SUPPLEMENTAL MATERIAL

The protection conferred by *HSD17B13* rs72613567 polymorphism on risk of steatohepatitis and fibrosis may be limited to selected subgroups of patients with NAFLD.

[1]EduardoVilar-Gomez, [2]Carlos J. Pirola, [3]SilviaSookoian, [4]Laura A. Wilson, [1]TiebingLiang, [1]NagaChalasani.

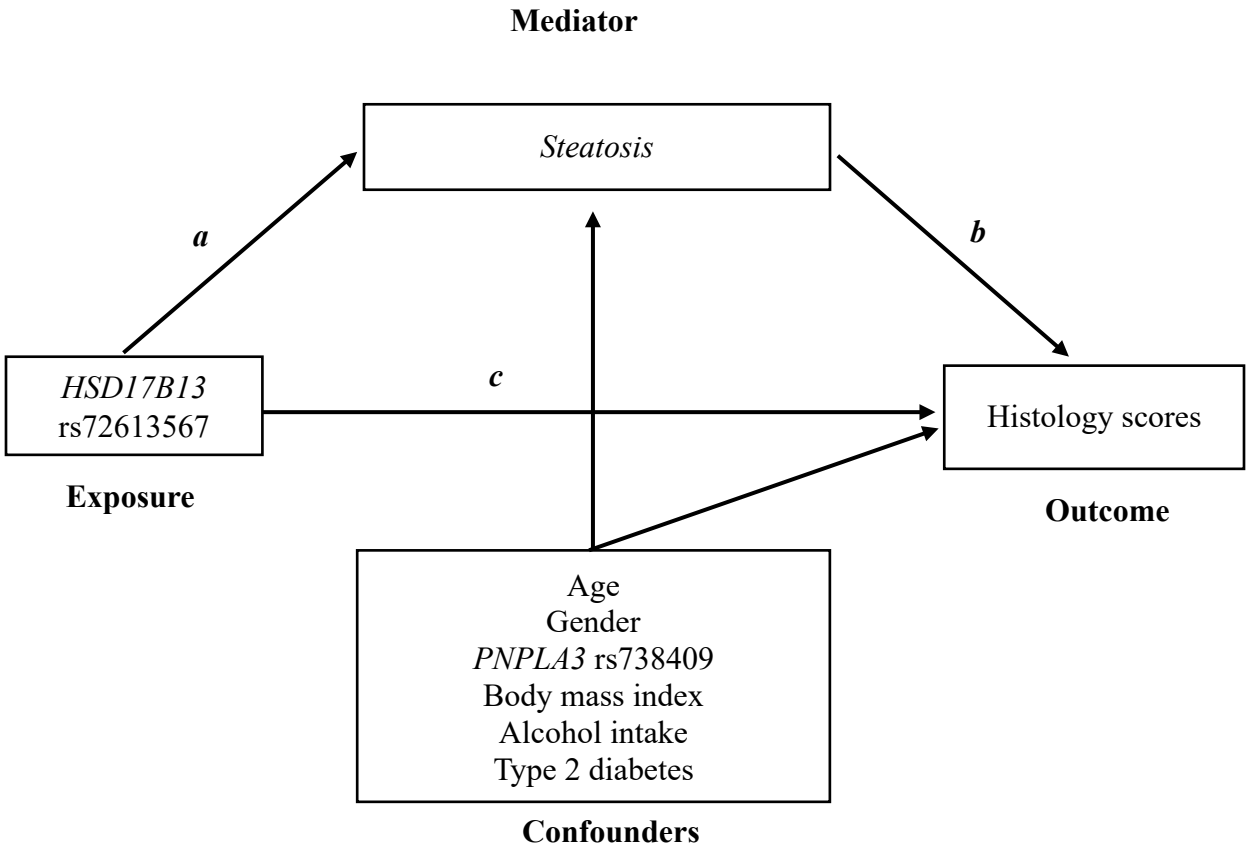
[1]Department of Medicine, Indiana University School of Medicine, Indianapolis, IN.
[2]Molecular Genetics and Biology of Complex Diseases and [3]Department of Clinical and Molecular Hepatology, Institute of Medical Research (IDIM), University of Buenos Aires-National Scientific and Technical Research Council (CONICET), Ciudad Autonoma de Buenos Aires, Argentina.
[4]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Table of contents Pages

MEDIATION ANALYSIS

The statistical rationale behind the mediation analysis

Mediation tests whether the effects of exposure on outcome operate through a third variable, mediator. In this way, mediators explain the causal relationship between two variables or "how" the relationship works.



