SUPPLEMENTARY DATA

SUGP1 is a novel regulator of cholesterol metabolism

DETAILED METHODS

Cell culture

HepG2, Hep3B and Huh7 cells were obtained from the American Type Culture Collection (ATCC) and grown in EMEM (ATCC) supplemented with 10% heat inactivated fetal bovine serum (FBS) (Hyclone). Cultures were maintained at 37°C with 5% CO_2. LDL was isolated as previously described [1]. Cells were exposed in replicate to conditions of sterol depletion (10% lipoprotein deficient serum and 2 μM activated simvastatin), and after 24 hours, 50 μg/ml LDLC was added for an additional 24 hours. Immortalized lymphoblastoid cell lines (n=54) derived from donors of the Cholesterol and Pharmacogenetics (CAP) clinical trial that either carried the minor allele of rs10401060 (CT) or homozygous non-carriers (TT) were grown at 37°C, 5% CO_2 in RPMI Medium 1640 (Invitrogen) supplemented with 10% fetal bovine serum (Hyclone), 500 units/ml Penicillin/Streptomycin and 2nM GlutaMAX (Invitrogen).

siRNA transfections and gene expression measurements

SUGP1, ATP13A1 and *MAU2* knock-down were achieved by 48 hour transfection of 80,000 HepG2 or Huh7 cells/well in 12-well plates using either the Ambion Silence Select siRNAs (SUGP1-1: s33721, SUGP1-3: s33722, ATP13A1: s32750, MAU2: s225955) or non-targeting control (AM4611) using the pSPORT transfection reagent according to the manufacturer's protocol. Transfection time length was optimized through a time course analysis (16-48hr), which showed that the greatest level of knock-down was achieved 48hr post-transfection. 48hr after knock-down, RNA was isolated using the PureLink RNA mini kit (Invitrogen), and cDNA was synthesized using the high capacity cDNA archive kit (Applied Biosystems). Transcripts were quantified using TaqMan and SYBR Green assays as shown in **Table S1**. Each real time PCR reaction was performed in triplicate on an ABI PRISM 7900 Sequence Detection System with standard reagents all purchased from Applied Biosystems. Values were normalized to *CLPTM*, quantified as a loading control as previously described [2].

Rescue experiments

HepG2 cells were transfected in replicate with SUGP1-1 siRNA or NTC as described above. 48 hours post-transfection, SUGP1-1 siRNA treated cells were re-plated in 6-well plates at a density of 0.2x10E6 cells/well and transfected with 2μg of the SUGP1 overexpression plasmid (pCMV-SUGP1) or an empty vector control (pCMV), both purchased from Origene, and incubated for an additional 48 hours.

Cellular phenotyping

Cell culture media was collected from all samples at time of harvest, and APOB and APOE were quantified in triplicate by sandwich-style ELISA. Samples with a coefficient of variation greater than 15% were subject to repeat measurement. HMGCR enzyme activity was quantified as the rate of incorporation of 14C-HMG-CoA into 14C-mevalonate and normalized to total cellular protein as previously described [2]. To quantify rate of DiI-LDL uptake, cells were incubated for 3 hours with 10μg DiI-LDL, washed three times, and the geometric mean of the fluorescence values in the 10,000 gated events were determined by FACS. To measure cholesterol synthesis, 48 hr-post-knock-down cells were given 10 mM [1-^{13}C] sodium acetate (Sigma) for an additional 24 hr. At the end of the incubation, cells were washed with PBS, collected, and cholesterol quantification was performed as previously described [3]. Briefly, stigmasterol was added as internal standard to each sample. Sterols were then saponified at 55^{o}C for 45 min with KOH ethanolic solution, and extracted with hexane. The organic layers were dried under nitrogen, and sterols were converted to their trimethylsilyl derivatives by adding N,O-bis(trimethylsilyl)trifluoroacetamide (Sigma) and incubated at 55^{o}C for 45 min. The derivatives were solvate with cyclohexane and analyzed by GC/MS using an Agilent 5975 instrument. Total ^{13}C-labeled

cholesterol from each sample was quantified with correction for natural enrichment of ^{13}C, normalized by cellular protein, and expressed as μg cholesterol per mg protein.

In vivo study

Sugp1 was PCR-amplified from the pCMV-Entry vector (NM_027481.2, Origene) containing the expression construct. The amplified product and the pLIVE-MCS destination vector (Mirus Bio) were restriction enzyme digested and ligated by T4 ligase (New England Biolabs). The resulting expression construct (*Sugp1*-pLIVE) and empty vector (pLIVE-MCS with no insert) were purified using the Endofree Plasma Mega kit (Qiagen).

Sugp1-pLIVE was injected (25ug) into the tail vein of 6-wk old CD-1 male mice (n=10, Charles River Laboratories) using the TransIT EE hydrodynamic gene delivery system (Mirus Bio), as previously described [4]. In a second set of CD-1 male mice (n=10), the empty vector was injected to serve as the negative control. Five mice per construct each were euthanized at 7 and 28 days after injection by cardiac puncture. Animal studies were approved by the Children's Hospital Oakland Research Institute Animal Care and Use Committee and conform to all regulatory standards. Whole blood was collected in EDTA-coated microtubes (Becton-Dickinson) and livers were harvested and snap frozen in liquid nitrogen. Plasma was collected by centrifugation of whole blood at 3000 rpm at 4°C for 15 minutes. Plasma cholesterol was quantified by LIASYS. Hepatic RNA was extracted and cDNA was synthesized as previously described above. Transcripts were quantified using TaqMan and SYBR Green assays as shown in **Table S1**. Values were normalized to the geomean of *Clptm1* and *Rplp0*, quantified as loading controls, as described [2]. Hepatic Sugp1 protein levels were quantified by Western blot analysis as described below using mouse Sugp1 and Gadph antibodies (Santa Cruz Biotechnology).

