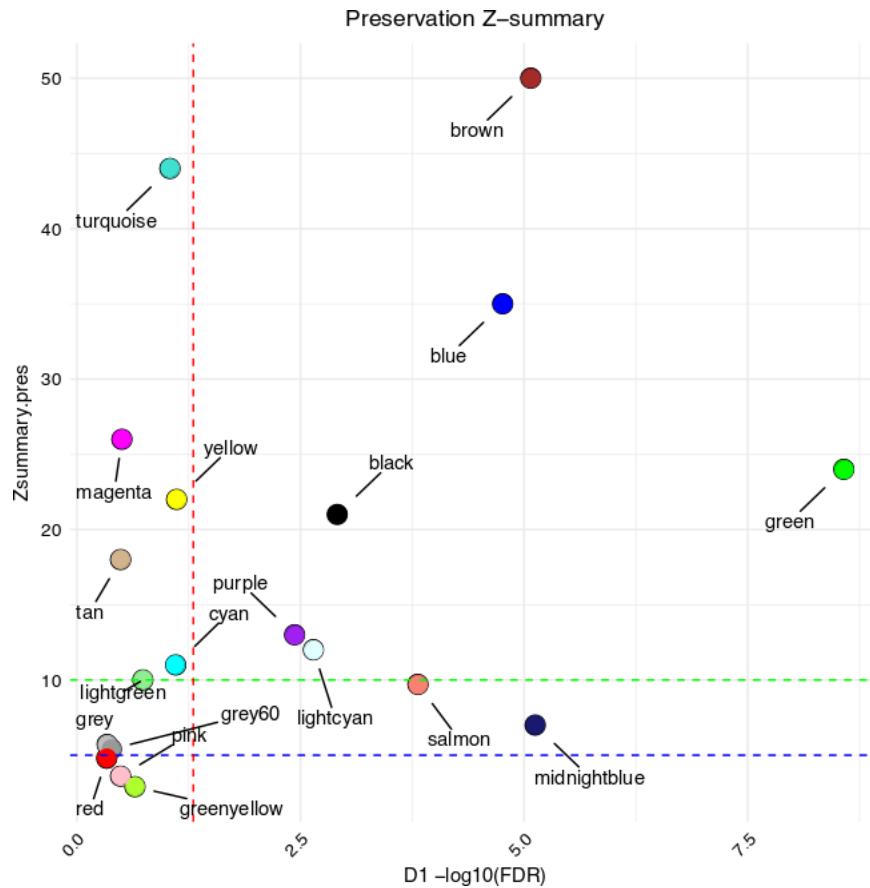
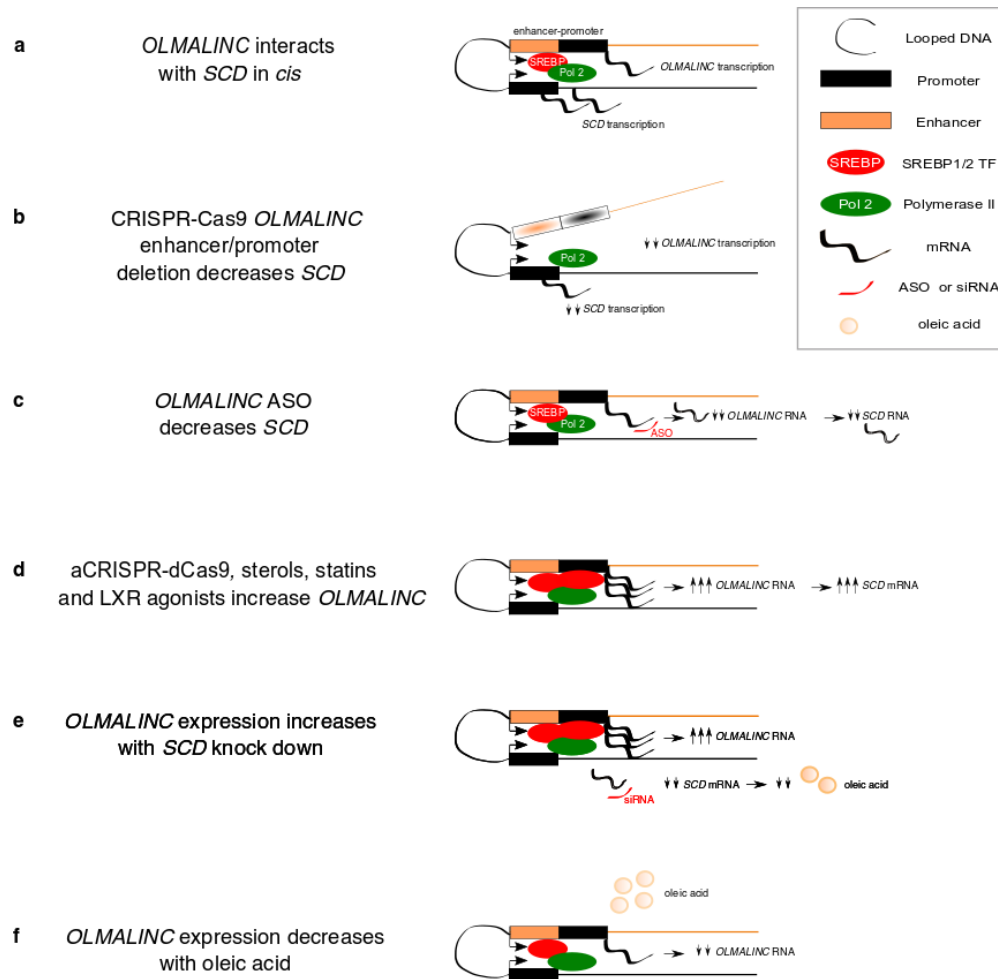**Figure 8-[Supplementary figure S1]- Most KOBS liver WGCNA modules are preserved in the GTEx liver RNA-seq data (n=96).** We considered a Z-score > 10 strongly preserved based on the previous guidelines [98]. The blue and green horizontal dotted lines indicate the summary preservation Z-score at 5 and 10, respectively, and the red dotted line indicates the significant threshold for a trait association at FDR<0.05.
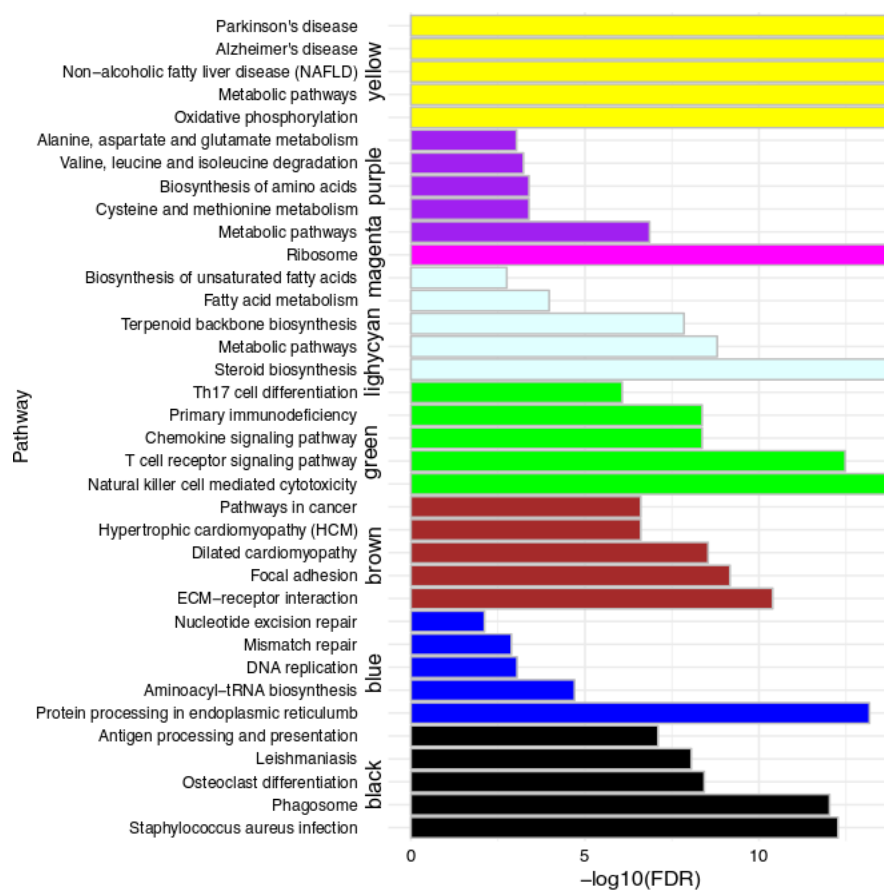
Figure 9-[Supplementary figure S2]- The statin-associated liver WGCNA network module is enriched for steroid biosynthesis, fatty acid and other metabolic pathways. We show the top 5 significant preserved pathways (FDR<0.05) of each module. The cyan module is excluded due to no enrichments.