In the diagram:

*The **total effect** of exposure on the outcome (c'=c + ab)*

*The **direct effect** of exposure on the outcome after controlling for mediators (c)*

*The **indirect effect** of exposure on the outcome (a + b)*

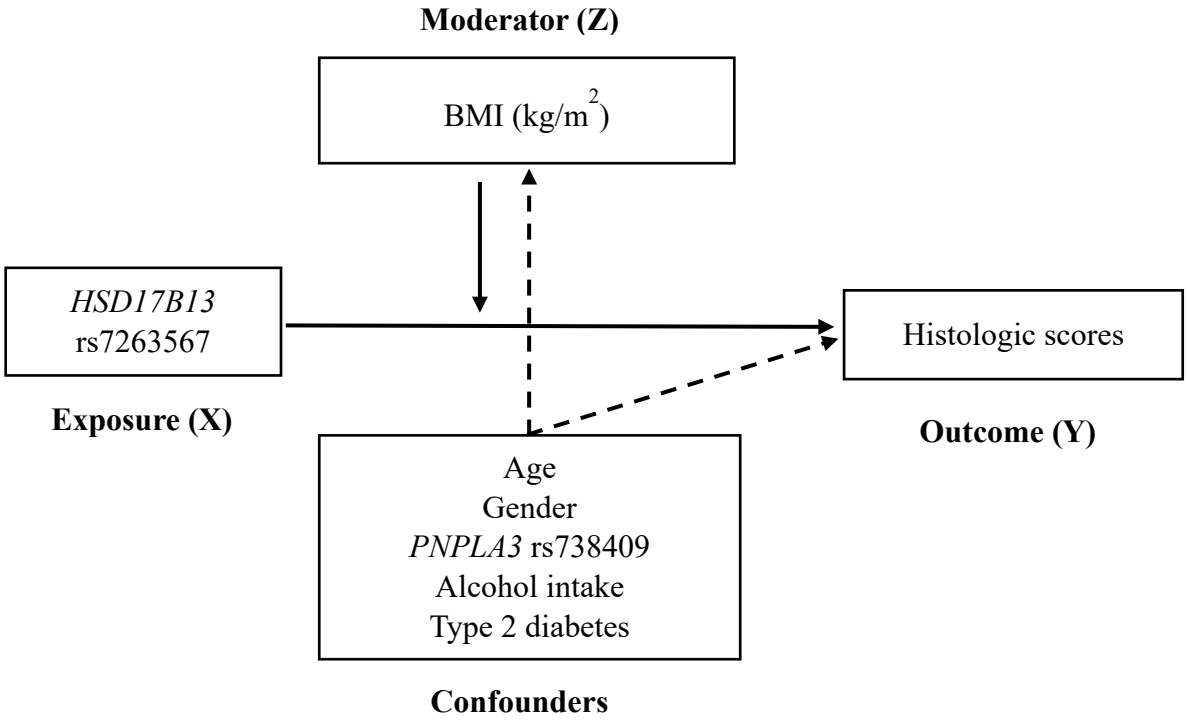
The indirect (mediating) effect is represented by the path from exposure to outcome via the mediator (paths a and b). An indirect (mediated) effect expresses the fraction of the exposure effect that is mediated through a specific mediator.

