



Supplementary Fig. 1 Short hairpin RNA-mediated knockdown of Tm6sf2 in mice. (a) AAV vector alone or AAV expressing a shRNA directed against mTm6sf2 mRNA were injected into the tail veins of 8 week old chow-fed C57Bl/6J male mice ($n=5$ /group). After two weeks, the mice were killed and the liver, white adipose tissue (WAT) and small intestine were harvested. Levels of Tm6sf2 mRNA were measured using Real-Time PCR and normalized to m36B4 levels. The differences in mean expression levels were compared using a two sample t-test. (b) AAV expressing 3 different shRNAs or AAV vector alone were injected intravenously into 8-week old chow-fed C57BL/6J male mice ($n=6$ /group). All three shRNAs were directed against mouse Tm6sf2, but only shRNA8 knocked down the levels of Tm6sf2 mRNA with high efficiency in cultured cells. Two weeks after the injection, mice were fasted for 4 h, the livers were harvested and the levels of TM6SF2 mRNA were measured using Real-Time PCR. Values were normalized to m36B4, and expressed relative to the mean value of the AAV-only treated mice. Plasma was collected in the same experiment and TG and cholesterol levels measured as described in Methods. Differences in mean expression levels were compared using a two-sample T-tests. Ct, cycle threshold value.

Supplementary Table 1 Characteristics of Dallas Heart Study participants stratified by *TM6SF2* (E167K) genotype.

Characteristic	EE ^a	EK	KK	P-value ^b
Total n	4,151	423	13	-
Female, n (%)	2,389 (57.6)	225 (53.2)	9 (69.2)	0.16
Ethnicity ^c , n (%)				
Non-Hispanic African-American	2,203 (53.1)	160 (37.8)	2 (15.4)	2.1×10^{-10}
Non-Hispanic European-American	1,169 (28.2)	174 (41.1)	11 (84.6)	3.9×10^{-11}
Hispanic	676 (16.3)	69 (16.3)	0 (0)	0.35
Other	103 (2.5)	20 (4.7)	0 (0)	0.029
Age, years (mean)	45 ± 11	44 ± 11	50 ± 12	0.73
Body-mass index, kg/m ²	29.6 (26-35)	28.5 (26-34)	31.8 (28-34)	0.95
Diabetes ^d , n (%)	523 (12.6)	37 (8.7)	2 (16.7)	0.14
Alcohol consumption ^e >20 g/day, n (%)	334 (8.1)	48 (11.4)	1 (7.7)	0.058
Lipid-lowering medication (statins), n (%)	359 (8.8)	33 (8.0)	0 (0)	0.17
HDL-cholesterol (mg/dL)	50.6 ± 14.9	49.9 ± 14.0	50.2 ± 18.8	0.98
HOMA-IR (U) ^f	3.0 (2-5)	2.9 (2-5)	4.6 (2-8)	0.52
Fasting glucose (mg/dL) ^g	91 (84-98)	91 (85-99)	94 (86-104)	0.15
Gamma-glutamyl transpeptidase (IU/L)	23 (17-36)	24 (16-38)	22 (12-28)	0.51
Bilirubin (mg/dL)	0.5 (0.4-0.7)	0.5 (0.4-0.7)	0.6 (0.5-0.8)	0.92
HTGC (%) ^h	3.5 (1.9-6.7)	4.4 (2.2-9.9)	15.7 (8.8-18.4)	5.6×10^{-7}

Values are median (interquartile range), mean ± s.d., or number (%). ^aMissense variant (E167K) in *TM6SF2*. ^bP-values were calculated using linear regression for continuous variables (age, body-mass index, glucose, HOMA-IR) and logistic regression for categorical variables (gender, ethnicity, diabetes diagnosis, alcohol consumption, and use of lipid-lowering drugs). Regression models were adjusted for age, gender, ethnicity, and BMI (where appropriate). ^cEthnicity was self-reported as described in the Methods. ^dDiabetes was defined as a self-reported diagnosis of diabetes, use of prescription medication, fasting glucose ≥ 126 mg/dL or non-fasting glucose ≥ 200 mg/dL or HbA1c ≥ 6.5%. ^eAlcohol consumption was self-reported. ^fHOMA-IR, homeostatic model assessment of insulin resistance was calculated as described in the Methods. ^gDiabetics excluded. ^hP-value after adjustment for HOMA-IR, alcohol consumption and covariates listed above.