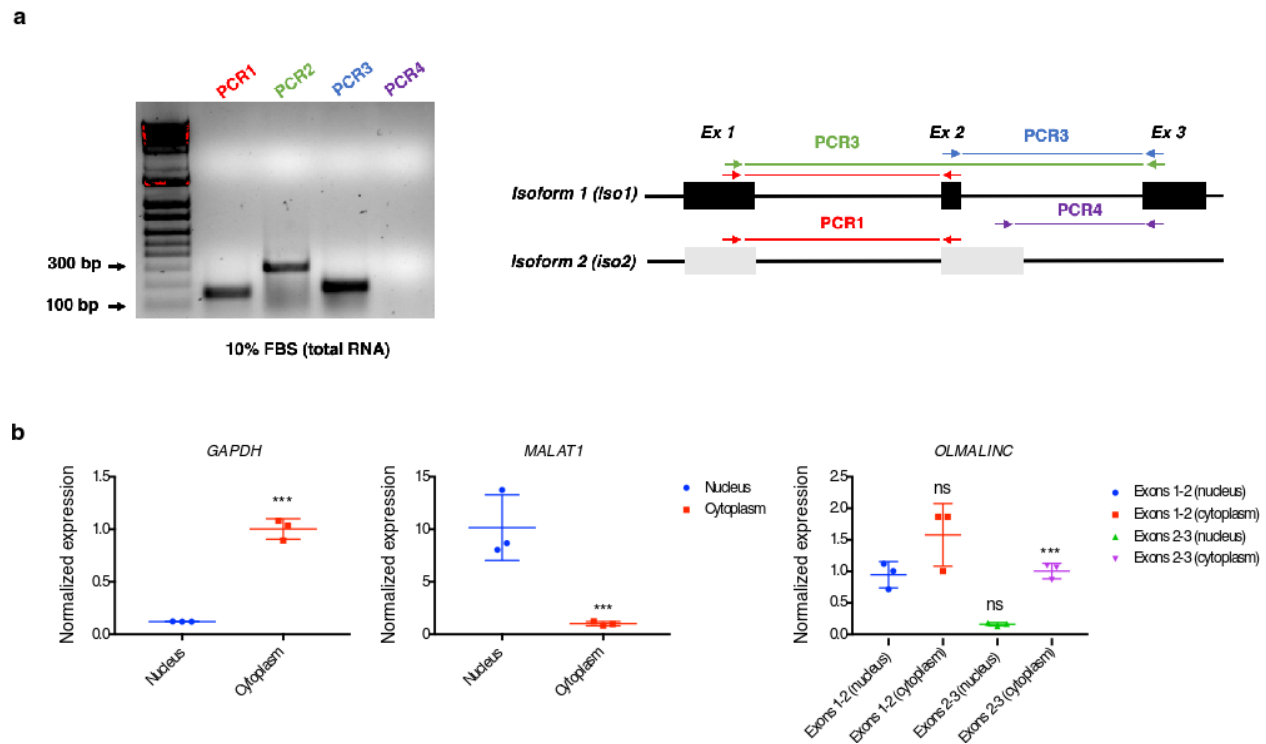**Figure 10-[Supplemental figure S3]-** *OLMALINC* gene expression across different human tissues

in the GTEx cohort (expression shown in TPMs).

Figure 11-[Supplemental figure S4]- **Overview and schematic representation showing how**
***OLMALINC* regulates *SCD* in *cis*. A**, The *OLMALINC* enhancer/promoter regions interact with
SCD promoter/enhancer regions at the DNA level by forming a DNA-DNA looping interaction.
B, Deletion of the *OLMALINC* promoter/enhancer regions by CRISPR-Cas9 gene editing
decreases *OLMALINC* and *SCD* gene expression. **C**, ASO targeting *OLMALINC* in HepG2 cells
causes a decrease in *OLMALINC* and *SCD* gene expressions. **D**, *OLMALINC*, similarly to *SCD*, is
responsive to sterols, statins and LXR agonists. Endogenous over-expression of *OLMALINC* by
aCRISPR-dCas9 increases *SCD* gene expression. **E**, *OLMALINC* gene expression increases with
SCD siRNA knock-down. **F**, *OLMALINC* gene expression decreases after treating HepG2 cells
with the SCD enzyme by-product oleic acid.

