American Heart Association.

Circulation Research

Manuscript Submission and Peer Review System

Disclaimer: The manuscript and its contents are confidential, intended for journal review purposes only, and not to be further disclosed.

URL: https://circres-submit.aha-journals.org/

Manuscript Number: CIRCRES/2020/318141R2

Title: Posttranscriptional regulation of the LDL Receptor in humans by the U2-spliceosome and its interactors

Authors:
Paolo Zanoni (University of Zurich)
Grigorios Panteloglou (University Hospital Zurich)
Alaa Othman (University Hospital Zurich)
Joel Haas (Institut Pasteur de Lille)
Roger Meier (ETH Zurich)
Antoine Rimbert (INSERM, UMR1087, CNRS, UMR 6291, Université de Nantes, CHU de Nantes, l'institut du thorax, CIC-thorax, Nantes, F-44000 France)
Marta Futema (St. George's University of London)
Yara Abou-Khalil (INSERM U1148 - Paris University and Saint Joseph university of Lebanon)
Simon Nørrelykke (ETH Zurich)
Andrzej Rzepiela (ETH Zurich)
Szymon Stoma (ETH Zurich)
Michael Stebler (ETH Zurich)
Freerk Dijk (University Medical Center Groningen)
Melinde Wijers (University of Groningen, University Medical Center Groningen)
Justina Wolters (University of Groningen, University Medical Center Groningen)
Nawar Dalila (Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital)
Nicolette Huijkman (University of Groningen, University Medical Center Groningen)
Marieke Smit (University of Groningen, University Medical Center Groningen)
Antonio Gallo (Groupe Hospitalier Pitié-Salpêtrière, Assistance Publique Hôpitaux de Paris)
Valerie Carreau (Pitie Salpetriere Hospital)
Anne Philippi (Cochin Institute, INSERM U1016)
Jean-Pierre Rabès (APHP-Hôpital Ambroise Paré)
Catherine Boileau (INSERM U1148)
Michele Visentin (University of Zurich)
Luisa Vonghia (University Hospital Antwerpen)

Circulation Research Online Submission: https://circres-submit.aha-journals.org
Circulation Research Homepage: https://www.ahajournals.org/journal/res
Circulation Research Editorial Office
Email: CircRes@circresearch.org

Jonas Weyler (Antwerp University Hospital)
Sven Francque (Antwerp University Hospital)
An Verrijken (Antwerp University Hospital/University of Antwerp)
Ann Verhaegen (University of Antwerp)
Luc Van Gaal (University Hospital Antwerp)
Adriaan van der Graaf (University Medical Center Groningen)
Belle van Rosmalen (Academic Medical Centre - University of Amsterdam)
Jérôme Robert (University of British Columbia)
Srividya Velagapudi (University of Zurich)
Mustafa Yalcinkaya (Columbia University)
Michaela Keel (University of Zurich)
Silvija Radosavljevic (University Hospital of Zurich)
Andreas Geier (University Hospital WÃ1/4rzburg)
Anne Tybjærg-Hansen (Rigshospitalet, Copenhagen University Hospital)
Mathilde Varret (LVTS INSERM U1148)
Lucia Rohrer (University Hospital Zürich, Zurich, Switzerland)
Steve Humphries (British Heart Foundation Laboratories, University College of London)
Bart Staels (Inserm U1011)
Bart van de Sluis (University Medical Center Groningen)
Jan Albert Kuivenhoven (University of Groningen, University Medical Center Groningen)
Arnold Von Eckardstein (University Hospital Zurich)

For Circulation Research Peer Review
distribute. Destroy after use

Posttranscriptional regulation of the LDL Receptor in humans by the U2-spliceosome and its interactors

Paolo Zanoni[1,2,*], Grigorios Panteloglou[1,2,*], Alaa Othman[3], Joel T. Haas[4], Roger Meier[5], Antoine Rimbert[6], Marta Futema[7], Yara Abou Khalil[8,9], Simon F. Norrelykke[5], Andrzej J. Rzepiela[5], Szymon Stoma[5], Michael Stebler[5], Freerk van Dijk[10], Melinde Wijers[6], Justina C. Wolters[6], Nawar Dalila[11], Nicolette C. A. Huijkman[6], Marieke Smit[6], Antonio Gallo[12], Valérie Carreau[12], Anne Philippi[13], Jean-Pierre Rabès[8,14,15], Catherine Boileau[8,16], Michele Visentin[17], Luisa Vonghia[18,19], Jonas Weyler[18,19], Sven Francque[18,19], An Verrijken[19,20], Ann Verhaegen[19,20], Luc Van Gaal[19,20], Adriaan van der Graaf[10], Belle V. van Rosmalen[21], Jerome Robert[1], Srividya Velagapudi[1,2], Mustafa Yalcinkaya[1,2], Michaela Keel[1,2], Silvija Radosavljevic[1,2], Andreas Geier[22], Anne Tybjaerg-Hansen[11], Mathilde Varret[8], Lucia Rohrer[1,2], Steve E. Humphries[23], Bart Staels[4], Bart van de Sluis[6], Jan Albert Kuivenhoven[6], and Arnold von Eckardstein[#,1,2]

* These authors contributed equally to this work

Author affiliations:

1. Institute for Clinical Chemistry, University and University Hospital Zurich, Zurich, 8091 Switzerland.
2. Center for Integrative Human Physiology, University of Zurich, Zurich, 8057 Switzerland.
3. Institute of Molecular Systems Biology, ETH Zurich, 8093, Zurich, Switzerland.
4. University of Lille, Inserm, CHU Lille, Institut Pasteur de Lille, U1011- EGID, 59000 Lille, France.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

5. Scientific center for optical and electron microscopy (ScopeM), ETH Zurich, 8093, Zurich, Switzerland.

6. Department of Pediatrics, Section Molecular Genetics, University of Groningen, University Medical Center Groningen, 9713 AV Groningen, the Netherlands.

7. Cardiology Research Centre, Molecular and Clinical Sciences Research Institute, St George's, University of London, London, UK.

8. LVTS-INSERM UMRS 1148 and University of Paris, CHU Xavier Bichat, 75018, Paris, France.

9. Laboratory of Biochemistry and Molecular Therapeutics (LBTM), Faculty of Pharmacy and Pôle technologie Santé (PTS), Saint-Joseph University, 1104 2020, Beirut, Lebanon.

10. University of Groningen, University Medical Center Groningen, Department of Genetics, 9713AV, Groningen, The Netherlands

11. Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital, Faculty of Health and Medical Sciences, University of Copenhagen, 2100 Copenhagen, Denmark.

12. AP-HP, Endocrinology and Metabolism Department, Human Research Nutrition Center, Pitié-Salpêtrière Hospital, 75013 Paris, France.

13. Université de Paris, Faculté de Médecine Paris-Diderot, UMR-S958, Paris, France. Current address: Université de Paris, Institut Cochin, INSERM U1016, CNRS UMR-8104, Paris, France.

14. AP-HP, Université Paris-Saclay, 75004, Paris, France.

15. UFR Simone Veil des Sciences de la Santé, UVSQ, 78180 Montigny-Le-Bretonneux, France.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

16. AP-HP, Genetics Department, CHU Xavier Bichat, Université de Paris,75018 Paris, France.
17. Department of Clinical Pharmacology and Toxicology, University Hospital Zurich, 8091, Zürich, Switzerland
18. Department of Gastroenterology and Hepatology, Antwerp University Hospital, 2650 Antwerp, Belgium.
19. Laboratory of Experimental Medicine and Paediatrics, Faculty of Medicine, University of Antwerp, 2610 Antwerp, Belgium.
20. Department of Endocrinology, Diabetology and Metabolism, Antwerp University Hospital, 2650 Edegem, Belgium.
21. Department of Surgery, Academic Medical Center, University of Amsterdam, 1105 AZ, Amsterdam, The Netherlands
22. Division of Hepatology, Department of Medicine II, University Hospital Würzburg, 97080 Würzburg, Germany.
23. Cardiovascular Genetics, Institute of Cardiovascular Science, University College London, WC1E 6JJ London, United Kingdom.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Actual affiliations at the moment of submission

- P.Z. is affiliated to the Institute of Medical Genetics, University of Zürich, 8952 Schlieren, Switzerland
- A.R. is affiliated to Inserm UMR 1087 / CNRS UMR 6291 IRS-UN, 44007, Nantes, France
- M.Y. is affiliated to the Division of Molecular Medicine, Department of Medicine, Columbia University, New York, 10032, NY, USA
- S.V. is affiliated to the Center for Molecular Cardiology, University of Zurich, 8952 Schlieren, Switzerland

Short title: U2 spliceosome regulates LDL receptor

Contact information:

Arnold von Eckardstein; Institute for Clinical Chemistry, University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland; +41 44 255 22 60; arnold.voneckardstein@usz.ch

Word count: 7858 excluding title page and abstract

Keywords: LDL-cholesterol, LDL receptor, spliceosome, intron retention, RNA interference, genome wide screening

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Abstract

Rationale: The low-density lipoprotein receptor (LDLR) in the liver is the major determinant of LDL-cholesterol levels in human plasma. The discovery of genes that regulate the activity of LDLR helps to identify pathomechanisms of hypercholesterolemia and novel therapeutic targets against atherosclerotic cardiovascular disease.