Identification of novel SUGP1 splice variants

RNA was isolated from Human hepatoma cell lines and reverse transcribed into cDNA for PCR. SUGP1 exon 7 forward primer 5' AAG TTG GCC AGG TTC ATA GC 3' and SUGP1 exon 10 reverse primer 5' TGG TCT TAT CCA TCT CCA TGC 3' were used for touchdown PCR program with Phusion High-Fidelity DNA polymerase (New England BioLabs) to detect novel splice variants underlying rs10401969. The PCR products were subject to 0.8% agarose gel for electrophoresis, and the bands were gel extracted by Qiagen gel extraction kit for Sanger sequencing.

RNA electrophoretic mobility gel shift assay (RNA EMSA)

3' end-biotinylated RNA oligonucleotides were synthesized from Sigma-Aldrich. Purified His-tagged human recombinant HNRNPA1 and SRSF1 proteins were purchased from Prospec and ProteinOne. RNA gel shift assay was carried out using LightShift Chemiluminescent RNA EMSA kit (Thermo Scientific). RNA oligos with either the SUGP1 rs10401969 "T" or "C" allele were incubated for 30 minutes in the presence or absence of either His-HNRNPA1 or His-SRSF1 protein in 1x binding buffer, 5% glycerol, 2 μg tRNA, and 6.25nM biotin-labeled RNA oligos. 5x loading dye was added to each reaction, loaded onto a 4-20% Novex® TBE Gels (Life Technologies), and run at 100 V for 2 hours in 0.5x Novex® TBE Running Buffer (Invitrogen) at 4°C. The RNA-protein complexes were transferred from the gel to a nylon membrane (Thermo Scientific) at 35V for 40 minutes in a 4°C cold room, and crosslinked using the Stratalinker UV crosslinker (Stratagene). Biotin-labeled RNA oligos were detected by the Chemiluminescent Nucleic Acid detection Module (Thermo Scientific), and quantified using the GelQuant.NET software.

SUGP1 and HMGCR mini-gene construct

To create the *SUGP1* and *HMGCR* mini-genes, the segments of the 1.2 Kbp-*SUGP1* and 2 Kbp-*HMGCR* genes were amplified from *SUGP1* intron 7 to 8 using 5' CACCGTTCTGATGTTCATTGTCTCAA 3' and 5'

CACACCATTCTCCTGCCTC 3' and *HMGCR* introns 16 to 19 using 5' CACCGTGAGTGACTGGATGGATAAT 3' and 5' CTATGGTAGAAAATGACAAAGT -3' forward and reverse primers, cloned into the pENTR™TOPO® vector (Invitrogen), and then subcloned into the pDEST exon-trap vector provided by Dr. Stephan Stamm using the Gateway® LR clonase™ enzyme mix (Invitrogen). The plasmid was DNA sequence verified. HepG2 and Huh7 cells were transfected with either *SUGP1* mini-genes alone or *SUGP1* mini-genes and pCMV Human HNRNPA1 overexpression plasmid for 48 hours. Then, cells were harvested to measure mini-genes transcripts. For HMGCR mini-gene assays, HepG2 and Huh7 cells were first transfected with either one of the two *SUGP1* targeting siRNA or NTC for 24 hours, after which cells were subsequently transfected with the *HMGCR* mini-gene for an additional 24 hours.

Identification of novel HMGCR splice variants

Lymphoblastoid cell lines (LCLs) from 8 African American individuals (7 female, 1 male) were treated with 2 μM simvastatin or sham buffer for 24 hours. RNA was extracted, pooled, and paired-end RNA-seq libraries were prepared following Illumina's protocol with a few modifications. Modifications included a double purification of mRNA from total RNA using the MicroPoly(A)Purist Kit (Ambion) and reduced fragmentation time (30 seconds instead of 5 minutes) to obtain a longer library fragment size. Libraries were sequenced on an Illumina GAII machine with read lengths of 76bp and 101bp. The resulting sequence reads were aligned to hg19 using TopHat v1.2.0 and Bowtie v0.12.7 [5,6]. Splice junctions detected by TopHat were compared to Ensembl (www.ensembl.org) v61 annotations to identify potentially novel junctions in *HMGCR*. From this analysis two novel splice variants were identified with truncated versions of exons 11 and 18. In each case, multiple unique RNA-seq reads were detected that crossed these novel junctions in libraries from multiple different tissue types, providing direct evidence of their existence. We also observed indirect evidence of these novel splice junctions in the form of paired end reads with insert sizes that were inconsistent with all known transcripts, but consistent with the 93bp and 45bp deletions predicted of the 11b and 18b splice variants, respectively. We noted that many of these reads also contained portions of the normally adjacent exons; for example, sequences containing 11b had portions of exons 10 and 12, while sequences containing 18b matched exons 17 and 19. Lastly, we used BLAT to ensure that these sequences did not align to other areas of the known genome or transcriptome, and verified the novel junctions using RT-PCR with Sanger sequencing

HMGCR transcript stability

HepG2 (n=6) and Huh7 (n=1) were reverse transfected with either one of two siRNAs targeting *SUGP1* or a non-targeting negative control (NTC) for 48 hours using siPORT. 1μg/ml Actinomycin D was added, and aliquots were collected after 0, 2, 4, 6, 24, 30 and 48 hours. Transcript half-life was calculated as previously described [7].