76

```
                              LXRE/DR4
                          ←——————    ←————
Human FAS           5'-gcgaTGACCGgcagTAACCCcggc-3'
Mouse SREBP1c       5'-acagTGACCGccagTAACCCcagc-3'
OLMALINC            5'-gtctgTGACCGgcagTAACCCcggc-3'
```

Figure 12-[Supplemental figure S5]- We identified an LXR responsive element (LXRE-DR4) in

the annotated *OLMALINC* promoter similarly to other LXR responsive genes, including *FAS* and

SREBP1c. This region is within the RXRA ChIP-seq binding site identified in ENCODE.

Figure 13-[Supplemental figure S6]- *OLMALINC* **has three exons expressed in HepG2 cells and demonstrates differential cellular localization between the cytoplasm and nucleus based on exonic expression. A**, PCR (and Sanger sequencing of isolated PCR fragments) confirms that *OLMALINC* expresses exons 1-3 in HepG2 cells under standard conditions. **B**, RT-qPCR confirms expression of exons 1-3 of *OLMALINC* and demonstrates a higher cytoplasmic expression for exons 2-3 when normalized to cytoplasmic expression for exons 1-2. *MALAT1* was used as a nuclear positive control and *GAPDH* as a cytoplasmic positive control.

78

Synteny between human chromosome10 and mouse chromosome 19
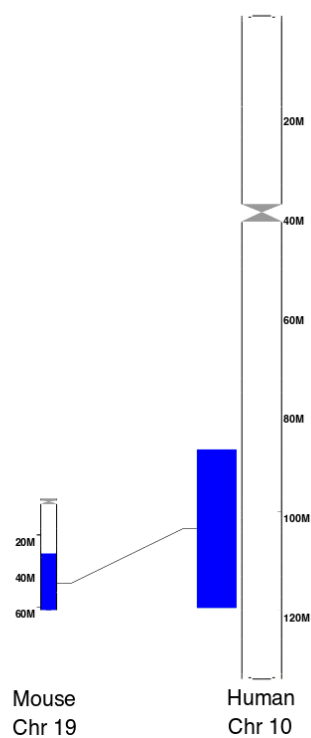
Figure 14-[Supplemental figure S7]- Synteny between the mouse chromosome 19 and human chromosome 10. The highlighted blue region demonstrates the syntenic regions between the two chromosomes.
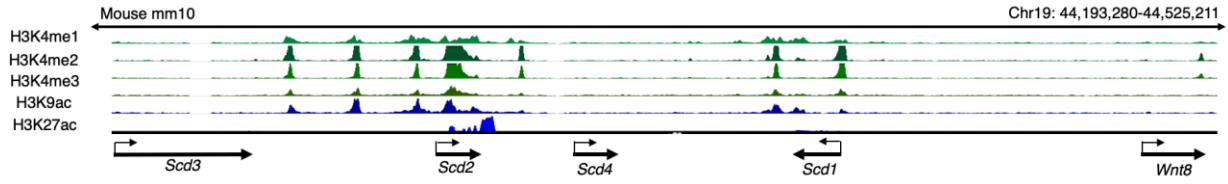
Figure 15-[Supplemental figure S8]- *SCD* **genes on mouse chromosome 19.** There is no evidence of histone methylation marks in the region flanking *SCD1* and *WTN8B* genes where *OLMALINC* is localized in the human chromosome 10, to suggest, an active transcription start site.
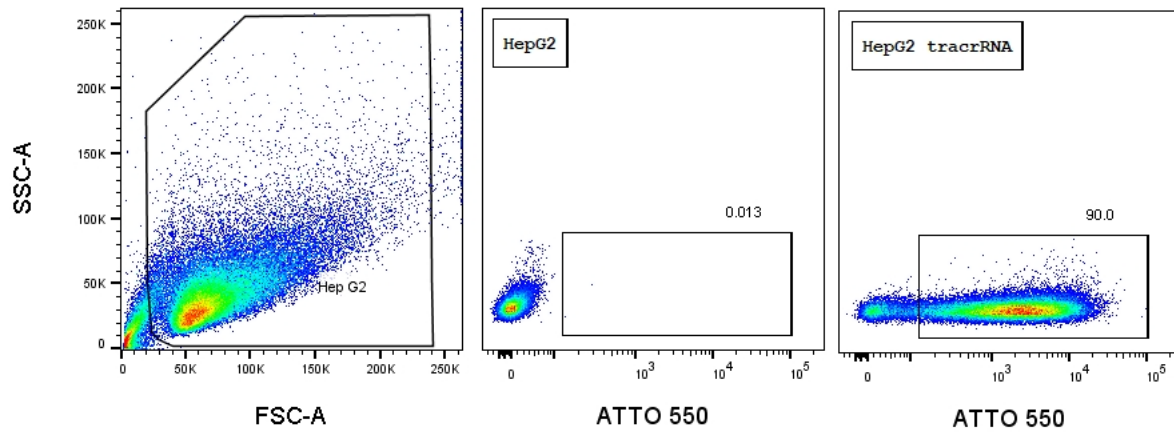
Figure 16-[Supplemental figure S9]- Evaluation of efficiency of HepG2 cells treated with CRISPR-Cas9 fluorescently labeled tracRNA ATTO-550 and *OLMALINC* promoter/enhancer gRNAs. Counting of labeled cells after 24-h transfection demonstrates an efficiency of 84% with Cas9 treated cells with corresponding gRNAs.

Supplemental table S1- Genes that correlate with *OLMALINC* liver expression in the KOBS liver RNA-seq cohort. *OLMALINC* liver expression correlated with 6182 genes in the KOBS liver RNA-seq cohort, passing FDR<0.05. The correlation analyses were conducted using a linear regression model (see Methods).

 (see attached excel file for Supplemental table 1)

Supplemental table S2- *OLMALINC* liver expression correlates with serum triglycerides in the KOBS cohort. Serum triglycerides correlate significantly (passing the Bonferroni corrected p-value cut-point of P<0.007 for the 7 tested traits) with *OLMALINC* liver expression. The correlation analyses were performed using linear and logistic regression models, and the quantitative traits were adjusted for age and sex as well as inverse normal transformed to avoid outlier effects.

Clinical phenotype	ß estimate	Standard error	p-value
NAFLD	0.210	0.126	0.097
NASH	0.069	0.167	0.677
Liver fibrosis	0.131	0.124	0.292
Liver steatosis	0.210	0.126	0.097
Type 2 Diabetes	0.237	0.127	0.064
Total cholesterol (mmol/L)	-0.06	0.069	0.366
Triglycerides (mmol/L)	0.275	0.084	***0.001***

Supplemental table 3S- siRNA and ASO sequences.