Methods and results: In a genome-wide RNA interference screen, the knock-down of 54 genes led to a significant inhibition of LDL uptake. Fifteen of these genes encode for proteins involved in splicing, especially components or interactors of the U2-spliceosome. Knocking down any one of 11 out of 15 genes resulted in the selective retention of intron 3 of *LDLR*. The transcript is translated into an LDLR fragment, which lacks 88% of the full length LDLR and is detectable neither in non-transfected cells nor in human plasma. Surprisingly, the intron 3 retention transcript is expressed in considerable amounts in human liver and in blood cells. Its hepatic expression is increased in non-alcoholic fatty liver disease as well as after bariatric surgery. Its expression in blood cells correlates with LDL-cholesterol and age. Single nucleotide polymorphisms and three rare variants of one spliceosome gene, *RBM25*, are associated with LDL-cholesterol in the population and familial hypercholesterolemia, respectively. Compared to overexpression of wild type RBM25, overexpression of the three rare RBM25 mutants in Huh-7 cells led to lower LDL uptake.

Conclusions: We identified a novel mechanism of post-transcriptional regulation of LDLR activity in humans and associations of genetic variants of RBM25 with LDL-cholesterol levels.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Introduction

Hypercholesterolemia is a causal and treatable risk factor of atherosclerotic cardiovascular diseases (ASCVD)[1]. The most important determinant of LDL-cholesterol (LDL-C) levels in plasma is the hepatic removal of circulating LDL by binding to the LDL receptor (LDLR) for subsequent endocytosis and degradation[2]. The expression of LDLR is tightly regulated by transcription factors, proteasomal and lysosomal degradation, endosomal recycling, and cleavage at the cell surface[1,2]. The unravelling of this complex regulation led to the development of drugs that effectively lower plasma levels of cholesterol and, as the consequence, risk of ASCVD[1].

To identify novel regulators of LDL uptake into the liver, we performed an image-based genome-wide RNA interference (RNAi) screen in Huh-7 human hepatocarcinoma cells. Fifteen out of 54 genes significantly reducing LDL uptake upon knockdown encode for proteins involved in pre-mRNA splicing. The majority of them are either core components or interactors of the U2-spliceosome[3]. By functionally validating this finding *in vitro* as well as in human tissues, we provide evidence that a functional U2 spliceosome is needed for the expression of full length LDLR and, hence, determining LDLR activity in humans.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Methods

Data Availability

The authors declare that all data and methods supporting the findings of this study are available in the Data Supplement or from the corresponding authors on reasonable request.

A detailed description of materials and methods is provided in the text and Major Resources Table of the Online Supplement

Results

The U2-spliceosome and its interactors are rate-limiting for LDL endocytosis

For the genome wide RNAi screen of genes limiting uptake of LDL or HDL, Huh-7 human hepatocarcinoma cells were reverse-transfected using three different siRNA oligonucleotides against each of the 21,584 different human genes. To control efficacy and specificity of transfection, each plate contained wells with cells transfected with siRNAs against *PLK1* whose knockdown results in cell death, and *LDLR,* respectively. Based on results of time and dose finding experiments, the cells were exposed 72 hours post transfection to 33 μg/ml each of Atto594-labelled LDL and Atto655-HDL for 4 hours. As background controls, wells with cells transfected with a non-targeting siRNA, were incubated in the absence of fluorescent lipoproteins. After washing, fixation, and staining of the nuclei with Hoechst 33258, the plates were imaged at 4x and 20x with two twin wide-field automated microscopes. Nuclei, the relative cytoplasm, and fluorescent LDL-containing vesicles were identified through automated image analysis (Figure 1A). Transfection efficiency was very high (Supplemental Figure Ia). Analysis and validation of HDL image data will be subject of a separate report. For the uptake of fluorescent LDL, the five best performing assay features (foci count per cell, foci mean intensity, cytoplasm granularity 1 and 2, cytoplasm median intensity) showed a

For Circulation Research Peer Review. Do not distribute. Destroy after use.

high degree of correlation. Therefore and because of the widest dynamic range based on Z'-factor values from control wells, we identified gene hits by the Redundant siRNA Activity (RSA) analysis of data from the median cytoplasm intensity feature. Z'-factor values for median cytoplasm intensity in each assay plate for both the background (median 0.00, interquartile range [IQR] -023 to 0.20) and positive control (median -0.56, IQR -0.99 to -0.20) clustered mostly around the 0-line, indicating a suboptimal but analytically exploitable signal-to-noise ratio (Supplemental Figure Ib). Dimensionality reduction of main assay features did not significantly alter the outcome (Supplemental Figures Ic and Id). At an RSA p-value cutoff of $p < 10^{-3}$, interference with 54 and 37 genes decreased and increased LDL uptake, respectively (Table 1, Supplemental Table I). By contrast to the findings of a previous genome wide CRISPR-based screening in Huh7 cells[4], our list does not include LDLR or its modulators such as SCAP, MBTPS1, or IDOL/MYLIP except AP2M1, which is an essential contributor to clathrin mediated endocytosis (table 1). Gene Ontology (GO) enrichment analysis showed significant clustering for genes whose loss of function decreased LDL uptake (Supplemental Table II). Functional clustering of these genes with the STRING tool revealed four major groups: the ribosome (N = 7), the proteasome (N = 8), the spliceosome (N = 15), and vesicular transport (N = 5) (Figure 1B). Out of the 15 spliceosome genes, six encode for core components of the U2 spliceosome, namely *SF3A1*, *SF3A2*, *SF3B1*, *SF3B2*, *SF3B5* and *SF3B6*. Other proteins, interact with the U2-spliceosome either directly (AQR, ISY1 and RBM25) or indirectly (RBM22)[3].

To confirm the role of the U2 spliceosome in LDL endocytosis *in vitro*, we performed ^{125}I-LDL cell association assays in Huh-7 and HepG2 cells. *SF3B4* was also included in these experiments as it is part of the U2 spliceosome and barely missed the RSA p-value cut-off ($p = 1.4*10^{-3}$). Knockdown was achieved using 4 pooled siRNA molecules against each hit gene acquired from vendors other than that of the siRNAscreening library, namely Dharmacon or Sigma instead of Ambion. (see Major Resource Table). For RBM25 we replaced

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Dharmacon's siRNAs with those from Sigma because of their presumable off-target effects on LDLR protein expression (Supplemental Figures IIId and IIIe). Knockdown of each of these genes significantly decreased the specific cell association of ^{125}I-LDL with both Huh-7 and HepG2 cells (Figure 1C, Supplemental Figure IIb). The association of ^{125}I-LDL was equally decreased by knockdown of *SF3B1* (-45±5%), *SF3A2* (-47±6%), *AQR* (-45±6%), and *LDLR* (-43±8%) (Figure 1C). RNA interference with RBM25 reduced the specific cellular association of ^{125}I-LDL and fluorescent LDL by 27% ±8% and 52%±5%, respectively (Figure 1D and Supplemental Figure IIIf). Of note, the specific cell association of ^{125}I-HDL was unaltered or even increased upon knockdowns of *AQR* and *SF3A1* in either Huh-7 or HepG2 cells (Supplemental Figures IIc and IId).