MODERATION ANALYSIS

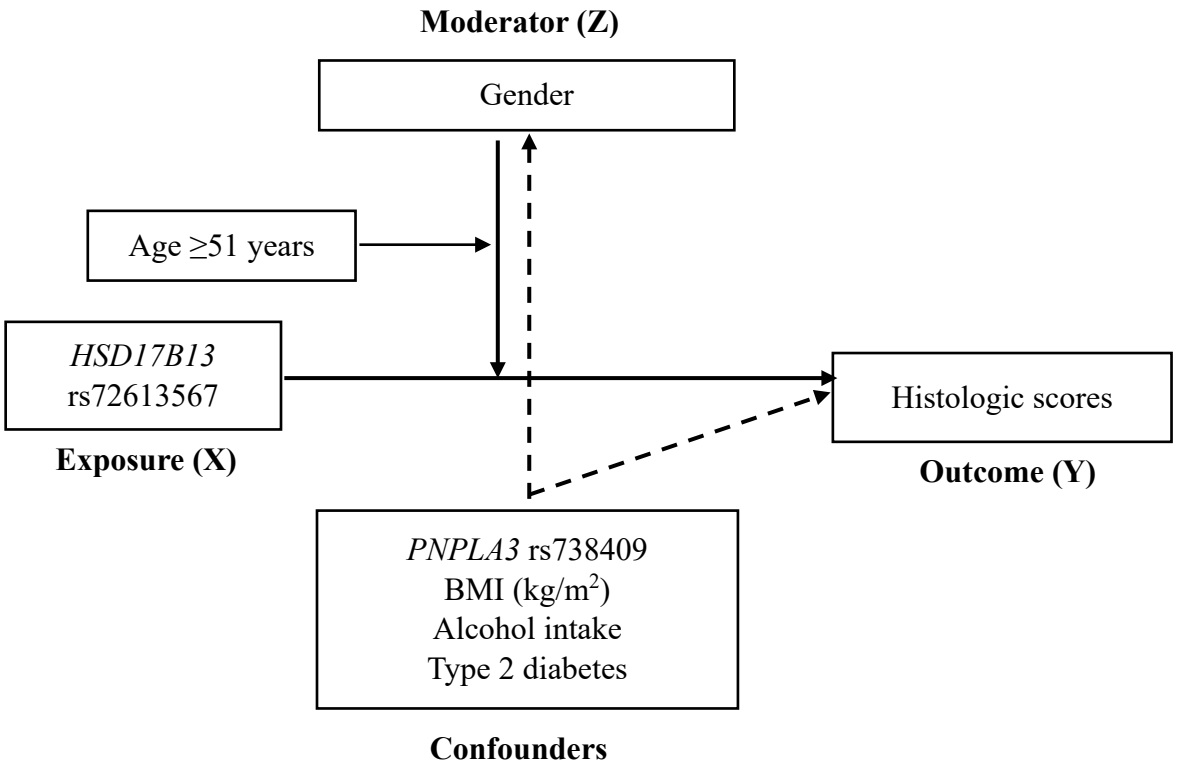
The statistical rationale behind the moderation analysis

Moderation analysis also allows to test for the influence of a third variable, Z, on the relationship between variables X and Y. Rather than testing a causal link between these other variables, moderation tests for when or under what conditions an effect occurs. Moderators can strengthen, weaken, or reverse the nature of a relationship. In practice, the analysis adds an interaction term (X*Z) to the model.

Model 1 (single model) (an example including BMI as a moderator)



Model 3 (includes two moderators: gender and age ≥51)



Note: this model allows to explore if moderator effects of gender on the relationship between *HSD17B13* rs72613567 and liver histology severity might differ at levels of a second moderator (age <51 or ≥51 years).

***PNPLA3* rs738409 AND *HSD17B13* rs72613567 GENOTYPING**

The *HSD17B13* rs72613567 is an insertion variant (-/A insertion) near the donor splice site of exon 6, predicted to result in a splice donor variant with altered function. The *HSD17B13* rs72613567 assay is from Integrated DNA Technologies (Coralville, IA) (catalog number 171165943). Probes were labeled with VIC or FAM fluorescent dye for allelic differentiation.