OLMALINC ASO	/52MOErA/*/i2MOErT/*/i2MOErG/*/i2MOErT/*/i2MOErC/ *A*C*A*T*G*C*A*T*C*C*/i2MOErG/*/i2MOErT/*/i2MOErG/*/i2MOErT/*/32 MOErG/
Control ASO	/52MOErG/*/i2MOErC/*/i2MOErG/*/i2MOErA/*/i2MOErC/*T*A*T*A*C*G*C* G*C*A*/i2MOErA/*/i2MOErT/*/i2MOErA/*/i2MOErT/*/32MOErG/
SCD siRNA	CCAGAGGAGGTACTAGAAATT
SREBP1 siRNA	CAGCTTATCAACAACCAAGACAGTG
SREBP2 siRNA	GCCTTTGATATACCAGAATTT

Supplemental table 4S- Primer sequences used for RT-qPCR.

Primer	Forward sequence (5'→3')	Reverse sequence (5'→3')
36B4 (RPLP0)	CCACGCTGCTGAACATGCT	TCGAACACCTGCTGGATGAC
GAPDH	GGGTGTGAACCATGAGAAGT	CCTTCCACGATACCAAAGTT
OLMALINC exons 1-2	CCAGGAGTCAGCAAAACACA	CTGGGTCTTCAGCACCAAAT
OLMALINC exons 2-3	CATGTGACATTTGGTGCTGA	CTTGGACTCAGAGGCCTGAC
SCD	TGCCCACCTCTTCGGATATC	GATGTGCCAGCGGTACTCACT
SREBP2	GACGCCAAGATGCACAAGTC	ACCAGACTGCCTAGGTCGAT
SREBP1a	TCAGCGAGGCGGCTTTGGAGCAG	CATGTCTTCGATGTCGGTCAG
SREBP1c	CGCTCCTCCATCAATGACA	TGCGCAAGACAGCAGATTTA
HMGCS1	GATGTGGGAATTGTTGCCCTT	ATTGTCTCTGTTCCAACTTCCAG
LDLR	AGGCTGTGGGCTCCATCGCCTA	AGTCAGTCCAGTACATGAAGCCA
MALAT1	GGTAACGATGGTGTCGAGGTC	CCAGCATTACAGTTCTTGAACATG

Supplemental table 5S- *OLMALINC* **sgRNA sequences for CRISPR-Cas9 knock out (the region column specifies the region upstream or downstream of the** *OLMALINC* **promoter/enhancer regions).** Combination of the two different upstream and downstream guide RNAs were tested to determine efficiency of knock-out. The combination of (*) showed the highest efficiency and used in the final experiments.

Region	Strand	Sequence	PAM
Upstream	(-)	*GATTGTATCCACAAGTCTGA	TGG
	(-)	AATAACACCTGCTTCGGATG	TGG
Downstream	(+)	*AAAGCTGGGATAGTCACGGT	GGG
	(+)	GGAAGCTATTGTTACCACTC	TGG

Supplemental table 6S- *OLMALINC* **primers used to validate the enhancer/promoter deletions.**

Region	Forward sequence (5'→3')	Reverse sequence (5'→3')
Outside enhancer/promoter	GGCAAGCTGCTATAAACTGGA	CCTCCCAAATGCCTCTCAGC
Within enhancer/promoter	AGCCTTCCCACTTTCAGGAC	AGCCTTCCCACTTTCAGGAC

Chapter 4

Concluding remarks and future directions

Our data suggest that having a concomitant diagnosis of NAFLD with chronic HCV does not affect HCV cure in a large, diverse VA patient cohort. We observe that HbA1c improves independently of the HCV genotype or changes in a patient's T2D medication regimen, suggesting a virus-dependent effect of improved glycemic control. Given the era of DAAs and high HCV cure rates, it is also important to understand the prevalence of NAFLD in the post SVR12 patient population to elucidate how concomitant fatty liver affects liver fibrosis progression and HCC risk. Some studies have estimated that 50% of patients who reached SVR12 have underlying NAFLD [120]. Our preliminary data using the national VHA Corporate Data Warehouse demonstrate that NAFLD and/or NASH are underdiagnosed (18%) in the post SVR12 patient population using the ICD9 and ICD10 codes. Identifying this patient population and their clinical outcomes for mortality, liver-related complications and HCC will be important since the natural history of NAFLD in this patient population is not understood and will affect how these patients are treated in clinic[55].

We also show that high BMI, and potentially T2D, predict NAFLD cirrhosis in a NAFLD-associated HCC cohort. There are ethnic differences not only in how patients respond to HCV treatment but also in the NAFLD and NAFLD-associated HCC, where non-Hispanic patients are more likely to develop HCC in a non-cirrhosis background. Leveraging these differences can help us create population and patient-specific screening tools to better address the heterogeneity of these complex diseases. Current general population guidelines do not reflect these nuances and are thus not optimal in the era of precision health. Through our UCLA study, we also demonstrate that etiologies of HCC affect overall and recurrence free survival. Although dyslipidemia appeared to have a trend towards protection of NAFLD cirrhosis, the differences were not significant, likely due to small sample sizes and thus, they warrant further investigation in larger populations.

Through our detailed transcriptomics analyses of the human liver RNA-sequencing, we identified novel genes that are statin responsive that should be further evaluated, especially in the context of the potential benefits of statins on the non-cirrhosis patient population. Targeting of these novel lipid genes, such as *OLMALINC*, may have synergistic effects on the cholesterol and lipid pathways. Although not assessed further in our work, we observed that *OLMALINC* gene expression was significantly associated with *SREBP2* gene expression (ß=0.67, FDR=3.96×10^{-34}) (and SREBP2 target genes), the main transcription factor for the cholesterol synthesis pathway. How *OLMALINC* regulates *SREBP2* is of interest and needs additional investigation given the non-cholesterol lowering beneficial effects of statin therapy, which *OLMALINC* may have. For instance, we observed that *OLMALINC* knock down and knock out (through ASO and CRISPR-Cas9, respectively) decreased cell viability at ~2-3 weeks in culture. This was previously observed by Lui et al. independently in a brain U87 glioblastoma multiforme cell line [107]. Given that our work was conducted in the cancerous HepG2 cell line and that SCD has been shown to play a role in liver fibrosis and tumorigenesis through the activation of the wnt pathway [121-123], these effects will need to be studied in a non-cancer or primary hepatocyte cell lines. We currently have on-going studies repeating similar experiments in the non-cancer immortalized human hepatocyte cell line, Fa2N4. Differences observed in *SCD* regulation by *OLMALINC* between the HepG2 and Fa2N4 cell-lines could be attributed to the cancer and non-cancer origins of these cell-lines.