SUGP1 Western Blot analysis

Lymphoblastoid cell lines (LCLs) were lysed using Cell Lytic lysis buffer (Sigma) and loaded on a 4-12% Tris-Glycine Gel (Invitrogen), and proteins were transferred onto a PDVF membrane using the iBLOT gel transfer system (Invitrogen). SUGP1 protein was identified after overnight incubation with an anti-SUGP1 antibody (Sigma) diluted 1:800 and a 1 hour incubation with an anti-goat-HRP antibody diluted 1:20,000. Blots were then stripped and reprobed with an anti- β-actin antibody. Band densities were quantified by Image J and normalized to β-actin (Santa Cruz Biotechnology), adjusted for date of blot and quantile normalized prior to statistical analysis.

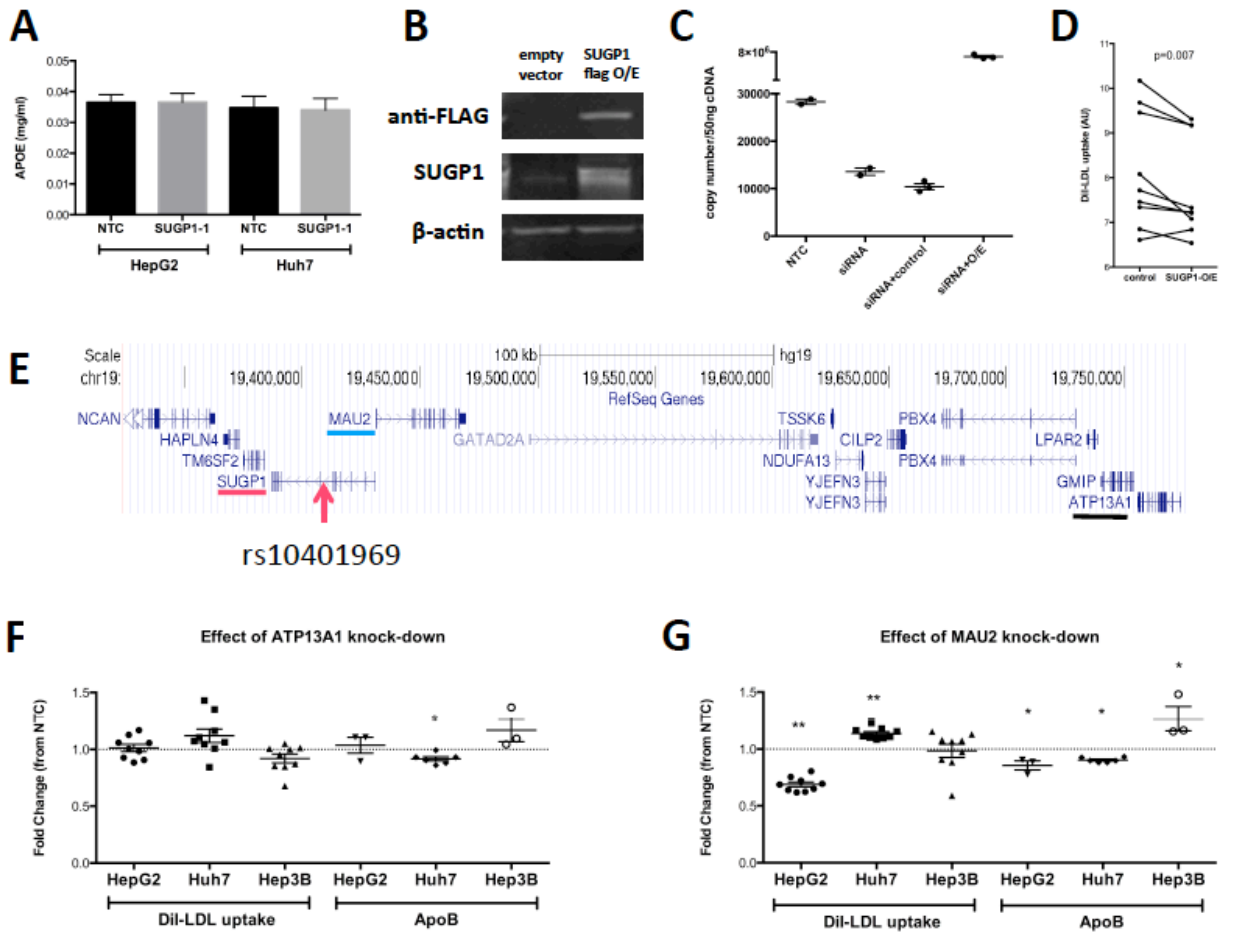

Figure S1. Related to Figure 1. A) HepG2 and Huh7 were reversed transfected with Silence Select siRNAs against *SUGP1* (or a non-targeting control (NTC) siRNA for 48 hr, after which APOE in the cell culture media was quantified by ELISA (n=6 per cell line/condition). **B)** HepG2 cells (n=3) were reversed transfected with the SUGP1-1 siRNA or non-targeting control siRNA for 48 hours, after which SUGP1-1 siRNA treated cells were then transfected with either an SUGP1 overexpression plasmid (siRNA+O/E) or empty vector control (siRNA+control). SUGP1 protein overexpression was validated by western blot. **C)** *SUGP1* transcript levels from the same experiment were quantified by qPCR and normalized to *CLPTM*. **D)** Huh7 cells (n=9) were reversed transfected with either an SUGP1 overexpression plasmid or empty vector control, and after 48hr DiI-LDL uptake was quantified by FACS in 10,000 gated cells. **E)** Genomic region tagged by rs10401969 that is associated with variation in plasma LDLC by GWAS. **F and G)** HepG2, Huh7, and Hep3B were reversed transfected with Silence Select siRNAs against *ATP13A1*, *MAU2,* or a non-targeting control (NTC) siRNA. After 48 hours cells were incubated with DiI-LDL, and magnitude of LDL uptake was quantified as the mean fluorescence per cell in 10,000 gated events was quantified by FACS. *p<0.05, **p<0.001