The PNPLA3 rs738409 (C>G) is a coding sequence variant, resulting in an Ile to Met substitution. The *PNPLA3* rs738409 assay is from Life Technologies Corporation (Grand Island, NY) (catalog number 4351379, assay ID number C_7241_10). *PNPLA3* rs738409 was genotyped using the TaqMan SNP genotyping system. Each reaction contained 5 µL 2X TaqMan Universal PCR Mastermix, No AmpErase UNG, 3.750 µL water, 0.250 µL 40X Assay Mix, and 1 µL sample DNA. *HSD17B13* rs72613567 was genotyped using the rhAmp genotyping system. Each reaction contained 5 µL 2X rhAmp Mastermix, 2.2 µL water, 0.250 µL rhAmp Reporter Mix with Reference, 0.500 µL 20x Assay Mix, and 2 µL sample DNA. Eight controls were included on each 96-well plate: two no-template controls, two replicate heterozygous samples, and two replicates of each of the homozygous samples. Because genotyping was determined by endpoint reading, the PCR was carried out in standard Applied Biosystems thermocyclers (AB2720, SimpliAmp, and Veriti, ThermoFisher). The PCR products were analyzed in an ABI PRISM® 7300 Sequence Detection System (SDS) instrument (ThermoFisher). SDS Software 1.3.1 was used to convert the raw data to pure dye components and to plot the results of the allelic discrimination on a scatter plot of Allele X versus Allele Y. The genotype of each sample was determined by the fluorescence levels of the VIC and FAM reporter dyes, and samples of the same genotype clustered together.

HARDY-WEINBERG EQUILIBRIUM AND SAMPLE SIZE CALCULATIONS

In individuals without clinically significant fibrosis, *HSD17B13* rs72623567 genotypes were in Hardy-Weinberg Equilibrium (-/-), (-/A) and (A/A), 347, 188, 36, respectively; estimated disequilibrium coefficient (D) = 0.0112, SE = 0.0077; Hardy-Weinberg Equilibrium Test: Pearson chi2 (1) = 2.322 P= .127, likelihood-ratio chi^2 (1) = 2.249, P= .134 or exact significance prob = .123.

Statistical power calculation for the association of *HSD17B13* rs72623567. The statistical power of our study was calculated based on the following assumptions: 500 cases (presence of clinically significant fibrosis) and 500 controls (absence of clinically significant fibrosis), even for a prevalence of clinically significant fibrosis of 0.40, an allele frequency of 0.20 in non-Hispanic whites, an OR of 1.30 (or its inverse 0.77), and a significant level of 0.05. Under an additive or dominant genetic model, our sample size has a statistical power of 98% and 89%, respectively.

Suppl. Table 1. Role of steatosis as a mediator between *HSD17B13* rs72613567 (-/A) + (A/A) with (-/-) as the reference group and histologic scores. Results based on mediation analysis.

Histologic scores	**Bootstrapped estimates (n=10,000)**			
	Mediator = steatosis (0-3)			
Lobular inflammation (0-3)	Estimate*	95% CI	SEs	P value
Total effect of rs72613567 (-/A)+(A/A)	-0.13	-0.22 to -0.06	0.04	<0.01
Direct effect of rs71623567 (-/A)+(A/A)	-0.16	-0.25 to -0.09	0.04	<0.001
Indirect effect of rs72613567 (-/A)+(A/A)	0.03	0.05 to 0.01	0.01	<0.01
Ballooning degeneration (0-2)	Estimate*	95 CI	SEs	P value
Total effect of rs72613567 (-/A)+(A/A)	-0.10	-0.004 to -0.20	0.05	0.039
Direct effect of rs71623567 (-/A)+(A/A)	-0.12	-0.22 to -0.02	0.05	0.014
Indirect effect of rs72613567 (-/A)+(A/A)	0.02	0.002 to 0.04	0.009	0.025
Steatohepatitis (0-2)	Estimate*	95% CI	SEs	P value
Total effect of rs72613567 (-/A)+(A/A)	-0.12	-0.22 to -0.03	0.04	<0.01
Direct effect of rs71623567 (-/A)+(A/A)	-0.16	-0.25 to -0.07	0.05	<0.001
Indirect effect of rs72613567 (-/A)+(A/A)	0.04	0.01 to 0.06	0.01	<0.01
Fibrosis (0-4)	Estimate*	95% CI	SEs	P value
Total effect of rs72613567 (-/A)+(A/A)	-0.17	-0.31 to -0.02	0.07	0.023
Direct effect of rs71623567 (-/A)+(A/A)	-0.17	-0.31 to -0.02	0.07	0.026
Indirect effect of rs72613567 (-/A)+(A/A)	-0.0004	-0.02 to 0.02	0.01	0.956

Abbreviations: SEs, standard errors.