To further investigate and address the molecular basis of the non-cirrhosis HCC, we will be sequencing human liver biopsy samples from HCC and adjacent non-cancer NASH liver tissue. Due to the heterogeneity that exists in HCC [124,125], our experiments will be conducted using single cell RNA-sequencing, which has successfully been utilized in other cancers and has allowed the identification of new cell-type sub-populations not previously known from bulk RNA-sequencing.

Single cell nuclei extraction and sequencing have also been successfully conducted in our laboratory using frozen human subcutaneous adipose tissue (Alvarez et al., manuscript in preparation). Understanding which genes are differentially expressed between the HCC and non-HCC liver in NASH patients will provide insight into which pathways need closer evaluation, including those involved in the mevalonate pathway [126].

The NAFLD, NASH and HCC epidemics are alarming and have already placed large economic burdens on healthcare [49,127]. Understanding their clinical and epidemiological factors and characteristics in different health care settings and populations will help us frame the biologically-relevant questions and subsequently identify the involved mechanisms. Thus, combining multidisciplinary approaches to study these patients will allow for new discoveries and advance the field of NAFLD, NASH and HCC.

References:

1. Denniston MM, Jiles RB, Drobeniuc J, et al. Chronic hepatitis C virus infection in the United States, National Health and Nutrition Examination Survey 2003 to 2010. *Ann Intern Med.* 2014;160(5):293-300.

2. Niederau C, Lange S, Heintges T, et al. Prognosis of chronic hepatitis C: results of a large, prospective cohort study. *Hepatology.* 1998;28(6):1687-1695.

3. Butt AA, Wang X, Moore CG. Effect of hepatitis C virus and its treatment on survival. *Hepatology.* 2009;50(2):387-392.

4. Morgan RL, Baack B, Smith BD, Yartel A, Pitasi M, Falck-Ytter Y. Eradication of hepatitis C virus infection and the development of hepatocellular carcinoma: a meta-analysis of observational studies. *Ann Intern Med.* 2013;158(5 Pt 1):329-337.

5. Backus LI, Belperio PS, Shahoumian TA, Loomis TP, Mole LA. Real-world effectiveness of ledipasvir/sofosbuvir in 4,365 treatment-naive, genotype 1 hepatitis C-infected patients. *Hepatology.* 2016;64(2):405-414.

6. Jadoon NA, Shahzad MA, Yaqoob R, Hussain M, Ali N. Seroprevalence of hepatitis C in type 2 diabetes: evidence for a positive association. *Virol J.* 2010;7:304.

7. Mason AL, Lau JY, Hoang N, et al. Association of diabetes mellitus and chronic hepatitis C virus infection. *Hepatology.* 1999;29(2):328-333.

8. Allison ME, Wreghitt T, Palmer CR, Alexander GJ. Evidence for a link between hepatitis C virus infection and diabetes mellitus in a cirrhotic population. *J Hepatol.* 1994;21(6):1135-1139.

9. Hui JM, Sud A, Farrell GC, et al. Insulin resistance is associated with chronic hepatitis C virus infection and fibrosis progression [corrected]. *Gastroenterology.* 2003;125(6):1695-1704.

10. Hickman IJ, Powell EE, Prins JB, et al. In overweight patients with chronic hepatitis C, circulating insulin is associated with hepatic fibrosis: implications for therapy. *J Hepatol.* 2003;39(6):1042-1048.

11. El-Serag HB, Tran T, Everhart JE. Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma. *Gastroenterology.* 2004;126(2):460-468.

12. Adami HO, Chow WH, Nyren O, et al. Excess risk of primary liver cancer in patients with diabetes mellitus. *J Natl Cancer Inst.* 1996;88(20):1472-1477.

13. Furutani M, Nakashima T, Sumida Y, et al. Insulin resistance/beta-cell function and serum ferritin level in non-diabetic patients with hepatitis C virus infection. *Liver Int.* 2003;23(4):294-299.

14.   Knobler H, Zhornicky T, Sandler A, Haran N, Ashur Y, Schattner A. Tumor necrosis factor-alpha-induced insulin resistance may mediate the hepatitis C virus-diabetes association. *Am J Gastroenterol.* 2003;98(12):2751-2756.

15.   Estes C, Razavi H, Loomba R, Younossi Z, Sanyal AJ. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology.* 2018;67(1):123-133.

16.   Dominitz JA, Boyko EJ, Koepsell TD, et al. Elevated prevalence of hepatitis C infection in users of United States veterans medical centers. *Hepatology.* 2005;41(1):88-96.

17.   Nelson KM. The burden of obesity among a national probability sample of veterans. *J Gen Intern Med.* 2006;21(9):915-919.

18.   Miller DR, Safford MM, Pogach LM. Who has diabetes? Best estimates of diabetes prevalence in the Department of Veterans Affairs based on computerized patient data. *Diabetes Care.* 2004;27 Suppl 2:B10-21.

19.   Yoshida EM, Sulkowski MS, Gane EJ, et al. Concordance of sustained virological response 4, 12, and 24 weeks post-treatment with sofosbuvir-containing regimens for hepatitis C virus. *Hepatology.* 2015;61(1):41-45.

20.   Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology.* 2007;45(4):846-854.

21.   Vallet-Pichard A, Mallet V, Nalpas B, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology.* 2007;46(1):32-36.

22.   Lenters-Westra E, Schindhelm RK, Bilo HJ, Groenier KH, Slingerland RJ. Differences in interpretation of haemoglobin A1c values among diabetes care professionals. *Neth J Med.* 2014;72(9):462-466.

23.   Little RR, Rohlfing CL, Sacks DB, National Glycohemoglobin Standardization Program Steering C. Status of hemoglobin A1c measurement and goals for improvement: from chaos to order for improving diabetes care. *Clin Chem.* 2011;57(2):205-214.

24.   Rutter MK, Wilson PW, Sullivan LM, Fox CS, D'Agostino RB, Sr., Meigs JB. Use of alternative thresholds defining insulin resistance to predict incident type 2 diabetes mellitus and cardiovascular disease. *Circulation.* 2008;117(8):1003-1009.

25.   Karve S, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. Prospective validation of eight different adherence measures for use with administrative claims data among patients with schizophrenia. *Value Health.* 2009;12(6):989-995.