Loss of U2-spliceosome genes and their interactors causes selective retention of *LDLR* intron 3 (IVS3)

To unravel the mechanism through which the U2-spliceosome and its interactors regulate LDL endocytosis, we applied RNA sequencing to Huh-7 cells, which were transfected with either siRNAs against eleven U2-spliceosome genes or a non-targeting control siRNA. Sequences can be accessed by codes PRJEB46899 and PRJEB46898 in the data bank of the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/support) 72h after transfection, we measured both expression at the gene level and alternative exon usage in polyA-selected transcripts. Knockdown of all eleven genes except *RBM25* induced a marked increase in the retention of intron 3 of *LDLR* in mature transcripts without altering the expression of the *LDLR* full length transcript (Figure 2A, Supplemental Figure IV). This effect was confirmed in Huh-7 cells by RT-PCR upon knockdown of *AQR, SF3B1, or RBM25* by employing a primer set that was previously used to study the effects of the rare LDLR c.313+1, G>A intronic variant, which leads to *LDLR* loss of function by constitutively

For Circulation Research Peer Review. Do not distribute. Destroy after use.

promoting intron 3 retention[5] (Supplemental Figure Va). By contrast to the RNA sequencing (Supplemental Figure IV), RT-PCR unravelled increased expression of the LDLR IVS3 retention transcript upon knock-down of RBM25, albeit not as much as with knock-down of SF3B1 and AQR (Supplemental Figures Vb and Vc).

Among all intronic or exonic sequences in the transcriptome, the expression of the intron 3 retaining *LDLR* transcript was altered most strongly. Upon knockdown of *SF3B1, AQR, or SF3A2*, the retained intronic sequence of *LDLR* ranked at the top of each respective dataset when the exon-level expression data was plotted against each other (Figure 2B). The degree of intron 3 retention upon knocking down U2-spliceosome genes was significantly correlated with the decrease in ^{125}I-LDL cell association, ($r = -0.73$, $p = 1.4*10^{-2}$, Figure 2C).

To investigate reasons for intron 3 retention in *LDLR*, we transfected HEK293T cells with two minigenes containing different portions of the *LDLR* genomic sequence flanked by two artificial exons (Figure 3A). The first minigene (MG_1) encoding only for exon 3 of *LDLR* and the adjacent intronic regions cloned between two artificial exons (SD6 and SA2), displayed very low if any RNA sequencing reads mapping to the first ~130bp of intron 3. On the contrary, upon expression of the whole genomic sequence between the 3'-end of intron 2 and the 5'-end of intron 4 of *LDLR* (MG_2) an increased number of reads mapped to the first section of intron 3. This indicates incomplete splicing of intron 3 when the physiological exon 4 acceptor site and the branch point site (BPS) were present in the larger minigene MG_2 (Figure 3B). The acceptor splice site of exon 4 of *LDLR* hence appears to be poorly defined. The bioinformatic analysis of the portion of intron 3 neighbouring exon 4 by the U2 branchpoint prediction algorithm SVM-BP-finder (http://regulatorygenomics.upf.edu/Software/SVM_BP/)[6] identified one plausible U2-spliceosome dependent BPS located 30 bp upstream of the acceptor site (Supplemental Table III). The *gtgat* pentamer in the centre of the *cggtgatgg* branchpoint sequence was associated with very low U2 binding energy and occurs at low frequency in the branchpoint database[6].

For Circulation Research Peer Review. Do not distribute. Destroy after use.

We discarded another predicted branchpoint 124bp upstream of the acceptor site as the subsequent AG-exclusion zone does not reach up to the acceptor. Contrary to exon 4 of human *LDLR*, exon 4 of murine *Ldlr* contains a strong and frequently recurring branchpoint 33 bp upstream of the acceptor site (Figure 3C). This finding is in accordance with intron 3 of *Ldlr* being barely detectable at the RNA level by RT-PCR in mouse liver (data not shown). Taken together, these data suggest that the BPS of intron 3 in human *LDLR* is poorly defined and therefore very sensitive to alternative splicing.

Selective intron 3 retention limits LDLR cell surface abundance

The transcript with intron 3 retention encodes for a prematurely truncated proteoform of LDLR because the 5'-end of intron 3 encodes for 12 novel amino acids followed by a stop codon. Including the signal peptide, this theoretical 116 amino acid residues long and 12.7 kDa large 'LDLRret fragment' encompasses the complete first and large part of the second class A domains (labelled as L1 and L2 in Figure 4A[7]) but lacks all other domains, including the transmembrane portion of LDLR. Western blots probed with an antibody against the C-terminus of LDLR revealed 60±30% and 61±13% lower LDLR protein levels upon knockdown of AQR and *SF3B1*, respectively *(*Figures 4B and 4C). A similar decrease in LDLR protein was seen upon knock-down of RBM25 with siRNAs from Sigma (-68%±10%), whereas the knock-down of RBM25 with the siRNA of Dharmacon led to an increase in LDLR protein (+122%±109%), presumably due to off target effects (Supplemental Figures IIId and IIIe). Flow cytometry experiments on alive Huh-7 cells after *SF3B1* and *AQR* knockdown showed a -87%±1% and –61%±4%, respectively, lower cell surface abundance of LDLR (Figure 4D). The knock-down of RBM25 with siRNAs from Sigma and Dharmacon decreased the cell surface abundance of LDLR by 53%±6% and 21%±5% , respectively, as compared to scrambled siRNAs from the same manufacturers (Supplemental Figure IIIg).

For circulation Research Peer Review. Do not distribute. Destroy after use.

To investigate whether cells produce and secrete the LDLRret fragment, we overexpressed a C-terminally HA-tagged version of the LDLRret fragment in HEK293T cells. 48 hours after transfection, the HA-tagged LDLRret fragment was detectable in the cell lysates (Figure 4E) as well as in undiluted cell culture media (Figure 4F). The proteasomal inhibitor MG-132 decreased cellular LDLRret protein levels (Figure 4E) suggesting that the LDLRret fragment is not catabolized through the proteasome. We also overexpressed an untagged version of the LDLRret fragment in HEK293T cells. Targeted mass spectrometry recorded a peptide, which is present in both the full-length protein and in LDLRret, over its basal endogenous level in HEK293T cell lysates (Supplemental Figure VI) but not in human plasma (data not shown).

A large proportion of *LDLR* transcripts in human liver and blood cells retains intron 3

To investigate its physiological or pathological relevance, we quantified *LDLR* intron 3 retention in liver biopsies as well as in peripheral blood cells by three different methods, and explored associations with non alcoholic fatty liver disease (NAFLD), demographic measures, lipid traits, and therapeutic interventions.

RT-PCR of mRNAs of liver tissue from 17 patients with benign liver tumours and nine patients with suspected NAFLD, found the *LDLR* intron 3 retention transcript expressed at considerable and interindividually variable amounts (Figure 5A). Taking the sum of the full length and intron 3 retention transcripts of *LDLR* as the reference, 43 % (range 23% to 85%) of the transcripts retained intron 3 (Figure 5A).

The bioinformatics analysis of RNA sequencing data on liver samples of 13 healthy non-obese subjects,, 12 obese subjects without NAFLD, 15 patients with NAFLD, and 15 patients with non-alcoholic steatophepatitis (NASH) (Gene Expression Omnibus, accession number GSE126848)[8] found 14 different LDLR transcripts (Supplemental Figure VII). Four transcripts showed the largest interindividual variation, namely LDLR-201 and LDLR-208,

For Circulation Research Peer Review. Do not distribute. Destroy after use.

encoding full length LDLR as well as LDLR-206, which corresponds to the retained intron 3 transcript, and the likewise futile LDR-214. Interestingly, the median concentration of LDLR-206 was substantially higher in patients with NAFLD or NASH than in normal weight or obese subjects without NAFLD. The median percentages of LDLR-206 reads relative to total reads from all transcripts of *LDLR* gene increased significantly from 1.8% (range 0.7% to 4.2%) and 1.7% (0.4% to 3.7%) in normal weight and obese subjects without NAFLD, respectively, to 5.8% (1.1 to 26.7%) and 5.0% (0.9% to 29.0%) in patients with NAFLD and NASH, respectively (figure 5B). Of the two most abundant full length encoding LDLR transcripts, LDLR-208 decreased significantly (figure 5C) while the expression of LDLR-201 did not change (supplemental figure VII).