All analyses were adjusted by age, gender, *PNPLA3* rs738409, body mass index, alcohol intake, and type 2 diabetes.

* Bootstrapped β coefficients.

Negative estimates [bootstrapped β coefficients]) indicate that *HSD17B13* rs72613567 (-/A) + (A/A) or steatosis degrees are inversely associated with NAFLD severity.

Positive estimates [bootstrapped β coefficients]) indicate that steatosis degrees are positively associated with NAFLD severity.

The reference genotype for all comparative analyses is *HSD17B13* rs72613567 (-/-).

Note: This causal mediation analysis confirms the following findings:

A) A direct and negative "protective" effect of *HSD17B13* rs72613567 (-/A) + (A/A) genotypes on inflammation and fibrosis.

B) The *HSD17B13* rs72613567 (-/A) + (A/A) "protective" effect on lobular inflammation, ballooning degeneration, and steatohepatitis seems to be mediated by steatosis degree in a positive direction (indirect effects show positive estimates). This indicates that steatosis degrees associate positively with histologic parameters related to inflammation.

C) Steatosis degrees did not mediate the "protective" effect of *HSD17B13* rs72613567 (-/A) + (A/A) on fibrosis. Total and direct effects of *HSD17B13* rs72613567 (-/A) + (A) on fibrosis were identical (-0.17). The indirect effect was -0.0004 with a P value of .956 which may suggest the "protective" effect of rs72613567 (-/A) + (A/A) on fibrosis is direct and not significantly mediated by steatosis.

Suppl. Table 2. Conditional effect of *HSD17B13* rs72613567 (-/A) or (A/A) genotypes (additive model) on aminotransferase levels and histology scores at levels of *PNPLA3* rs738409 genotypes (CC) or (GC+GG).

	Univariate analysis				Multivariable analysis†			
Moderators	Conditional effect*	Standard error	95% confidence interval	P value	Conditional effect*	Standard error	95% confidence interval	P value
rs738409 (CC)	***Steatohepatitis categories (0-2)***							
rs72613567 (-/A)	-0.18	0.09	-0.35 to -0.003	0.050	-0.15	0.08	-0.33 to 0.02	0.079
rs72613567 (A/A)	-0.34	0.15	-0.65 to -0.03	0.032	-0.34	0.15	-0.64 to -0.04	0.024
rs738409 (GC)+(GG)								
rs72613567 (-/A)	-0.09	0.06	-0.21 to 0.03	0.142	-0.08	0.06	-0.19 to 0.04	0.213
rs72613567 (A/A)	-0.07	0.13	-0.33 to 0.19	0.597	-0.12	0.12	-0.38 to 0.12	0.317
rs738409 (CC)	***Fibrosis (0-4)***							
rs72613567 (-/A)	-0.50	0.14	-0.78 to -0.21	<0.01	-0.44	0.13	-0.71 to -0.18	<0.01
rs72613567 (A/A)	-0.50	0.25	-0.99 to -0.003	0.049	-0.49	0.23	-0.95 to -0.03	0.038
rs738409 (GC)+(GG)								
rs72613567 (-/A)	-0.05	0.10	-0.25 to 0.14	0.592	-0.01	0.09	-0.19 to 0.17	0.907
rs72613567 (A/A)	0.09	0.21	-0.32 to 0.52	0.648	0.01	0.19	-0.37 to 0.40	0.950

***** Represents the effect of the exposure (*HSD17B13* rs72613567 (-/A) or (A/A) genotypes) on outcomes at values of the moderator (i.e., *PNPLA3* rs738409 genotypes (CC) or (GC) + (GG).

† Analysis adjusted for type 2 diabetes mellitus, non-heavy alcohol intake, age, gender, and body mass index.