26.   Heinze G, Schemper M. A solution to the problem of separation in logistic regression. *Stat Med.* 2002;21(16):2409-2419.

27.   Maiti T, Pradhan V. Bias reduction and a solution for separation of logistic regression with missing covariates. *Biometrics.* 2009;65(4):1262-1269.

28.   Wilder JM, Jeffers LJ, Ravendhran N, et al. Safety and efficacy of ledipasvir-sofosbuvir in black patients with hepatitis C virus infection: A retrospective analysis of phase 3 data. *Hepatology.* 2016;63(2):437-444.

29.   Konerman MA, Lok AS. Hepatitis C Treatment and Barriers to Eradication. *Clin Transl Gastroenterol.* 2016;7(9):e193.

30.   Backus LI, Boothroyd DB, Phillips BR, Mole LA. Predictors of response of US veterans to treatment for the hepatitis C virus. *Hepatology.* 2007;46(1):37-47.

31.   Su F, Green PK, Berry K, Ioannou GN. The association between race/ethnicity and the effectiveness of direct antiviral agents for hepatitis C virus infection. *Hepatology.* 2017;65(2):426-438.

32.   Louie V, Latt NL, Gharibian D, et al. Real-World Experiences With a Direct-Acting Antiviral Agent for Patients With Hepatitis C Virus Infection. *Perm J.* 2017;21.

33.   Yasuda SU, Zhang L, Huang SM. The role of ethnicity in variability in response to drugs: focus on clinical pharmacology studies. *Clin Pharmacol Ther.* 2008;84(3):417-423.

34.   Burroughs VJ, Maxey RW, Levy RA. Racial and ethnic differences in response to medicines: towards individualized pharmaceutical treatment. *J Natl Med Assoc.* 2002;94(10 Suppl):1-26.

35.   Ge D, Fellay J, Thompson AJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature.* 2009;461(7262):399-401.

36.   Rosen HR, Miner C, Sasaki AW, et al. Frequencies of HCV-specific effector CD4+ T cells by flow cytometry: correlation with clinical disease stages. *Hepatology.* 2002;35(1):190-198.

37.   Wojcik G, Latanich R, Mosbruger T, et al. Variants in HAVCR1 gene region contribute to hepatitis C persistence in African Americans. *J Infect Dis.* 2014;209(3):355-359.

38.   Akamatsu S, Hayes CN, Ochi H, et al. Association between variants in the interferon lambda 4 locus and substitutions in the hepatitis C virus non-structural protein 5A. *J Hepatol.* 2015;63(3):554-563.

39.   Hum J, Jou JH, Green PK, et al. Improvement in Glycemic Control of Type 2 Diabetes After Successful Treatment of Hepatitis C Virus. *Diabetes Care.* 2017;40(9):1173-1180.

40.   Adinolfi LE, Nevola R, Rinaldi L, Romano C, Giordano M. Chronic Hepatitis C Virus Infection and Depression. *Clin Liver Dis.* 2017;21(3):517-534.

41. Desbois AC, Cacoub P. Diabetes mellitus, insulin resistance and hepatitis C virus infection: A contemporary review. *World J Gastroenterol.* 2017;23(9):1697-1711.

42. Conjeevaram HS, Wahed AS, Afdhal N, et al. Changes in insulin sensitivity and body weight during and after peginterferon and ribavirin therapy for hepatitis C. *Gastroenterology.* 2011;140(2):469-477.

43. Corey KE, Kartoun U, Zheng H, Shaw SY. Development and Validation of an Algorithm to Identify Nonalcoholic Fatty Liver Disease in the Electronic Medical Record. *Dig Dis Sci.* 2016;61(3):913-919.

44. Husain N, Blais P, Kramer J, et al. Nonalcoholic fatty liver disease (NAFLD) in the Veterans Administration population: development and validation of an algorithm for NAFLD using automated data. *Aliment Pharmacol Ther.* 2014;40(8):949-954.

45. Younossi Z, Stepanova M, Ong JP, et al. Nonalcoholic Steatohepatitis Is the Fastest Growing Cause of Hepatocellular Carcinoma in Liver Transplant Candidates. *Clin Gastroenterol Hepatol.* 2019;17(4):748-755 e743.

46. Kanwal F, Kramer JR, Mapakshi S, et al. Risk of Hepatocellular Cancer in Patients With Non-Alcoholic Fatty Liver Disease. *Gastroenterology.* 2018;155(6):1828-1837 e1822.

47. Adams LA, Lindor KD. Nonalcoholic fatty liver disease. *Ann Epidemiol.* 2007;17(11):863-869.

48. Brunt EM, Tiniakos DG. Histopathology of nonalcoholic fatty liver disease. *World J Gastroenterol.* 2010;16(42):5286-5296.

49. Kaplan DE, Chapko MK, Mehta R, et al. Healthcare Costs Related to Treatment of Hepatocellular Carcinoma Among Veterans With Cirrhosis in the United States. *Clin Gastroenterol Hepatol.* 2018;16(1):106-114 e105.

50. Piscaglia F, Svegliati-Baroni G, Barchetti A, et al. Clinical patterns of hepatocellular carcinoma in nonalcoholic fatty liver disease: A multicenter prospective study. *Hepatology.* 2016;63(3):827-838.

51. Stine JG, Wentworth BJ, Zimmet A, et al. Systematic review with meta-analysis: risk of hepatocellular carcinoma in non-alcoholic steatohepatitis without cirrhosis compared to other liver diseases. *Aliment Pharmacol Ther.* 2018;48(7):696-703.

52. Simon TG, King LY, Chong DQ, et al. Diabetes, metabolic comorbidities, and risk of hepatocellular carcinoma: Results from two prospective cohort studies. *Hepatology.* 2018;67(5):1797-1806.

53. Tong MJ, Rosinski AA, Huynh CT, Raman SS, Lu DSK. Long-term survival after surveillance and treatment in patients with chronic viral hepatitis and hepatocellular carcinoma. *Hepatol Commun.* 2017;1(7):595-608.

54.   Singal AG, Mittal S, Yerokun OA, et al. Hepatocellular Carcinoma Screening Associated with Early Tumor Detection and Improved Survival Among Patients with Cirrhosis in the US. *Am J Med.* 2017;130(9):1099-1106 e1091.

55.   Serper M, Kaplan DE, Shults J, et al. Quality measures, all-cause mortality, and healthcare utilization in a national cohort of Veterans with cirrhosis. *Hepatology.* 2019.