We also investigated the expression of LDLR transcripts in liver biopsies of 155 obese non-diabetic subjects[9] by using Affymetrix Human Gene 2.0 ST arrays (see Supplemental Table IV for clinical and biochemical characteristics). The signal intensities from a probe located in intron 3 of *LDLR* were significantly higher than the other intronic *LDLR* probes located in introns 2, 4 and 15 and comparable to probes located in coding exons (Figure 5D). The percent intensities of the IVS3 probe relative to the sum of all LDLR probes ranged from 7.5% to 82%. Intron 3 retention correlated significantly only with *SF3B1* ($r = 0.26$, $p = 1.5*10^{-2}$) , while no U2-spliceosome gene showed any significant correlation with overall *LDLR* expression (Supplemental Table V). Relative intensities of neither the intron 3 probe nor any other of the 24 *LDLR* probes showed significant correlations with plasma levels of total, HDL- or LDL-cholesterol (Supplemental Figures VIIIa, VIIIb and VIIIc, Supplemental Table VI). Correlations with histological NAFLD stages were inverse by trend but significant only for ballooning, a marker of apoptosis or degeneration of hepatocytes (Spearman $r = 0.15$ $p = 6.0*10^{-2}$) (Supplemental Figure VIIId). Intron 3 relative probe intensity did not correlate with BMI (Supplemental Figure VIIIe) .However, in a subgroup of 21 patients who underwent a second liver biopsy after bariatric surgery (median follow-up time = 13 months, IQR = [12,

For Circulation Research Peer Review. Do not distribute. Destroy after use.

15]),the proportion of the intron 3 retention transcript relative to the full length *LDLR* transcript increased significantly after surgery ($p = 4.9*10^{-3}$) (Supplemental Figure VIIIf; supplemental table IV). This increase was even more pronounced in eleven patients with NASH at baseline but no NASH at follow-up ($p = 1.8*10^{-2}$, Supplemental Figure VIIIg)..

Finally, we analyzed the RNA-sequencing data in whole blood samples from 2,462 subjects of the Dutch BIOS consortium[10]. The *LDLR* ENST00000557958 transcript, predicted to retain intron 3, was detectable in all subjects and represented 21%±7% of the total *LDLR* transcripts. The ENST00000557958 transcript levels significantly correlated with age ($r = 0.25$, $p = 2.3*10^{-36}$, Figure 6A) and less strongly with LDL-C ($r = 0.089$, $p = 9.8*10^{-6}$, Figure 6B). The latter correlation lost its statistical significance after adjusting for age, suggesting age itself as the main driver of the association between ENST00000557958 levels and LDL-C. ENST00000252444, the only transcript encoding for full length *LDLR* and expressed in blood cells in all subjects in this dataset , was also positively correlated with age ($r = 0.19$, $p = 2.2*10^{-20}$, Figure 6C) but not with LDL-C ($r = -0.033$, $p = 1.0*10^{-1}$, Figure 6D). Correlation of neither transcript with BMI was statistically significant.

Single Nucleotide Polymorphisms in RBM25 are associated with lower LDL-Cholesterol

The analysis of whole exome sequencing (WES) data of 40,468 UK Biobank subjects[11] did not unravel any significant association between our spliceosome hit genes and LDL-C or any other clinical lipid trait (Supplemental Table VII). However, constraints data from the gnomAD database indicate a strong intolerance to functional genetic variation for our U2-spliceosome genes, with a probability of intolerance to loss of function (pLI)[12] of 0.91±0.17 (mean± SD) (Supplemental Table VIII). The analysis of SNPs of 11 U2-spliceosome hit genes in 361,194 participants of UK Biobank found 24 SNPs of *RBM25*

For Circulation Research Peer Review. Do not distribute. Destroy after use.

significantly associated with lower levels of LDL-C (Figure 7A) and apoB (Supplemental Figure IXa).

In Europeans, four SNPs in introns or downstream of the *RBM25* coding sequence including the lead SNP rs17570658 and two upstream SNPs are in almost complete LD (Supplemental Figure IXb). , With $R^2 > 0.8$ no other SNP of RBM25 is in strong LD (i.e.). A meta-analysis of eight studies with 455'537 samples (https://cvd.hugeamp.org/variant.html?variant=rs17570658) and data of the Copenhagen City Heart and General Population Studies[13] according to METAL[14] showed the association of rs17570658 with LDL-C (Z Score = -41.81, P = $2.9*10^{-5}$, supplemental table IX).

RBM25 is widely expressed in many tissues, but expression is relatively low in liver (GTeX https://gtexportal.org/home/, data not shown). rs17570658 shows strong association with *RBM25* expression in 15 different tissues including skeletal muscle and arteries (Figure 7B) as well as adipose and mammary tissue, lung, oesophagus, kidney, and skin. Carriers of the rare allele have higher mean RBM25 mRNA concentration, which is compatible with higher LDLR activity and lower LDL-C levels.

Impaired LDL uptake by cells expressing rare RBM25 mutants found in patients with familial hypercholesterolemia

In the UK10K study, *RBM25* was also among the genes identified to harbour an excess of rare novel variants in 71 patients with familial hypercholesterolemia who are negative for mutations in *LDLR*, *APOB* and *PCSK9*, the known FH-causing genes[15]. We re-analyzed the burden of variants in the *RBM25* gene, using previously published WES data from 71 FH patients negative for mutations in *LDLR*, *APOB* and *PCSK9*, and 56,352 European data provided by the gnomAD study[12]. Missense, splice site, frameshift, and stop-gained variants identified by WES in both FH cases and gnomAD were filtered to select those with MAF<$1.0*10^{-4}$. After filtering, three *RBM25* variants were found in the FH cohort and

For Circulation Research Peer Review. Do not distribute. Destroy after use.

163 in the gnomAD Europeans cohort. (Supplemental Table X). Two variants, p.I152F (c.454A>T) and p.A455D (c.1364C>A), were not found in any publicly available sequencing database and hence appear unique to the FH cohort. The third variant, p.L17P (c.50T>C) (rs1167173761), was found in one European individual in the gnomAD cohort (MAF=9 $*10^{-6}$, allele count = 1/251402). The comparison of variant numbers in FH cases vs. gnomAD using a binomial test demonstrated the enrichment of rare variants in *RBM25* in the FH cohort ($p = 1.0*10^{-3}$). Within the UK10K cohort, no other U2-spliceosome gene was found to carry a rare presumable LOF mutation.

We investigated the functional consequences of overexpressing the three FH-associated RBM25 mutants in Huh7 cells. Overexpression of all constructs RBM25 was confirmed by qPCR (Supplemental Figures Xa and XIa) and - for wild type RBM25 - Western blotting (Supplemental Figure Xb). The overexpression of neither wild type RBM25 nor any RBM25 mutant in Huh7 cells caused significant changes in the expression of full length or IVS3 retention transcripts of LDLR (Supplemental Figures Xc, Xd, XIb, and XIc). Compared to empty vector, overexpression of wild type RBM25 in Huh7 cells changed neither the cell surface abundance of LDLR nor LDL uptake significantly (Supplemental Figures Xe and Xf). Comparisons with cells overexpressing wild type RBM25 revealed minor decreases of LDLR cell surface levels but more pronounced or even significant decreases of Atto655-LDL-uptake of cells overexpressing the RBM25 mutants p.L17P (-15%+16%), I152F (-23%+12%), or p.A455D (-28%+12%, $p = 2.6*10^{-2}$) (Supplemental Figure XIe).

Discussion

Through genome-wide siRNA screening, we discovered that the U2-spliceosome as well as some interacting proteins, control LDLR levels and LDL uptake in liver cells by modulating the selective retention of intron 3 of *LDLR*. The intron 3 retaining *LDLR* transcript encodes a truncated and most probably non-functional receptor. In several cohorts of healthy

For Circulation Research Peer Review. Do not distribute. Destroy after use.

individuals and patients, we observed considerable interindividual variation of *LDLR's* IVS3 retention in liver as well as in peripheral blood cells. Finally, we obtained initial evidence that rare genetic variants as well as SNPs associated with its expression levels in the U2-spliceosome-associated gene *RBM25* are related to LDL-C levels in humans. Taken together, our findings suggest intron 3 retention of *LDLR* as a novel mechanism regulating LDLR activity and thereby plasma levels of LDL-C.