The reference genotype for all comparative analyses is *HSD17B13* rs72613567 (-/-).

Suppl Table 3. Conditional effect of *HSD17B13* rs72613567 (-/A) or (A/A) genotypes (additive model) on aminotransferase levels and histology scores at levels of moderators (BMI < 35 or ≥ 35 kg/m^2).

	Univariate analysis				Multivariable analysis†			
Moderators	Conditional effect*	Standard error	95% confidence interval	P value	Conditional effect*	Standard error	95% confidence interval	P value
BMI <35	***Steatohepatitis categories (0-2)***							
rs72613567 (-/A)	-0.05	0.07	-0.19 to 0.08	0.426	-0.03	0.06	-0.16 to 0.11	0.674
rs72613567 (A/A)	-0.05	0.13	-0.32 to 0.22	0.719	-0.05	0.13	-0.31 to 0.21	0.691
BMI ≥35								
rs72613567 (-/A)	-0.20	0.07	-0.36 to -0.06	<0.01	-0.20	0.07	-0.35 to -0.05	<0.01
rs72613567 (A/A)	-0.33	0.15	-0.63 to -0.04	0.027	-0.39	0.14	-0.67 to -0.10	<0.01
BMI <35	***Fibrosis (0-4)***							
rs72613567 (-/A)	-0.12	0.11	-0.34 to 0.010	0.277	-0.07	0.10	-0.28 to 0.13	0.494
rs72613567 (A/A)	-0.38	0.22	-0.81 to 0.050	0.083	-0.32	0.20	-0.72 to 0.09	0.123
BMI ≥35								
rs72613567 (-/A)	-0.31	0.12	-0.56 to -0.07	0.012	-0.26	0.11	-0.48 to -0.03	0.026
rs72613567 (A/A)	0.11	0.24	-0.36 to 0.59	0.635	0.01	0.22	-0.43 to 0.45	0.948

Abbreviations: BMI, body mass index.

*Represents the effect of the exposure (*HSD17B13* rs72613567 (-/A) or (A/A) genotypes) on outcomes at values of the moderator (i.e., BMI <35 or ≥35). The reference genotype for all comparative analyses is *HSD17B13* rs72613567 (-/-).

† Analysis adjusted for type 2 diabetes mellitus, non-heavy alcohol intake, age, gender, and *PNPLA3* rs738409.

Suppl. Table 4. Conditional effect of rs72613567 (-/A) or (A/A) genotypes (additive model) on aminotransferase levels and histology scores at levels of moderators (Age < 45 or ≥ 45 years).

	Univariate analysis				Multivariable analysis†			
Moderators	Conditional effect*	Standard error	95% confidence interval	P value	Conditional effect*	Standard error	95% confidence interval	P value
Age <45	***Steatohepatitis categories (0-2)***							
rs72613567 (-/A)	0.04	0.09	-0.14 to 0.23	0.645	0.05	0.09	-0.13 to 0.24	0.569
rs72613567 (A/A)	-0.13	0.19	-0.51 to 0.25	0.506	-0.19	0.19	-0.56 to 0.18	0.321
Age ≥45								
rs72613567 (-/A)	-0.18	0.06	-0.30 to -0.06	<0.01	-0.17	0.06	-0.28 to -0.05	<0.01
rs72613567 (A/A)	-0.19	0.12	-0.42 to 0.04	0.110	-0.24	0.11	-0.46 to -0.01	0.041
Age <45	***Fibrosis (0-4)***							
rs72613567 (-/A)	-0.02	0.15	-0.33 to 0.28	0.872	0.01	0.14	-0.28 to 0.30	0.946
rs72613567 (A/A)	-0.22	0.31	-0.83 to 0.39	0.487	-0.28	0.29	-0.87 to 0.29	0.333
Age ≥45								
rs72613567 (-/A)	-0.25	0.09	-0.45 to -0.07	<0.01	-0.24	0.09	-0.42 to -0.05	0.010
rs72613567 (A/A)	-0.12	0.18	-0.49 to 0.24	0.504	-0.18	0.17	-0.53 to 0.16	0.310

***** Represents the effect of the exposure (*HSD17B13* rs72613567 (-/A) or (A/A) genotypes) on outcomes at values of the moderator (i.e., age <45 or ≥45).The reference genotype for all comparative analyses is *HSD17B13* rs72613567 (-/-).