56.   Kanwal F, Singal AG. Surveillance for HCC: Current Best Practice and Future Direction. *Gastroenterology.* 2019.

57.   Kim TK, Lee KH, Jang HJ, et al. Analysis of gadobenate dimeglumine-enhanced MR findings for characterizing small (1-2-cm) hepatic nodules in patients at high risk for hepatocellular carcinoma. *Radiology.* 2011;259(3):730-738.

58.   Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology.* 1996;24(2):289-293.

59.   Hester CA, Rich NE, Singal AG, Yopp AC. Comparative Analysis of Nonalcoholic Steatohepatitis- Versus Viral Hepatitis- and Alcohol-Related Liver Disease-Related Hepatocellular Carcinoma. *J Natl Compr Canc Netw.* 2019;17(4):322-329.

60.   Wakai T, Shirai Y, Sakata J, Korita PV, Ajioka Y, Hatakeyama K. Surgical outcomes for hepatocellular carcinoma in nonalcoholic fatty liver disease. *J Gastrointest Surg.* 2011;15(8):1450-1458.

61.   Hernandez-Alejandro R, Croome KP, Drage M, et al. A comparison of survival and pathologic features of non-alcoholic steatohepatitis and hepatitis C virus patients with hepatocellular carcinoma. *World J Gastroenterol.* 2012;18(31):4145-4149.

62.   Welzel TM, Graubard BI, Zeuzem S, El-Serag HB, Davila JA, McGlynn KA. Metabolic syndrome increases the risk of primary liver cancer in the United States: a study in the SEER-Medicare database. *Hepatology.* 2011;54(2):463-471.

63.   Phan J, Ng V, Sheinbaum A, et al. Hyperlipidemia and Nonalcoholic Steatohepatitis Predispose to Hepatocellular Carcinoma Development Without Cirrhosis. *J Clin Gastroenterol.* 2019;53(4):309-313.

64.   Thrift AP, Natarajan Y, Liu Y, El-Serag HB. Statin Use After Diagnosis of Hepatocellular Carcinoma Is Associated With Decreased Mortality. *Clin Gastroenterol Hepatol.* 2019.

65.   Butt AA, Yan P, Bonilla H, et al. Effect of addition of statins to antiviral therapy in hepatitis C virus-infected persons: Results from ERCHIVES. *Hepatology.* 2015;62(2):365-374.

66.   Lee TY, Wu JC, Yu SH, Lin JT, Wu MS, Wu CY. The occurrence of hepatocellular carcinoma in different risk stratifications of clinically noncirrhotic nonalcoholic fatty liver disease. *Int J Cancer.* 2017;141(7):1307-1314.

67. Fan JG, Kim SU, Wong VW. New trends on obesity and NAFLD in Asia. *J Hepatol.* 2017;67(4):862-873.

68. Kim D, Kim WR. Nonobese Fatty Liver Disease. *Clin Gastroenterol Hepatol.* 2017;15(4):474-485.

69. Kubota K, Ina H, Okada Y, Irie T. Growth rate of primary single hepatocellular carcinoma: determining optimal screening interval with contrast enhanced computed tomography. *Dig Dis Sci.* 2003;48(3):581-586.

70. Cheng SJ, Freeman RB, Jr., Wong JB. Predicting the probability of progression-free survival in patients with small hepatocellular carcinoma. *Liver Transpl.* 2002;8(4):323-328.

71. Northup PG, Intagliata NM, Shah NL, Pelletier SJ, Berg CL, Argo CK. Excess mortality on the liver transplant waiting list: unintended policy consequences and Model for End-Stage Liver Disease (MELD) inflation. *Hepatology.* 2015;61(1):285-291.

72. Grohmann M, Wiede F, Dodd GT, et al. Obesity Drives STAT-1-Dependent NASH and STAT-3-Dependent HCC. *Cell.* 2018;175(5):1289-1306 e1220.

73. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther.* 2011;34(3):274-285.

74. 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(1):73-84.

75. Browning JD, Horton JD. Molecular mediators of hepatic steatosis and liver injury. *J Clin Invest.* 2004;114(2):147-152.

76. Min HK, Kapoor A, Fuchs M, et al. Increased hepatic synthesis and dysregulation of cholesterol metabolism is associated with the severity of nonalcoholic fatty liver disease. *Cell Metab.* 2012;15(5):665-674.

77. Chiappini F, Coilly A, Kadar H, et al. Metabolism dysregulation induces a specific lipid signature of nonalcoholic steatohepatitis in patients. *Sci Rep.* 2017;7:46658.

78. Mannisto VT, Simonen M, Soininen P, et al. Lipoprotein subclass metabolism in nonalcoholic steatohepatitis. *J Lipid Res.* 2014;55(12):2676-2684.

79. Ntambi JM. Regulation of stearoyl-CoA desaturase by polyunsaturated fatty acids and cholesterol. *J Lipid Res.* 1999;40(9):1549-1558.

80. Walle P, Takkunen M, Mannisto V, et al. Fatty acid metabolism is altered in non-alcoholic steatohepatitis independent of obesity. *Metabolism.* 2016;65(5):655-666.

81. Puri P, Wiest MM, Cheung O, et al. The plasma lipidomic signature of nonalcoholic steatohepatitis. *Hepatology.* 2009;50(6):1827-1838.

82. Ntambi JM, Miyazaki M, Stoehr JP, et al. Loss of stearoyl-CoA desaturase-1 function protects mice against adiposity. *Proc Natl Acad Sci U S A.* 2002;99(17):11482-11486.

83. Iruarrizaga-Lejarreta M, Varela-Rey M, Fernandez-Ramos D, et al. Role of Aramchol in steatohepatitis and fibrosis in mice. *Hepatol Commun.* 2017;1(9):911-927.

84. Safadi R, Konikoff FM, Mahamid M, et al. The fatty acid-bile acid conjugate Aramchol reduces liver fat content in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol.* 2014;12(12):2085-2091 e2081.

85. van Solingen C, Scacalossi KR, Moore KJ. Long noncoding RNAs in lipid metabolism. *Curr Opin Lipidol.* 2018;29(3):224-232.

86. Kopp F, Mendell JT. Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell.* 2018;172(3):393-407.

87. Kaikkonen MU, Adelman K. Emerging Roles of Non-Coding RNA Transcription. *Trends Biochem Sci.* 2018;43(9):654-667.

88. Sallam T, Jones MC, Gilliland T, et al. Feedback modulation of cholesterol metabolism by the lipid-responsive non-coding RNA LeXis. *Nature.* 2016;534(7605):124-128.