A previous siRNA screen also found U2-spliceosome genes to limit the uptake of LDL into EA.hy926 cells but the authors excluded them from further analysis and validation[16]. Basic cellular functionality of spliceosome genes may be the reason why U2- spliceosome genes were not found by a previous CRISPR-based screen as limiting factors for LDL uptake into Huh-7 cells[4]. As these authors discussed, CRISPR-based screens may overlook genes that are essential or confer a fitness advantage in culture, since guide RNAs targeting those genes will be progressively depleted from the pooled population[4].

As a preliminary mechanistic explanation, our minigene data as well as our *in silico* predictions suggest that the BPS in intron 3 of human LDLR is poorly defined and thereby highly sensitive to alterations in the activity of U2 splice factors. In this regard it is noteworthy that the rare c.313+1, G>A intronic variant leads to loss of *LDLR* function by constitutively promoting IVS3 retention[5].

Medina and colleagues previously found alternative splicing of *HMGCR*, *HMGCS1*, *MVK*, *PCSK9,* and *LDLR* to be mediated by the splice protein PTBP1 and regulated by cellular cholesterol levels[17]. Interestingly, PTBP1 works as an inhibitor of the U2AF splice component, and thus inhibits the recognition of 3' splice sites by the U2-spliceosome[18]. However, the knockdown of *PTBP1* resulted in very limited changes in the expression levels of the different splice forms[17], especially when compared to the drastic changes observed in our study.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

In our *in vitro* experiments, the knock-down of several U2-spliceosome genes and the resulting IVS3 retention compromised LDLR cell surface expression and LDL uptake as much as *LDLR* knockdown. The sensitivity of our mass spectrometric analysis only allowed detection of the tagged fragment after overexpression in the immortalized kidney cell line HEK293T. The artificial construct unlike an endogenously produced protein may have escaped nonsense-mediated decay. Nevertheless, we cannot rule out that the theoretical 116 amino acid long aminoterminal fragment of the differentially spliced *LDLR* is expressed *in vivo* and secreted. In fact, human plasma contains LDLR fragments, which are currently assumed to result from shedding of LDLR at the cell surface[19] but may also correspond to secreted alternative splice variants.

The relative expression of *LDLR's* IVS3 transcript in human liver varies strongly due to both analytical and biological reasons, namely between 0.4% and 29% upon RNA sequencing, between 7.5% and 81% upon chip array analysis, and between 23% and 85% upon RT-PCR. Very likely, RNA sequencing yielded the most realistic data, because this method recorded the different LDLR transcripts most comprehensively. The large interindividual variation of IVS3 expression recorded by each method indicates relevant regulatory mechanisms and consequences. We made controversial observations on the association of IVS3 retention with NAFLD. On the one hand, the percentage of IVS3 transcripts was significantly higher in 30 patients with NAFLD or NASH than in 25 normal weight and obese control subjects without NAFLD. On the other hand, the chip array analysis found significant increases of IVS3 transcripts after bariatric surgery, which rather causes regression of NAFLD. Larger studies are hence needed to answer the question how NAFLD influences the expression of functional and non-functional LDLR transcripts.

In peripheral blood cells but not in liver tissue, we found a significant correlation between plasma LDL-C levels and the IVS3 retention *LDLR* transcript, which was stronger than the correlation with the full-length *LDLR* transcript. Smaller sample size and narrower

For Circulation Research Peer Review. Do not distribute. Destroy after use.

range of LDL-C levels but also differences between tissues may be the reasons, why no significant correlations of LDL-C with any hepatic LDLR transcript expression were found..However, the associations of *RBM25* SNPs with differences in *RBM25* expression and lLDL-C levels and the higher than expected prevalence of rare *RBM25* loss-of-function variants in FH patients with no mutation in canonical FH genes suggest that the regulation of *LDLR* splicing by the U2-spliceosome contributes to the determination of LDL-C levels in humans.

The lack of association of hypercholesterolemia with rare variants of any other U2-spliceosome gene may reflect their intolerance to gross variation as suggested by pLI values close to 1. Also of note tour analysis of WES data of UK biobank only retrieved heterozygous mutations in U2-spliceosome genes whereas our knockdown experiments rather mimic homozygous conditions. Opposite effects on upstream regulators of *LDLR* may be another reason why the majority of SNPs and rare exome variants of the spliceosome genes do not show any association with LDL-C levels. The exclusive association of LDL-C with *RBM25* variants may also indicate that *RBM25* regulates LDL-C levels by mechanisms unrelated to the U2-spliceosome and intron 3 retention. In fact, *RBM25* also partakes in other spliceosomal subunits[20]. Of note, RNAi with *RBM25* had the weakest effects on *LDLR* splicing and overexpression of hypercholesterolemia associated RBM25 mutants in Huh7 cells resulted in lower LDL uptake without affecting the expression of the *LDLR* IVS3 transcript.

The correlation between ENST00000557958 expression in blood cells with age makes us hypothesize that age-related changes in the activity of the U2-spliceosome contributes to the increase in LDL-C that parallels ageing[21] but is not mechanistically understood. The functionality of the splicing process changes with ageing[22]. Somatic mutations or decreased expression of splice factor genes, notably SF3B1 and RBM25 have been implicated in age-related processes, including cancer[22,23]. The total number of alternatively spliced genes also increases with age[24]. Until recently, *SIRT1* is the only known gene involved in cholesterol

For Circulation Research Peer Review. Do not distribute. Destroy after use.

metabolism and atherosclerosis[25] whose alternative splicing may be disrupted with age[22]. One may speculate that either the epigenetic dysregulation of the activity of splice factor genes or the accumulation of somatic loss-of-function variants in liver cells may promote increases in LDL-C with age.

Our study has several strengths and limitations. First, our screening unravelled several novel candidate genes that regulate hepatic LDL uptake but missed canonical LDL uptake regulating genes such as *MYLIP*, *MBTPS1*, *PCSK9* or *SREBBP2*. A general reason is the not optimal signal to noise ratio of our screening. A specific reason for the missing of *MYLIP* or *PCSK9* is the optimization of our screening towards the discovery of loss of function effects. Second, our validation studies did not only confirm the limiting effect of U2-spliceosome genes on LDL uptake but unravelled a novel mechanism of LDL receptor regulation, namely IVS3 retention within an *LDLR* transcript which is translated into a truncated and non-functional receptor protein. In both human liver and peripheral blood cells, we demonstrate that this process happens at considerable quantity and interindividual variability, possibly influenced by aging and NAFLD. Third, *RBM25* was the only spliceosome gene affected by mutations associated with differences in LDL-C, perhaps because *RBM25* may tolerate loss of function better than other U2-spliceosome genes. However, we cannot rule out that RBM25 affects LDL metabolism beyond or even independently of *LDLR* splicing because both knockdown of *RBM25* and overexpression of loss of function mutants associated with hypercholesterolemia exerted in Huh7 cells stronger and more consistent effects on LDL uptake than on IVS3 retention in *LDLR*.

In conclusion, we identified IVS3 retention of *LDLR* uponloss of U2-spliceosome activity as a novel mechanism regulating LDLR activity in cells. The importance of this mechanism for the regulation of plasma LDL-C levels and thus determination of cardiovascular risk remains to be established by further studies.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Acknowledgements and Sources of Funding

This work was conducted as part of the "TransCard" project of the 7th Framework Program (FP7) granted by the European Commission, to J.A.K., A.T.H. and A.v.E. (number 603091) as well as partiallythe FP7 "RESOLVE" project (to J.T.H., B.S., A.V., S.F. L.V.G., and A.v.E.) and the European Genomic Institute for Diabetes (EGID, ANR-10-LABX-46 to BS). Additional work by A.v.E.'s team was funded by the Swiss National Science Foundation (31003A-160126, 310030-185109) and the Swiss Systems X program (2014/267 (MRD) HDL-X).

P.Z. received funding awards from the Swiss Atherosclerosis Society (AGLA and the DACH Society for Prevention of Cardiovascular Diseases.

G.P. received funding from the University of Zurich (Forschungskredit, grant no. FK-20-037).

J.A.K. is an Established Investigator from the Dutch Heart Foundation (2015T068).

The project was also supported by GeniusII (CVON2017-2020).