† Analysis adjusted for type 2 diabetes mellitus, non-heavy alcohol intake, *PNPLA3* rs738409, gender, and body mass index.

Suppl. Table 5. Conditional effect of *HSD17B13* rs72613567 (-/A) or (A/A) genotypes (additive model) on aminotransferase levels and histology scores at levels of moderators (*female or male*).

	Univariate analysis				Multivariable analysis†			
Moderators	Conditional effect*	Standard error	95% confidence interval	P value	Conditional effect*	Standard error	95% confidence interval	P value
Male	***Steatohepatitis categories (0-2)***							
rs72613567 (-/A)	-0.05	0.08	-0.21 to 0.11	0.530	-0.03	0.08	-0.19 to 0.12	0.643
rs72613567 (A/A)	0.07	0.17	-0.27 to 0.42	0.678	0.05	0.17	-0.28 to 0.39	0.747
Female								
rs72613567 (-/A)	-0.15	0.06	-0.28 to -0.02	0.021	-0.14	0.06	-0.27 to -0.01	0.029
rs72613567 (A/A)	-0.30	0.12	-0.55 to -0.06	0.013	-0.35	0.12	-0.58 to -0.11	<0.01
Male	***Fibrosis (0-4)***							
rs72613567 (-/A)	-0.22	0.13	-0.48 to 0.05	0.103	-0.17	0.12	-0.42 to 0.07	0.169
rs72613567 (A/A)	0.06	0.28	-0.49 to 0.63	0.812	0.10	0.26	-0.42 to 0.62	0.702
Female								
rs72613567 (-/A)	-0.18	0.10	-0.38 to 0.030	0.092	-0.14	0.09	-0.33 to 0.05	0.166
rs72613567 (A/A)	-0.27	0.19	-0.66 to 0.12	0.177	-0.32	0.18	-0.68 to 0.04	0.084

*Represents the effect of the exposure (*HSD17B13* rs72613567 (-/A) or (A/A) genotypes) on outcomes at values of the moderator (i.e., female or male).

† Analysis adjusted for type 2 diabetes mellitus, non-heavy alcohol intake, age, *PNPLA3* rs738409, and body mass index. The reference genotype for all comparative analyses is *HSD17B13* rs72613567 (-/-).

Suppl. Table 6. Conditional effect of *HSD17B13* rs72613567 (-/A) or (A/A) genotypes (additive model) on aminotransferase levels and histology scores at levels of moderators (type 2 diabetes [yes or no]).

	Univariate analysis				Multivariable analysis†			
Moderators	Conditional effect*	Standard error	95% confidence interval	P value	Conditional effect*	Standard error	95% confidence interval	P value
T2D (no)	***Steatohepatitis categories (0-2)***							
rs72613567 (-/A)	-0.08	0.06	-0.21 to 0.03	0.164	-0.07	0.06	-0.19 to 0.04	0.208
rs72613567 (A/A)	-0.10	0.13	-0.36 to 0.14	0.405	-0.13	0.12	-0.38 to 0.12	0.302
T2D (yes)								
rs72613567 (-/A)	-0.14	0.08	-0.31 to 0.03	0.100	-0.15	0.08	-0.32 to 0.02	0.090
rs72613567 (A/A)	-0.33	0.15	-0.63 to -0.02	0.033	-0.33	0.15	-0.63 to -0.04	0.028
T2D (no)	***Fibrosis (0-4)***							
rs72613567 (-/A)	-0.09	0.09	-0.28 to 0.10	0.359	-0.06	0.09	-0.25 to 0.12	0.509
rs72613567 (A/A)	-0.15	0.20	-0.55 to 0.25	0.464	-0.12	0.19	-0.51 to 0.26	0.267
T2D (yes)								
rs72613567 (-/A)	-0.34	0.13	-0.61 to -0.07	0.013	-0.32	0.13	-0.58 to -0.06	0.015
rs72613567 (A/A)	-0.28	0.24	-0.75 to 0.19	0.241	-0.27	0.23	-0.73 to 0.18	0.239

Abbreviations: T2D, type 2 diabetes.