89. Mannisto VT, Simonen M, Hyysalo J, et al. Ketone body production is differentially altered in steatosis and non-alcoholic steatohepatitis in obese humans. *Liver Int.* 2015;35(7):1853-1861.

90. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* 2013;45(6):580-585.

91. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology.* 2005;41(6):1313-1321.

92. Brunt EM. Histopathology of non-alcoholic fatty liver disease. *Clin Liver Dis.* 2009;13(4):533-544.

93. Simonen M, Mannisto V, Leppanen J, et al. Desmosterol in human nonalcoholic steatohepatitis. *Hepatology.* 2013;58(3):976-982.

94. Leeuw Jd. Gifi Methods for Optimal Scaling in R: The Package homals. *Journal of Statistical Software.* 2009;31(4):1-21.

95. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol.* 2016;34(5):525-527.

96.	Leek JT. svaseq: removing batch effects and other unwanted noise from sequencing data. *Nucleic Acids Res.* 2014;42(21).

97.	Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.

98.	Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? *PLoS Comput Biol.* 2011;7(1):e1001057.

99.	Tarling EJ, Edwards PA. ATP binding cassette transporter G1 (ABCG1) is an intracellular sterol transporter. *Proc Natl Acad Sci U S A.* 2011;108(49):19719-19724.

100.	Pan DZ, Garske KM, Alvarez M, et al. Integration of human adipocyte chromosomal interactions with adipose gene expression prioritizes obesity-related genes from GWAS. *Nat Commun.* 2018;9(1):1512.

101.	Mifsud B, Tavares-Cadete F, Young AN, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet.* 2015;47(6):598-606.

102.	Wingett S, Ewels P, Furlan-Magaril M, et al. HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res.* 2015;4:1310.

103.	Cairns J, Freire-Pritchett P, Wingett SW, et al. CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data. *Genome Biol.* 2016;17(1):127.

104.	Wang D, Garcia-Bassets I, Benner C, et al. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature.* 2011;474(7351):390-394.

105.	Kaikkonen MU, Niskanen H, Romanoski CE, et al. Control of VEGF-A transcriptional programs by pausing and genomic compartmentalization. *Nucleic Acids Res.* 2014;42(20):12570-12584.

106.	Rankin CR, Treger J, Faure-Kumar E, et al. Overexpressing Long Noncoding RNAs Using Gene-activating CRISPR. *J Vis Exp.* 2019(145).

107.	Liu SJ, Horlbeck MA, Cho SW, et al. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science.* 2017;355(6320).

108.	Engreitz JM, Haines JE, Perez EM, et al. Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature.* 2016;539(7629):452-455.

109.	Mills JD, Kavanagh T, Kim WS, et al. High expression of long intervening non-coding RNA OLMALINC in the human cortical white matter is associated with regulation of oligodendrocyte maturation. *Mol Brain.* 2015;8:2.

110.	Konermann S, Brigham MD, Trevino AE, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature.* 2015;517(7536):583-588.

111.   Shimano H, Shimomura I, Hammer RE, et al. Elevated levels of SREBP-2 and cholesterol synthesis in livers of mice homozygous for a targeted disruption of the SREBP-1 gene. *J Clin Invest.* 1997;100(8):2115-2124.

112.   Brown MS, Goldstein JL. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell.* 1997;89(3):331-340.

113.   Horton JD, Goldstein JL, Brown MS. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J Clin Invest.* 2002;109(9):1125-1131.

114.   Yan C, Chen J, Chen N. Long noncoding RNA MALAT1 promotes hepatic steatosis and insulin resistance by increasing nuclear SREBP-1c protein stability. *Sci Rep.* 2016;6:22640.

115.   Liu C, Yang Z, Wu J, et al. Long noncoding RNA H19 interacts with polypyrimidine tract-binding protein 1 to reprogram hepatic lipid homeostasis. *Hepatology.* 2018;67(5):1768-1783.

116.   Ebisuya M, Yamamoto T, Nakajima M, Nishida E. Ripples from neighbouring transcription. *Nat Cell Biol.* 2008;10(9):1106-1113.

117.   Kim TK, Shiekhattar R. Architectural and Functional Commonalities between Enhancers and Promoters. *Cell.* 2015;162(5):948-959.

118.   Ajmera VH, Cachay E, Ramers C, et al. Novel MRI assessment of treatment response in HIV-associated NAFLD: a randomized trial of an SCD1 inhibitor (ARRIVE Trial). *Hepatology.* 2019.

119.   Liu X, Burhans MS, Flowers MT, Ntambi JM. Hepatic oleate regulates liver stress response partially through PGC-1alpha during high-carbohydrate feeding. *J Hepatol.* 2016;65(1):103-112.

120.   Noureddin M, Wong MM, Todo T, Lu SC, Sanyal AJ, Mena EA. Fatty liver in hepatitis C patients post-sustained virological response with direct-acting antivirals. *World J Gastroenterol.* 2018;24(11):1269-1277.

121.   Bansal S, Berk M, Alkhouri N, Partrick DA, Fung JJ, Feldstein A. Stearoyl-CoA desaturase plays an important role in proliferation and chemoresistance in human hepatocellular carcinoma. *J Surg Res.* 2014;186(1):29-38.

122.   Yamashita T, Honda M, Takatori H, et al. Activation of lipogenic pathway correlates with cell proliferation and poor prognosis in hepatocellular carcinoma. *J Hepatol.* 2009;50(1):100-110.

123.   Lai KKY, Kweon SM, Chi F, et al. Stearoyl-CoA Desaturase Promotes Liver Fibrosis and Tumor Development in Mice via a Wnt Positive-Signaling Loop by Stabilization of Low-Density Lipoprotein-Receptor-Related Proteins 5 and 6. *Gastroenterology.* 2017;152(6):1477-1491.

124. Jovel J, Lin Z, O'Keefe S, et al. A Survey of Molecular Heterogeneity in Hepatocellular Carcinoma. *Hepatol Commun.* 2018;2(8):941-955.

125. Hou Y, Guo H, Cao C, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* 2016;26(3):304-319.

126. Jiang P, Mukthavaram R, Chao Y, et al. In vitro and in vivo anticancer effects of mevalonate pathway modulation on human cancer cells. *Br J Cancer.* 2014;111(8):1562-1571.

127. Sayiner M, Otgonsuren M, Cable R, et al. Variables Associated With Inpatient and Outpatient Resource Utilization Among Medicare Beneficiaries With Nonalcoholic Fatty Liver Disease With or Without Cirrhosis. *J Clin Gastroenterol.* 2017;51(3):254-260.