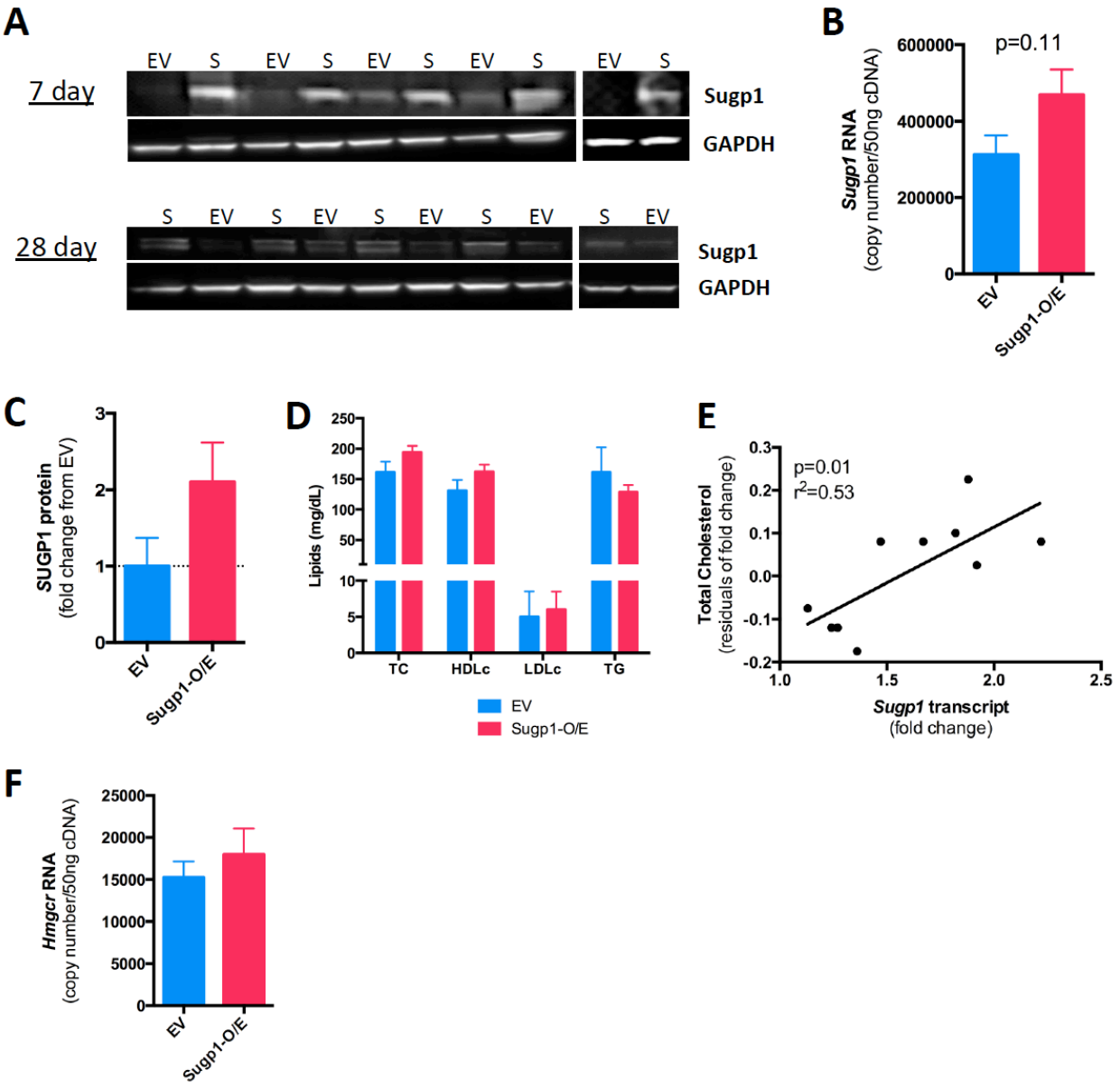

Figure S2. Related to Figure 2. An *Sugp1* overexpression or empty vector control was injected into the tail vein of male chow-fed CD-1 mice (n=10 animals/condition), and animals were sacrificed 7 days and 28 days post-injection. **A)** Hepatic Sugp1 and Gapdh protein levels were quantified by western blot. **B)** Hepatic *Sugp1* transcript **(B)** and protein **(C)** levels, as well as plasma lipids **(D)** were quantified 28 days post-injection. **E)** Correlation of inter-animal variation in magnitude of hepatic Sugp1 overexpression with increase in total cholesterol in the day 7 and 28 animals combined. Fold change in plasma cholesterol levels of Sugp1-O/E versus EV mice was calculated by averaging each Sugp1-O/E animal by the average of the EV injected animals per time point. Fold changes were adjusted by time point, the linear regulation of the residuals used to calculate a correlation with inter-animal variation in overexpression Sugp1 transcript levels. **F)** Hepatic *Hmgcr* transcript levels in control and Sugp1 overexpression mice 7 days post-injection.