The L.V.T.S. team is supported by Fondation Maladies Rares, PHRC (AOM06024), and the national project CHOPIN (CHolesterol Personalized Innovation) granted by the Agence Nationale de la Recherche (ANR-16-RHUS-0007).

Y.A.K. is supported by a grant from Ministère de l'Education Nationale et de la Technologie (France).

J.T.H. was supported by an EMBO Long Term Fellowship (ALTF277-2014). B.S. is a recipient of an ERC Advanced Grant (no. 694717). Both are also supported by PreciNASH (ANR 16-RHUS-0006).

Research at the Antwerp University Hospital was supported by the European Union: FP6 (HEPADIP Contract LSHM-CT-2005-018734)

S. F. has a senior clinical research fellowship from the Fund for Scientific Research (FWO) Flanders (1802154N).

For Circulation Research Peer Review. Do not distribute. Destroy after use.

S.E.H. received grants RG3008 and PG008/08 from the British Heart Foundation, and the support of the UCLH NIHR BRC. S.E.H. directs the UK Children's FH Register which has been supported by a grant from Pfizer (24052829) given by the International Atherosclerosis Society.

We acknowledge the use of data from BIOS-consortium (http://wiki.bbmri.nl/wiki/BIOS_bios) which is funded by BBMRI-NL (NWO project 184.021.007).

Flow cytometry was performed with equipment of the flow cytometry facility, University of Zurich.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Disclosures

The RNAi screening was performed at the ETH ScopeM facility (R.M., M.S., S.F.N., A.J.R, S.S.). As by contract, one third of the service costs had to be paid by grants of A.v.E. to cover part of the costs for personnel, infrastructure, and maintenance. S.E.H. received fees for Advisory Boards of Novartis and Amryt and is the Medical Director of a UCL spin-out company StoreGene that offers to clinicians genetic testing for patients with FH

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Supplemental Materials

Materials & Methods

Online Supplemental Figures I to XI

Online Supplemental Tables I to XII

Major Resource Table

Original Western Blots

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Figure Legends

Figure 1. Identification and validation of U2-spliceosome genes as limiting factors for the uptake of LDL by Huh-7 cells. A. Schematic representation of the genome-wide image-based siRNA screening and data analysis process. **B. Functional association networks** for genes decreasing LDL uptake upon siRNA-mediated knockdown. Genes with $P < 1.0*10^{-3}$ for median cytoplasm intensity were selected as top hits. Spheres represent single genes. Edges represent known and predicted gene-gene relationships such as protein-protein interactions, co-expression and homology. The graph was produced using the STRING online tool (http://string-db.org/). The superimposed coloured circles are used to highlight the main functional clusters. **C and D. Effects of RNA-interference with U2-spliceosome genes on cell association of ^{125}I-LDL in Huh-7 cells.** 72 hours after transfection with siRNAs from Ambion (*LDLR*), Sigma (*RBM25*), or Dharmacon (all other genes)**,** cells were incubated for 2 hours at 37°C in the presence of 33.3µg/ml of ^{125}I-LDL in the presence or absence of 40x excess unlabelled LDL. Specific cell association was calculated as the difference between the two conditions. Data are expressed as means ±SD of 2 quadruplicate experiments; Statistical analysis was performed using Kruskal-Wallis test with Dunn's multiple comparisons test between the non-targeting (scrambled) and each targeting siRNA (**C**) or Mann-Whitney test (one-sided) between each vendor's targeting and non-targeting (scrambled) siRNAs (**D**). The respective p values are shown above each condition.

Figure 2. Loss of U2-spliceosome genes causes intron 3 retention in LDLR. A. ***LDLR* Exon level expression upon *AQR* knockdown**.Expression of the *LDLR* exons was recorded by RNA sequencing of Huh-7 cells 72 hours after knockdown of *AQR*. Segments represent differential exon usage in each sector of the *LDLR* genomic sequence as identified by the DEXSeq algorithm and as summarized in the linear representation below the graph. Canonical

For Circulation Research Peer Review. Do not distribute. Destroy after use.

exons of the ENST00000252444 full-length transcript are shown below the graph. Normalized read counts are reported on the y axis. The black arrow indicates the location of ENSG00000130164:E009, corresponding to the first half of intron 3. Data represent the average of three independent experiments. **B. ENSG00000130164:E009 is most strongly upregulated upon RNA interference with spliceosome genes.** Log2 fold change in gene expression at the exon level for the whole transcriptome after knockdown of *AQR* (x axis), and *SF3B1* (y axis) and *SF3A2* (z axis) in Huh-7 cells. The red circle highlights the position of ENSG00000130164:E009 corresponding to the first half of intron 3. **C. Correlation between LDLR intron 3 retention and LDL cell association.** Correlation between the log2 fold change in ENSG00000130164:E009 expression level and the decrease in ^{125}I-LDL cell association (same data as in Figure 1C) upon knockdown of each U2-spliceosome hit gene. Cells treated with a non-targeting siRNA were used as reference. Cell association is expressed as mean±SD. r and p-value were calculated according to Spearman.

Figure 3. Determination of LDLR intron 3 splice patterns. A. Cloning strategy and structure of the minigenes. The upper part of this panel shows the genomic location of the two segments of the LDLR gene that were cloned in each minigene, while the lower half shows a simplified structure of the pSPL3 minigene used to express them. Genomic coordinates refer to the hg19 assembly. Note that, due to primer design, MG_1 is 1bp shorter at its 5' end, starting at chr19:11,212,960. **B. Characterization of the splice products.** The graphs represent the mean RNA sequencing coverage at the Exon3-Intron 3 junction in two replicate samples for each condition. Coverage data were normalized to the average coverage for exon 3. MG_1/MG_2 = short / long minigene. **C. *In silico* BPS predictions for the acceptor site of *LDLR* exon 4.** BP score: final score (svm_score) according to the SVM-BP-finder

For Circulation Research Peer Review. Do not distribute. Destroy after use.

algorithm for the putative BPS sequence highlighted in red. A BPS is considered valid when located close to the AG exclusion zone, with BP-score > 0 and with svm_score > 0.

Figure 4. Effect of loss of spliceosome function on LDLR protein expression A. Schematic structure of the LDLR protein. (modified from[7]). LDLRret: intron 3 retention fragment, LBD: Ligand binding domain; L1-L7: LDLR class A domain; EGFPH: Epidermal growth factor precursor homology domain; β: beta propeller; O: O-linked sugar repeat; A/B/C: EGF-type repeat; TM: transmembrane domain. The red line represents the location of the last canonical amino acid found also in the LDLRret fragment, followed by 12 novel amino acids and by a stop codon. **B and C. Effect of *SF3B1* and *AQR* knockdown on LDLR protein levels.** LDLR protein levels in Huh-7 cells 72 hours after *SF3B1* or *AQR* knockdown. **B** shows a representative Western blot. **C,** shows the relative densities of LDLR bands normalized to TATA-binding-protein (TBP, loading control) after knockdown of AQR or SF3B1 relative to the non-targeting control. Data are means ±SD of three independent experiments. **D. Effect of *SF3B1* and *AQR* knockdown on LDLR cell surface levels.** LDLR cell surface levels in alive Huh-7 cells were measured by flow cytometry 72 hours after knockdown of *SF3B1* or *AQR*. siRNAs against *LDLR* were used as positive controls. The data are normalized to a non-targeting control and means±SD of 3 independent experiments). Numbers in **C** and **D** are p-values obtained by Kruskal-Wallis test with Dunn's multiple comparisons test between the non-targeting (scrambled) and respective targeting siRNA. **E-F. Overexpressed LDLRret fragment is retrieved in cell lysates and cell culture medium.** 48 hours after transfection in HEK293T cells, the HA-tagged version of the LDLRret fragment was detected-by western blot in both total cell lysates **(E,F)** and media **(F)**. Lysates after 2 and more hours of incubation were obtained after treatment with the proteasome inhibitor MG132 as indicated by the labels in (**E**). EV = pcDNA3.1 empty vector. HA-frag = hemagglutinin-tagged LDLRret fragment.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Figure 5. *LDLR* intron 3 retention in human liver. A. Detection of intron 3 retention in human liver by RT-PCR. Transcripts encoding full-length *LDLR* or the IVS3 retention variant were measured by RT-PCR and normalized to GAPDH mRNA levels in healthy liver tissue of 17 patients with benign liver tumors and in liver biopsies of 9 patients with suspected NAFLD. Each bar shows the relative expression of the two *LDLR* transcripts in one subject. **B and C**. **Percent expression of the LDLR transcript LDLR-206 with retention of intron 3 (C) and a full length LDLR transcript LDLR-208 (D) relative to the sum of all 13 LDLR transcripts in livers of 12 healthy subjects or 13 obese patients without non-alcoholic fatty liver disease (NAFLD), 15 patients with NAFLD and 15 patients with non-alcoholic steatohepatitis (NASH).** Computational analysis of previously published RNA seq data (Gene Expression Omnibus, accession number GSE126848)[8]. For all transcripts, see supplemental figure VII. The dark and light blue lines within the violin plots represent means and medians, respectively. Numbers indicate p-values obtained by comparisons of indicated groups using the Kruskal-Wallis test and adjusted for multiple testing using the Bonferroni correction.. **D. Expression of *LDLR* exons and introns in human liver.** The violin plots show the normalized signal intensities for probes mapping to the 5'-UTR, 3'-UTR, the exons and some introns of the LDLR gene in 155 obese non-diabetic subjects. Dots indicate median values. Error bars span from the 2.5th to the 97.5th percentile. Intron 3 is highlighted in red while the other introns are shown in grey. The location of each probe is depicted in the diagram below.