Supplementary Table 2 Conditional association analysis for variants in the chromosome 19 region including *NCAN*, *TM6SF2*, *SUGP1*, *CILP2*, and *PBX4*

Locus	SNP	MAF (%)	Median Hepatic TG % by genotype			Beta	<i>P</i> Uncondi- tional	<i>P</i> _{SNP} condi- tional on rs58542926	<i>P</i> _{rs58542926} conditional on SNP
			0	1	2				
NCAN	rs2228603 (P92S)	3.7	3.5	3.9	12.0	0.076	0.002	0.79	1.3×10^{-5}
NCAN	rs150701551 (V548L)	0.64	3.6	3.3	-	0.033	0.57	0.52	5.5×10^{-8}
NCAN	rs10426537 (S838N)	3.9	3.6	2.8	3.2	-0.060	0.013	0.015	6.8×10^{-8}
NCAN	rs1064395	29.7	3.7	3.5	3.1	0.007	0.51	0.20	3.1×10^{-8}
TM6SF2	rs58542926 (E167K)	5.0	3.5	4.5	15.7	0.115	5.7×10^{-8}	NA	NA
SUGP1	rs10401969	12.4	3.6	3.6	3.3	0.035	0.014	0.26	7.2×10^{-7}
SUGP1	rs17751061 (R290H)	7.9	3.5	3.9	2.4	0.021	0.24	0.12	3.7×10^{-8}
SUGP1	rs138023808 (D104D)	0.7	3.6	3.3	-	0.027	0.62	0.56	5.5×10^{-8}
CILP2, PBX4	rs16996148	11.8	3.6	3.3	5.0	0.025	0.084	0.22	1.2×10^{-7}
CILP2, PBX4	rs17216525	5.3	3.5	4.2	18.1	0.073	0.0005	0.24	1.6×10^{-5}

Variants within 10 kb of the coding sequences of *NCAN*, *TM6SF2*, and *SUGP1* genotyped using the ExomeChip array, with a minor allele frequency of >0.5%, and other variants at the 19p13 locus reported in previous GWAS, are listed in the table. Genotypes were coded as 0, 1, and 2 for wild-type homozygotes, heterozygotes, and variant allele homozygotes, respectively. Each SNP was tested for association with hepatic TG content using linear regression including ancestry, age, gender, and BMI as covariates. Conditional analysis was also adjusted for *TM6SF2* rs58542926 genotype. NA, not applicable.

Supplementary Table 3 Association between *TM6SF2* (E167K) and hepatic triglyceride content (HTGC), liver enzymes and plasma lipid levels in Dallas Heart Study participants stratified by *TM6SF2* (E167K) genotype and (self-reported) ethnicity.

Trait	Ethnicity	EE ^a	EK	KK	P-value ^b
HTGC (%) mean	DHS	5.79 (0.14)	8.10 (0.60)	13.20 (2.63)	5.7 x 10 ⁻⁸
	Non-Hispanic African-American	4.78 (0.17)	6.52 (0.86)	2.15 (-)	0.0058
	Non-Hispanic European-American	5.86 (0.25)	8.63 (0.98)	15.04 (2.23)	2.7 x 10 ⁻⁶
HTGC (%) median	Hispanic	8.35 (0.41)	9.14 (1.38)	-	0.91
	DHS	3.46 (2-7)	4.49 (2-10)	15.70 (9-18)	5.2 x 10 ⁻⁵
	Non-Hispanic African-American	3.15 (2-5)	4.17 (2-7)	2.15 (-)	0.020
ALT (U)	Non-Hispanic European-American	3.50 (2-7)	4.29 (2-9)	16.87 (13-19)	0.0044 ⁶
	Hispanic	4.74 (3-11)	6.22 (3-13)	-	0.77
	Non-Hispanic African-American	22 ± 20.1	23.3 ± 16.8	13.0 ± NA	0.78
AST (U)	Non-Hispanic European-American	23.7 ± 16.4	26.2 ± 18.5	32.4 ± 28.4	0.084
	Hispanic	27.7 ± 25	30.7 ± 19.9	-	0.015
	Non-Hispanic African-American	24.1 ± 17.3	25.5 ± 27.7	18.0 ± NA	0.19
ALP (U)	Non-Hispanic European-American	23.9 ± 17.4	24.4 ± 12.1	27.3 ± 17.5	0.95
	Hispanic	25.7 ± 33	26.2 ± 12.6	-	0.23
	Non-Hispanic African-American	72.6 ± 29	71.4 ± 21.3	61.0 ± NA	0.063
LDL (mg/dL)	Non-Hispanic European-American	66.6 ± 20.5	64.3 ± 20.4	63.4 ± 17.8	0.87
	Hispanic	77.2 ± 26.3	71.3 ± 19.2	-	0.12
	Non-Hispanic African-American	108.1 ± 37	101.7 ± 34.8	61.0 ± NA	0.033
Triglyceride (mg/dL)	Non-Hispanic European-American	110.8 ± 34.9	108.7 ± 33.9	96.3 ± 41.2	0.22
	Hispanic	108.7 ± 33.1	104.8 ± 33.6	-	0.21
	Non-Hispanic African-American	107.2 ± 95.1	96.9 ± 55.2	67.0 ± NA	0.041
HDL (mg/dL)	Non-Hispanic European-American	136.2 ± 98.4	126.3 ± 76.6	132.2 ± 68	0.26
	Hispanic	149.6 ± 118.7	144.6 ± 160.8	-	0.46
	Non-Hispanic African-American	52.6 ± 15	52 ± 13.2	65.0 ± NA	0.85
HDL (mg/dL)	Non-Hispanic European-American	49.7 ± 15.4	49.5 ± 15.3	50.6 ± 19.2	0.76
	Hispanic	46.4 ± 11.5	46.9 ± 11.5	-	0.73

Values are median (interquartile range, 1st – 3rd quartile) or mean (s.e.m) for HTGC, and mean \pm s.d. for all other traits. P-values for comparison of means were calculated using linear regression adjusted for age, gender, BMI and ancestry. P-values for median HTGC were determined using the Jonckheere-Terpstra rank test for trend. Ethnicity was self-reported as described in the Methods.