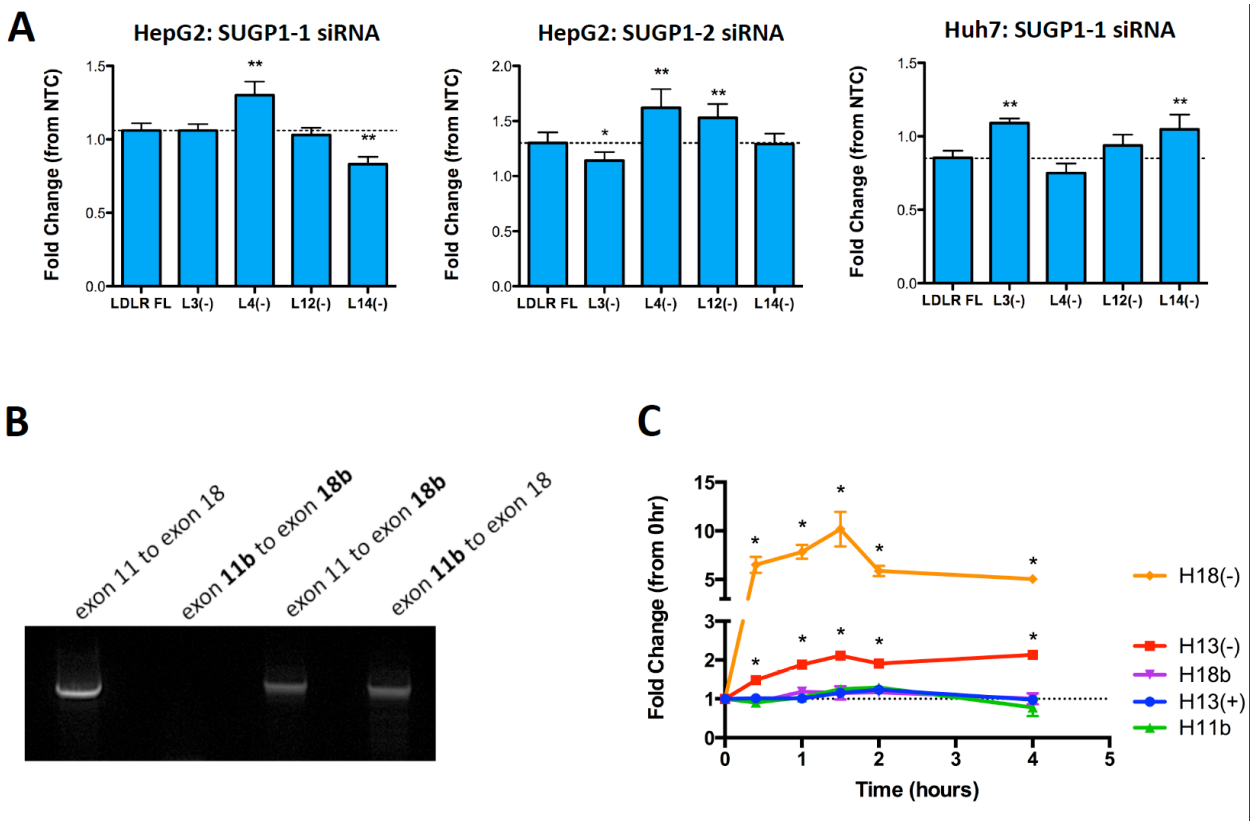

Figure S3. Related to Figure 3. A) HepG2 and Huh7 cells were reverse transfected with either one of two Silence Select siRNAs which target *SUGP1* (SUGP1-1: s33721 or SUGP1-2: s33722) or a non-targeting negative control (NTC). After 48 hours cells were collected and *LDLR* splice variants were quantified by real time qPCR using TaqMan assays (Table S1). All transfections were performed in replicate (n=12) FL=full length (contains exon 4), L3(-) = *LDLR* without exon 3, L4(-) = *LDLR* without exon 4, L12(-) = *LDLR* without exon 12, L14(-) = *LDLR* without exon 14. *p<0.1 **p<0.05 significantly different from FL based on a two-tailed unpaired t-test. **B)** RT-PCR on cDNA derived from a representative LCL (from an rs10401969 TT donor) was performed using forward primers than detect either the canonical *HMGCR* exon 11 (5'-CCCTTAGTGGCTGAAACAGA) or alternative exon 11b (5'-CCGAGAAAGAAAAGTACTGGTGA) with reverse primers that detect either the canonical *HMGCR* exon 18 (5'-CAAACAGGCTTGCTGAGGTAG) or alternative exon 18b (5'-TTGAACACCTAGCATCGTTCC). PCR products were run on a 1% gel with ethidium bromide. **C)** HepG2 cells were treated with 1 µg/ml cyclohexamide, an inhibitor of NMD, and cell pellets were harvested over four hours (n=3 experimental replicates). *HMGCR* splice variants were quantified by qPCR. *p<0.05 significantly different from H13(+) based on a two-tailed unpaired t-test.