. Figure 6. Correlations of the *LDLR* ENST00000557958 (A, B) and ENST00000252444 transcripts (C, D) in whole blood samples with age (A, C) and LDL-C levels (B, D). Data is from 2,462 subjects of the BIOS population[10]. ENST00000557958 represents the intron 3 retention transcript (**A, B**). ENST00000252444 (**C, D**) was the only full-length *LDLR*

For Circulation Research Peer Review. Do not distribute. Destroy after use.

transcript detected in all samples analysed. r-values and p-values refer to a Spearman correlation analysis. Linear regression lines and their 95% confidence intervals are shown in blue and gray, respectively.

Figure 7. Association between *RBM25* variants and LDL-C in the UK Biobank dataset. A. Association of GWAS SNPs from 11 spliceosome genes with LDL-C in the UK Biobank dataset. The dashed red horizontal line indicates the threshold for statistical significance after Bonferroni correction for multiple testing of 1360 variants within the genes of interest ($p=3.7*10^{-5}$). Effect size and directionality are reported on the x axis as beta value. **B. Association between the rs17570658 genotype and *RBM25* expression in different tissues.** Data shown for skeletal muscle and tibial artery (both empirical $p < 1.0*10^{-8,}$ corrected for multiple testing across genes using Storey's q value method[26,27]. The horizontal white lines reflect medians; the upper and lower borders of the grey boxes reflect the 75th and 25th percentiles, respectively.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

References

1. Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, Hegele RA, Krauss RM, Raal FJ, Schunkert H, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement fromthe European Atherosclerosis Society Consensus Panel. Eur. Heart J. 2017;38:2459–2472.
2. Zanoni P, Velagapudi S, Yalcinkaya M, Rohrer L, von Eckardstein A. Endocytosis of lipoproteins. Atherosclerosis. 2018;275:273–295.
3. Wilkinson ME, Charenton C, Nagai K. RNA Splicing by the Spliceosome [Internet]. Annu. Rev. Biochem. 2020;89:359–388.
4. Emmer BT, Sherman EJ, Lascuna PJ, Graham SE, Willer CJ, Ginsburg D. Genome-scale CRISPR screening for modifiers of cellular LDL uptake. PLoS Genet. 2021;17:e1009285.
5. Cameron J, Holla ØL, Kulseth MA, Leren TP, Berge KE. Splice-site mutation c.313+1, G>A in intron 3 of the LDL receptor gene results in transcripts with skipping of exon 3 and inclusion of intron 3. Clin. Chim. Acta. 2009;403:131–135.
6. Corvelo A, Hallegger M, Smith CWJ, Eyras E. Genome-wide association between branch point properties and alternative splicing. PLoS Comput. Biol. 2010;6:e1001016.
7. Surdo P Lo, Bottomley MJ, Calzetta A, Settembre EC, Cirillo A, Pandit S, Ni YG, Hubbard B, Sitlani A, Carfí A. Mechanistic implications for LDL receptor degradation from the PCSK9/LDLR structure at neutral pH. EMBO Rep. 2011;12:1300–1305.
8. Suppli MP, Rigbolt KTG, Veidal SS, Heebøll S, Eriksen PL, Demant M, Bagger JI, Nielsen JC, Oró D, Thrane SW, et al. Hepatic transcriptome signatures in patients with varying degrees of nonalcoholic fatty liver disease compared with healthy normal-weight individuals. Am. J. Physiol. - Gastrointest. Liver Physiol. 2019;316:G462–G472.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

9. Lefebvre P, Lalloyer F, Baugé E, Pawlak M, Gheeraert C, Dehondt H, Vanhoutte J, Woitrain E, Hennuyer N, Mazuy C, et al. Interspecies NASH disease activity whole-genome profiling identifies a fibrogenic role of PPARα-regulated dermatopontin. J. Clin. Invest. 2017;2:1–17.

10. Zhernakova D V., Le TH, Kurilshikov A, Atanasovska B, Bonder MJ, Sanna S, Claringbould A, Võsa U, Deelen P, Franke L, et al. Individual variations in cardiovascular-disease-related protein levels are driven by genetics and gut microbiome. Nat. Genet. 2018;50:1524–1532.

11. Cirulli ET, White S, Read RW, Elhanan G, Metcalf WJ, Tanudjaja F, Fath DM, Sandoval E, Isaksson M, Schlauch KA, et al. Genome-wide rare variant analysis for thousands of phenotypes in over 70,000 exomes from two cohorts. Nat. Commun. 2020;11:542.

12. Lek M, Karczewski KJ, Minikel E V., Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536:285–291.

13. Liu DJ, Peloso GM, Yu H, Butterworth AS, Wang X, Mahajan A, Saleheen D, Emdin C, Alam D, Alves AC, et al. Exome-wide association study of plasma lipids in >300,000 individuals. Nat. Genet. 2017;49:1758–1766.

14. Willer CJ, Li Y, Abecasis GR. METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26:2190–2191.

15. Futema M, Plagnol V, Li K, Whittall RA, Neil HAW, Seed M, Bertolini S, Calandra S, Descamps OS, Graham CA, et al. Whole exome sequencing of familial hypercholesterolaemia patients negative for LDLR / APOB / PCSK9 mutations. J. Med. Genet. 2014;51:537–544.

16. Kraehling JR, Chidlow JH, Rajagopal C, Sugiyama MG, Fowler JW, Lee MY, Zhang X, Ramírez CM, Park EJ, Tao B, et al. Genome-wide RNAi screen reveals ALK1

For Circulation Research Peer Review. Do not distribute. Destroy after use.

mediates LDL uptake and transcytosis in endothelial cells. Nat. Commun. 2016;7:13516.

17. Medina MW, Krauss RM. Alternative splicing in the regulation of cholesterol homeostasis. Curr. Opin. Lipidol. 2013;24:147–52.

18. Shao C, Yang B, Wu T, Huang J, Tang P, Zhou Y, Zhou J, Qiu J, Jiang L, Li H, et al. Mechanisms for U2AF to define 3′ splice sites and regulate alternative splicing in the human genome. Nat. Struct. Mol. Biol. 2014;21:997–1005.

19. Mayne J, Ooi TC, Tepliakova L, Seebun D, Walker K, Mohottalage D, Ning Z, Abujrad H, Mbikay M, Wassef H, et al. Associations between Soluble LDLR and Lipoproteins in a White Cohort and the Effect of PCSK9 Loss-of-Function. J. Clin. Endocrinol. Metab. 2018;103:3486–3495.

20. Carlson SM, Soulette CM, Yang Z, Elias JE, Brooks AN, Gozani O. RBM25 is a global splicing factor promoting inclusion of alternatively spliced exons and is itself regulated by lysine mono-methylation. J. Biol. Chem. 2017;292:13381–13390.

21. Balder JW, de Vries JK, Nolte IM, Lansberg PJ, Kuivenhoven JA, Kamphuisen PW. Lipid and lipoprotein reference values from 133,450 Dutch Lifelines participants: Age- and gender-specific baseline lipid values and percentiles. J. Clin. Lipidol. 2017;11:1055-1064.e6.