Supplementary Table 4 Characteristics of the Dallas BioBank participants stratified by *TM6SF2* (E167K) genotype.

Characteristic	EE ^a	EK	KK	P-value ^b
Total n	7,416	1,112	57	-
Female, n (%)	2,265 (30.5)	345 (31.0)	18 (31.6)	0.72
Age, years	53 \pm 11	54 \pm 11	54 \pm 10	0.069
Body-mass index, kg/m ²	26.2 (24-29)	26.1 (24-29)	26.1 (24-29)	0.71
Diabetes ^c , n (%)	173 (2.3)	27 (2.4)	2 (3.5)	0.62
Lipid-lowering medication (statins), n (%)	443 (6.0)	70 (6.3)	6 (10.5)	0.37
Fasting glucose (mg/dL) ^d	92 (87-98)	93 (88-99)	91 (86-96)	0.041
Bilirubin (mg/dL)	0.56 (0.40-0.76)	0.56 (0.42-0.74)	0.54 (0.39-0.76)	0.71
HDL-cholesterol (mg/dL)	57.1 \pm 17.2	57.8 \pm 16.7	56.3 \pm 19.3	0.14

Values are median (interquartile range), mean \pm s.d., or number (%). ^aMissense variant (E167K) in *TM6SF2*. ^bP-values were calculated using linear regression for continuous variables (age, body-mass index, glucose, HOMA-IR) and logistic regression for categorical variables (gender, ethnicity, diabetes diagnosis, alcohol consumption, and use of lipid-lowering drugs). Regression models were adjusted for age, gender, and BMI, where appropriate. ^cDiabetes was defined as a self-reported diagnosis of diabetes by a physician, use of prescription medication, fasting glucose \geq 126 mg/dL or non-fasting glucose \geq 200 mg/dL or HbA1c \geq 6.5%. ^eAlcohol consumption was self-reported. ^dDiabetics excluded.

Supplementary Table 5 Baseline characteristics of participants in the Copenhagen General Population Study and the Copenhagen City Heart Study combined (The Copenhagen Study), stratified by *TM6SF2* (E167K) genotype.

Characteristics	EE	EK ^a	KK	P-value ^b
Total n	61,279	11,700	553	-
Age, years	58 (47-67)	58 (47-67)	56 (45-66)	0.02
Female, n (%)	34,075 (56)	6,377 (55)	315 (57)	0.07
Body mass index, kg/m ²	26 (23-28)	26 (23-28)	25 (23-28)	0.14
Diabetes ^c	3,025 (5)	588 (5)	34 (6)	0.40
Physical activity ^c	28,858 (48)	5,459 (47)	267 (49)	0.60
Alcohol consumption ^c	10,874 (18)	2,141 (18)	103 (19)	0.31
Gamma-glutamyltransferase, U/L	28 (21-42)	28 (21-42)	28 (21-42)	0.16
Bilirubin, μ mol/L	10 (8-13)	10 (8-13)	11 (8-14)	0.87
Lipid-lowering therapy ^c (statins), n (%)	5,693 (9)	1,023 (9)	45 (8)	0.13
HDL (mg/dL)	62.2 \pm 19.7	62.1 \pm 19.6	62.6 \pm 19.4	0.67

Values are median (interquartile range), mean \pm s.d., or number (%). ^aMissense variant (E167K) in *TM6SF2*. ^bP-values were calculated using Kruskal-Wallis analysis of variance or Pearson's χ^2 -test. ^cDiabetes, physical activity in leisure time, alcohol consumption, and lipid-lowering therapy were self-reported, dichotomized, and defined as diabetes (self-reported disease, use of insulin, use of anti-diabetic medication, and/or non-fasting plasma glucose > 11.0 mmol/L) versus no diabetes, physical activity (four hours or more per week of light physical activity in leisure time versus less than four hours), alcohol consumption (>14/21 versus \leq 14/21 units per week in women/men; 1 unit=12g alcohol), and use of lipid-lowering drugs (yes/no), mainly statins (>97%).

Supplementary Table 6. Primers used for amplification using the polymerase chain reaction

Sequence Amplified		Forward Primer		Reverse Primer
TM6SF2-rs58542926	1	GCATGGCACCAGCAGGTA	2	CCTGCACCATGGAAGGCAAATA
Mouse TM6SF2 cDNA	3	CCTCGGTGGTGGACCTTG'	4	TCCTTGGTGTAGAAATCCATGAAG
Human TM6SF2 cDNA	5	GGCTGCCTATGCTCTCACCTT	6	TGCCTCCAGCAAACACCAA