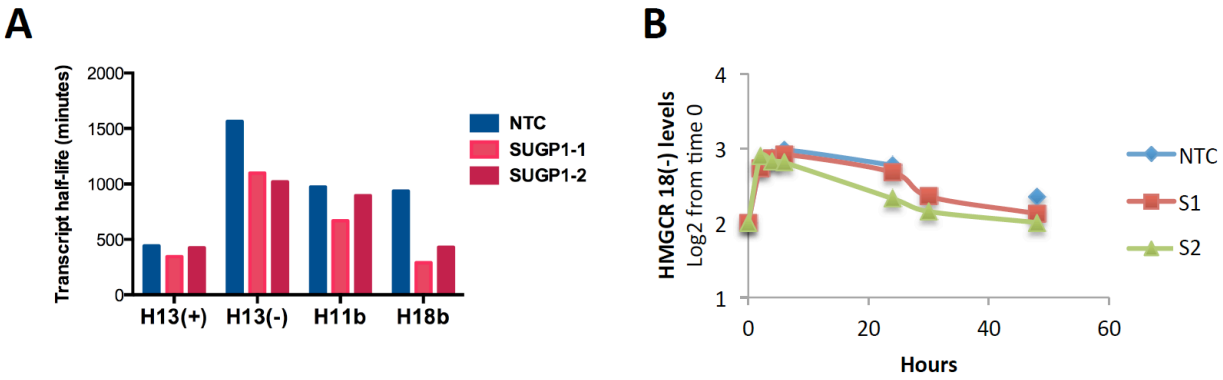

Figure S4. Related to Figure 4. Effect of SUGP1 knockdown on *HMGCR* transcript half-life. A) *HMGCR* transcript half-life quantified in Huh7 cells. Huh7 cells were transfected with one of two different siRNAs for SUGP1 (S1 or S2) or a non-targeting control (NTC). After 48 hours, actinomycin D was added, and *HMGCR* transcripts were quantified by qPCR over 24 hours. **B)** *HMGCR18(-)* levels were quantified in HepG2 (n=6) and Huh7 (n=1) cells over 48 hours after actinomycin D and siRNA treatment as described above. Data represents mean of log2 transformed values of the combined HepG2 and Huh7 datasets.

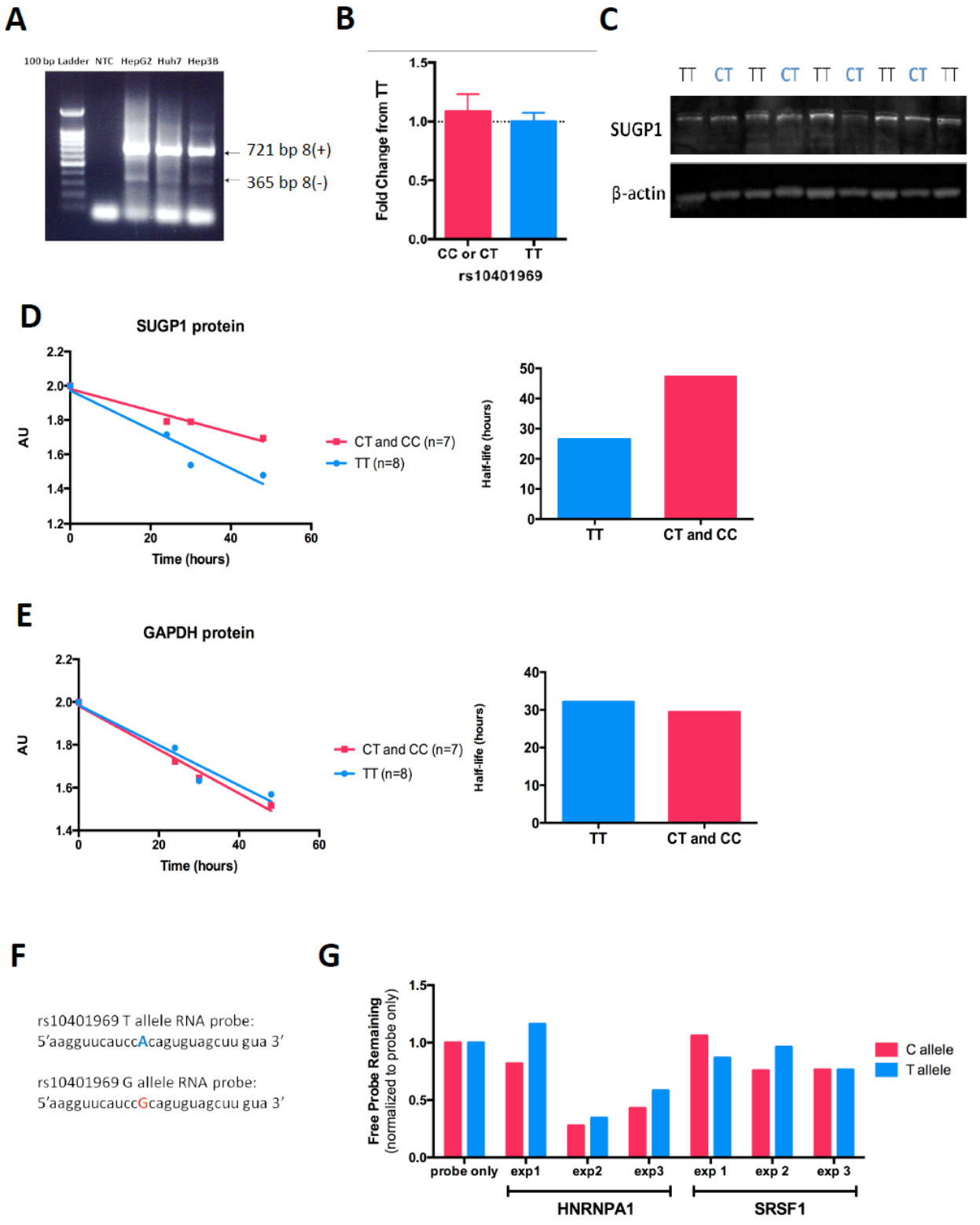

Figure S5. Related to Figure 5. A) Total RNA from Human hepatoma cell lines was isolated and reverse-transcribed into cDNA. PCR primers spanning exon 7 to 10 were used to detect any variants underlying rs10401969 through PCR amplification of cDNA coupled with 0.8% agarose gel electrophoresis. The 721