***** Represents the effect of the exposure (*HSD17B13* rs72613567 (-/A) or (A/A) genotypes) on outcomes at values of the moderator (i.e., type 2 diabetes [yes or no]).

The reference genotype for all comparative analyses is *HSD17B13* rs72613567 (-/-) in an additive model.

† Analysis adjusted non-heavy alcohol intake, *PNPLA3* rs738409, gender, body mass index, and age.

Suppl. Table 7. Effect of *HSD17B13* rs72613567 on histologic scores. Benjamini-Hochberg false discovery rate (FDR) to correct for multiple comparisons. Analysis under an additive genetic model.

Histologic scores (ordinal)	95% confidence interval of estimates	Q value*
Steatosis (0-3)	0.03 to 0.12	<0.001
Lobular inflammation (0-3)	-0.14 to -0.04	<0.001
Ballooning degeneration (0-2)	-0.09 to -0.02	0.037
Steatohepatitis (0-2)	-0.12 to -0.01	0.011
Fibrosis (0-4)	-0.09 to -0.008	0.036

*Q values calculated for multiple comparisons.

Analysis adjusted for *PNPLA3* rs738409, non-heavy alcohol intake, age, body mass index, type 2 diabetes mellitus, and gender.

The reference genotype for all comparative analyses *HSD17B13* rs72613567 (-/-).

Suppl. Table 8. Effect of *HSD17B13* rs72613567 (-/A) + (A/A) on histologic scores. Benjamini-Hochberg false discovery rate (FDR) to correct for multiple comparisons. Analysis under a dominant genetic model.

Risk factors	95% confidence interval of estimates	Q value*	95% confidence interval of estimates	Q value*
	Steatohepatitis (0-2)		**Fibrosis (0-4)**	
PNPLA3				
rs738409 (CC)	-0.21 to -0.02	0.031	-0.20 to -0.06	<0.001
rs738409 (GC) + (GG)	-0.10 to 0.02	0.175	-0.05 to 0.04	0.738
Female	-0.16 to -0.03	0.027	-0.10 to -0.001	0.049
Male	-0.09 to 0.06	0.692	-0.11 to 0.02	0.225
BMI≥35	-0.21 to -0.05	0.010	-0.12 to -0.004	<0.001
BMI<35	-0.09 to 0.04	0.470	-0.08 to 0.07	0.201
Age ≥45	-0.15 to -0.03	<0.01	-0.11 to -0.01	0.015
Age <45	-0.08 to 0.11	0.774	-0.08 to 0.07	0.814
T2D (yes)	-0.21 to -0.003	0.026	-0.15 to -0.02	0.015
T2D (no)	-0.10 to 0.02	0.303	-0.08 to 0.02	0.304

Abbreviations: BMI, body mass index; T2D, type 2 diabetes.

*****Q values calculated for multiple comparisons.

† Analysis adjusted for *PNPLA3* rs738409, non-heavy alcohol intake, and those covariates not tested as moderators.

The reference genotype for all comparative analyses *HSD17B13* rs72613567 (-/-).

SUPPLEMENTAL FIGURES LEGENDS

Suppl. Figure 1. Flow diagram through the study.

Suppl. Figure 2. Effects of *HSD17B13* rs72613567 genotypes (-/-), (-/A), and (A/A) on the risk of significant fibrosis or steatohepatitis at levels of moderators' cutoffs. Analysis performed under an additive genetic model of *HSD17B13* rs72613567.

A) *PNPLA3* rs738409 (CC) and (GC) + (GG).

B) Body mass index (<35 and ≥35) kg/m^2.

C) Age (<45 and ≥45) years.

D) Gender (female and male).

E) Type 2 diabetes (yes or no).

Note: Risk of steatohepatitis or significant fibrosis are calculated by taking *HSD17B13* rs72613567 (-/-) as the reference genotype.