base-pair (bp) band is the intact PCR product containing exon 8. The 365 bp band was gel extracted and Sanger sequence verified to be a PCR product spanning exon 7 to 10 without exon 8. **B)** *SUGP1* transcripts with and without exon 8 were quantified in human liver samples from rs10401969 C allele carriers (n=15) versus TT homozygotes (n=12) using qPCR. Values shown are the ratio of *SUGP1 8(-)/8(+)* normalized to the average value of the TT homozygotes. **C)** Representative western blot of SUGP1 and B-actin protein from LCLs with and without the rs10401069 "C" allele. **D** and **E)** LCLs were incubated for 24 hours under standardized conditions, after which cycloheximide (1ug/mL) was added, and cell aliquots were harvested after 0, 24, 30 and 48 hours. Cellular protein levels were quantified by Bradford, and SUGP1 and GAPDH protein were assayed by western blot. Image J was used to quantify gel images with values plotted against time. Protein half-life was calculated using the formula, $t1/2\ (h) = \ln 2/(2.303 \times \text{slope})$, and the slope was calculated from the mean of all replicates per genotype group. **F)** The sequences of biotin-labeled rs10401969 RNA probes containing either 'T' or 'C' allele are shown. **G)** Effects of rs10401969 on HNRNPA1 and SRSF1 binding to an *SUGP1* RNA probe were tested by RNA EMSA in three separate experiments. The intensity of the free probe after incubation with either HNRNPA1 or SRSF1 purified protein was quantified by GelQuant software and normalized by the free probe intensity of RNA probe only reactions.

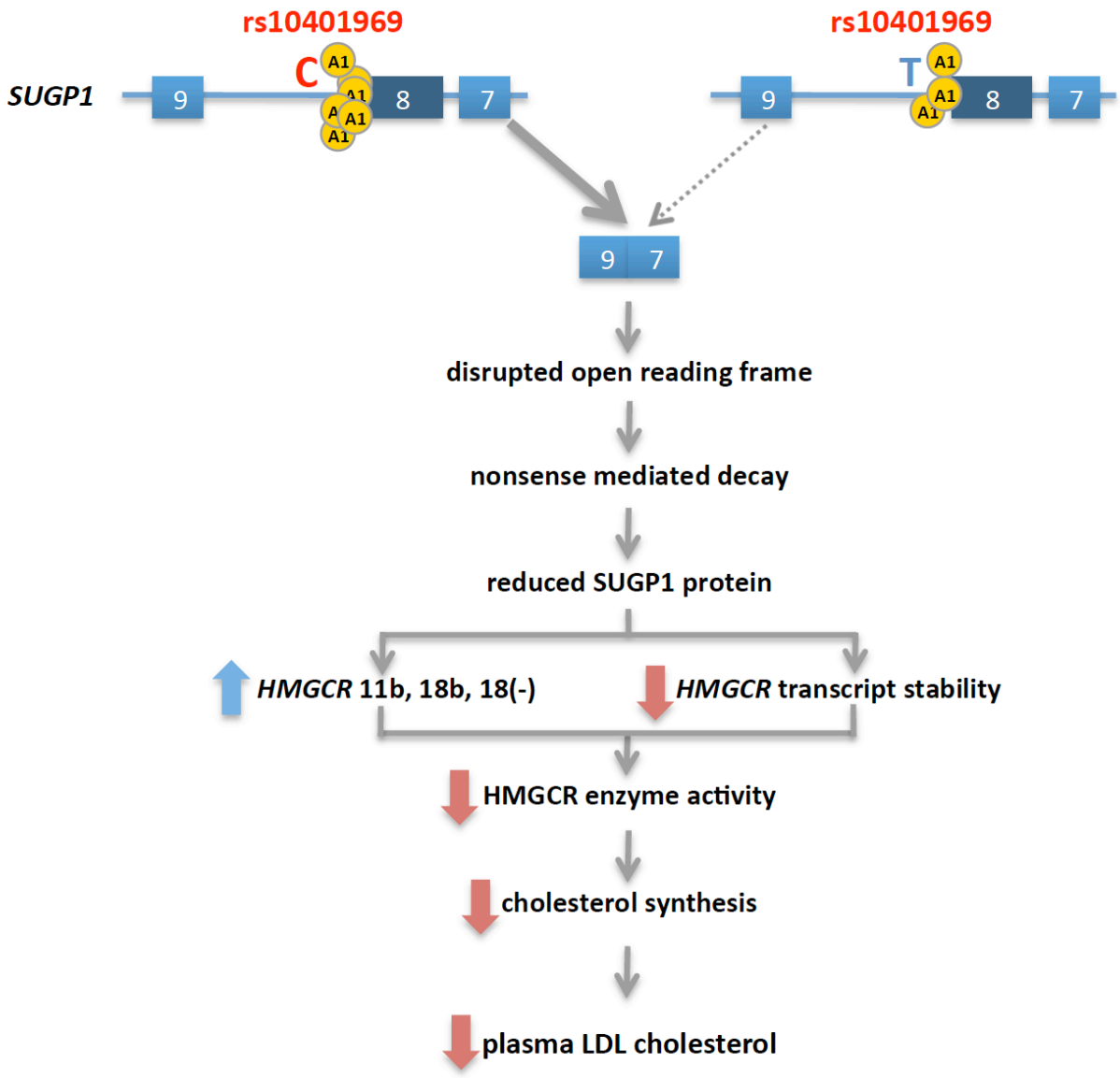

Figure S6. Related to discussion. Overall hypothetical model. Rs10401969 C allele promotes binding of HNRNPA1 (A1 in figure), a splicing suppressor, which stimulates *SUGP1* exon 8 skipping. Loss of exon 8 disrupts the open reading frame, and stimulates nonsense-mediated decay of the resulting transcript, leading to an overall reduction in SUGP1 protein. SUGP1 knock-down stimulates *HMGCR* alternative splicing as well as reduces *HMGCR* transcript stability, which leads to a reduction in HMGCR enzyme activity and thus lower cholesterol synthesis and plasma cholesterol levels.

Table S1. qPCR assays

Transcript	Dye	ABI product number	Forward Primer	Reverse Primer	Probe
ATP13A1	FAM	HS00220755_m1			
CLPTM	FAM	HS00171300			
HMGCR11b	SYBR		AAAGAAAAGTACTGGTGACACAGG	CATCACTAAGGAATTTTGCACCT	
HMGCR13(-)	FAM		CTCCAGTACCTACCTTACAGGGATT	GCTGCTGGCACCTCCA	CAAGCAAGGAGTAATTAT
HMGCR13(+)	FAM	HS01102991			
HMGCR18(-)	SYBR		ATTGCCTGTGGACAGATGCT	CTTTTGACAAGATGTCCTGCTG	
HMGCR18b	SYBR		GGAACGATGCTAGGTGTTCAA	CTTTTGACAAGATGTCCTGCTG	
LDLR full length	FAM		AAGGCTGTCCCCCCAAGA	GAGCTGTTGCACTGGAAGC	CTGCCCGATGGGA
LDLR12(-)	SYBR		GGCATCACCCTAGGACAAAGT	GGGTGAGGTTGTGGAAGAGAA	
LDLR14(-)	SYBR		TCACCCAGCCAAGAGAGG	TGGTTGTGTGCTGTGTCCTT	
LDLR3(-)	SYBR		GAGACGTGCTCTCCCCCAAGA	AACTGCCGAGAGATGCACTT	
LDLR4(-)	FAM		CATTAACGCAGCCAACTTCATC	AAGGCTGTCCTGTGGCCAC	AGTGCTCTGATGGAAAC
MAU2	FAM	HS00384096_m1			
mini-gene H18(-)	SYBR		ATTGCCTGTGGACAGATGCT	AAGGTCACGGGCCTCCAC	
mini-geneH18(+)	SYBR		CTGTTTGCAGATGCTAGGTG	AAGGTCACGGGCCTCCAC	
mini-gene H18b	SYBR		GGAACGATGCTAGGTGTTCA	AAGGTCACGGGCCTCCAC	
mini-gene SUGP1 8(+)	SYBR		GAAGCTCTCTACCTGGTGTG	TCCTCCAGCTTCTGTCGGTA	
mini-gene SUGP1 8(-)	SYBR		GAAGCTCTCTACCTGGTGTG	AGTTGTGCCACTTGTGGGT	
SUGP1 8(+)	SYBR		AAGTTGGCCAGGTTCATAGC	TCCTCCAGCTTCTGTCGGTA	
SUGP1 8(-)	SYBR		GCATTCAGTTCAGGACCTCA	GCTTGTGCTGCATGATCATG	
mouse assays					
Sugp1	FAM	Mm00550992_m1			
Clptm1	FAM	Mm01321458_m1			
mRplp01	SYBER		AACCCTGAAGTGCTCGACAT	GTACCCGATCTGCAGACACA	
mHmgcr13(+)	SYBER		ATTCCTTGGTGATGGGAGC	ATCGGCACCTGGTACTCCTT	

Supplemental References

1 Shen, M. M., Krauss, R. M., Lindgren, F. T. & Forte, T. M. Heterogeneity of serum low density lipoproteins in normal human subjects. *J Lipid Res* **22**, 236-244 (1981).

2 Medina, M. W., Gao, F., Ruan, W., Rotter, J. I. & Krauss, R. M. Alternative splicing of 3-hydroxy-3-methylglutaryl coenzyme A reductase is associated with plasma low-density lipoprotein cholesterol response to simvastatin. *Circulation* **118**, 355-362 (2008).

3 Serra, M., Matabosch, X., Ying, L., Watson, G. & Shackleton, C. Hair and skin sterols in normal mice and those with deficient dehydrosterol reductase (DHCR7), the enzyme associated with Smith-Lemli-Opitz syndrome. *The Journal of steroid biochemistry and molecular biology* **122**, 318-325 (2010).

4 Kim, M. J. *et al.* Functional characterization of liver enhancers that regulate drug-associated transporters. *Clinical pharmacology and therapeutics* **89**, 571-578 (2011).

5 Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111 (2009).

6 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).

7 Medina, M. W. *et al.* RHOA Is a Modulator of the Cholesterol-Lowering Effects of Statin. *PLoS genetics* **8**, e1003058 (2012).