22. Deschênes M, Chabot B. The emerging role of alternative splicing in senescence and aging [Internet]. Aging Cell. 2017;16:918–933.

23. Seiler M, Peng S, Agrawal AA, Palacino J, Teng T, Zhu P, Smith PG, Buonamici S, Yu L, Caesar-Johnson SJ, et al. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. Cell Rep. 2018;23:282-296.e4.

24. Rodríguez SA, Grochová D, Mckenna T, Borate B, Trivedi NS, Erdos MR, Eriksson M. Global genome splicing analysis reveals an increased number of alternatively

For Circulation Research Peer Review. Do not distribute. Destroy after use.

spliced genes with aging. Aging Cell. 2016;15:267–278.

25. Miranda MX, Van Tits LJ, Lohmann C, Arsiwala T, Winnik S, Tailleux A, Stein S, Gomes AP, Suri V, Ellis JL, et al. The Sirt1 activator SRT3025 provides atheroprotection in Apoe-/- mice by reducing hepatic Pcsk9 secretion and enhancing Ldlr expression. Eur. Heart J. 2015;36:51–59.

26. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, Mohammadi P, Park YS, Parsana P, Segrè A V., et al. Genetic effects on gene expression across human tissues. Nature. 2017;550:204–213.

27. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U. S. A. 2003;100:9440–9445.

28. Havel RJ, Eder HA, Bragdon JH. The distribution and chemical composition of ultracentrifugally separated lipoproteins in human serum. J. Clin. Invest. 1955;34:1345–1353.

29. Velagapudi S, Yalcinkaya M, Piemontese A, Meier R, Nørrelykke SF, Perisa D, Rzepiela A, Stebler M, Stoma S, Zanoni P, et al. VEGF-A regulates cellular localization of SR-BI as well as transendothelial transport of HDL but Not LDL. Arterioscler. Thromb. Vasc. Biol. 2017;37:794–803.

30. McFarlane AS. Efficient trace-labelling of proteins with iodine. Nature. 1958;182:53.

31. Kamentsky L, Jones TR, Fraser A, Bray MA, Logan DJ, Madden KL, Ljosa V, Rueden C, Eliceiri KW, Carpenter AE. Improved structure, function and compatibility for cellprofiler: Modular high-throughput image analysis software. Bioinformatics. 2011;27:1179–1180.

32. Parsons BD, Schindler A, Evans DH, Foley E. A Direct Phenotypic Comparison of siRNA Pools and Multiple Individual Duplexes in a Functional Assay. PLoS One. 2009;4:e8471.

33. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina

For Circulation Research Peer Review. Do not distribute. Destroy after use.

sequence data. Bioinformatics. 2014;30:2114–2120.

34. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

35. Liao Y, Smyth GK, Shi W. The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. 2013;41:e108.

36. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Res. 2012;22:2008–2017.

37. Fedoseienko A, Wijers M, Wolters JC, Dekker D, Smit M, Huijkman N, Kloosterhuis N, Klug H, Schepers A, van Dijk KW, et al. The COMMD family regulates plasma LDL levels and attenuates atherosclerosis through stabilizing the CCC complex in endosomal LDLR traffcking. Circ. Res. 2018;122:1648–1660.

38. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010;26:966–968.

39. Nisson PE, Watkins PC, Krizman DB. Isolation of exons from cloned DNA by exon trapping. Curr. Protoc. Hum. Genet. 2001;Chapter 6:Unit 6.1.

40. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. 2016;34:525–527.

41. Cirulli, E.T. and Washington NL. Helix Research. UK Biobank Exome rare variant analysis v1.3 [Internet]. Helix Blog.

42. Whiffin N, Minikel E, Walsh R, O'Donnell-Luria AH, Karczewski K, Ing AY, Barton PJR, Funke B, Cook SA, Macarthur D, et al. Using high-resolution variant frequencies to empower clinical genome interpretation. Genet. Med. 2017;19:1151–1158.

43. Birmingham A, Selfors LM, Forster T, Wrobel D, Kennedy CJ, Shanks E, Santoyo-

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Lopez J, Dunican DJ, Long A, Kelleher D, et al. Statistical methods for analysis of high-throughput RNA interference screens. Nat. Methods. 2009;6:569–75.

44. König R, Chiang C, Tu BP, Yan SF, DeJesus PD, Romero A, Bergauer T, Orth A, Krueger U, Zhou Y, et al. A probability-based approach for the analysis of large-scale RNAi screens. Nat. Methods. 2007;4:847–849.

45. Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. Science (80-.). 2000;290:2323–2326.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Table 1. Hit genes that induced upon knockdown in Huh-7 cells either a decrease (left column) or an increase (right column) in LDL uptake.

Decreased LDL uptake				Increased LDL uptake			
Gene	**Assay score[A] avg[B]**	**Assay score[A] SEM[C]**	**RSA p-value[D]**	**Gene**	**Assay score[A] avg[B]**	**Assay score[A] SEM[C]**	**RSA p-value[D]**
AP2M1	-3.103179681	0.346648222	$3.36*10^{-8}$	*PROX1*	6.53057396	0.631260417	$3.19*10^{-9}$
CHMP2A	-3.130900347	0.359445533	$2.51*10^{-7}$	*ITGAV*	7.431175355	1.519558432	$2.96*10^{-8}$
NFKB2	-2.59157417	0.136886566	$8.07*10^{-7}$	*TGFBR1*	3.464028514	0.397588943	$7.31*10^{-6}$
AQR[E]	-2.484868551	0.199482589	$4.57*10^{-6}$	*CDC37*	3.747034032	1.072191825	$2.35*10^{-5}$
PSMD11	-2.557101583	0.239773488	$4.77*10^{-6}$	*DTNBP1*	57.92944887	57.2617451	$4.46*10^{-5}$
SF3B2	-2.107311389	0.015210399	$4.81*10^{-6}$	*CYP27C1*	32.06817221	31.61438081	$8.92*10^{-5}$
RPL35	-2.346954606	0.150946677	$5.45*10^{-6}$	*PNPLA2*	2.279278207	0.266420784	$1.26*10^{-4}$
PSMD8	-2.988677308	0.491086915	$6.34*10^{-6}$	*C22orf39*	7.995494448	8.342242785	$1.78*10^{-4}$
SON	-2.164748153	0.201099955	$1.46*10^{-5}$	*TMEM133*	3.049034762	1.060165442	$1.84*10^{-4}$
COPA	-2.307675328	0.213018879	$1.61*10^{-5}$	*TMEM130*	6.466491664	5.317700355	$2.23*10^{-4}$
RBM25	-1.993998657	0.055265194	$1.92*10^{-5}$	*PM20D2*	2.155202336	0.176491898	$2.29*10^{-4}$
RBM22	-2.818121291	0.622885617	$3.36*10^{-5}$	*PET117*	3.001069341	1.652765441	$2.68*10^{-4}$
PSMD3	-2.21302629	0.224903034	$3.98*10^{-5}$	*CWF19L2*	3.806757511	4.571696977	$3.12*10^{-4}$
SF3B5	-2.285064158	0.25878823	$4.32*10^{-5}$	*ENY2*	2.420424347	0.514153659	$3.28*10^{-4}$
SF3B1	-2.267932169	0.253122099	$4.55*10^{-5}$	*NME4*	2.711491413	0.954425612	$3.39*10^{-4}$
SALL4	-1.937993523	1.13065979	$6.02*10^{-5}$	*ZC3H4*	4.545156994	3.551478266	$3.57*10^{-4}$
RPL5	-2.106905493	0.373542859	$7.40*10^{-5}$	*WASF2*	2.310874515	0.449202822	$3.61*10^{-4}$
CCDC180	-1.132459235	1.398333381	$9.52*10^{-5}$	*HELZ2*	2.546828237	0.984740435	$3.87*10^{-4}$
SF3B6	-2.277000896	0.332074616	$9.83*10^{-5}$	*RILP*	1.995550072	0.267567916	$4.23*10^{-4}$
HNRNPU	-1.724036435	0.093847304	$1.23*10^{-4}$	*MAT2A*	3.705559066	3.611891772	$4.91*10^{-4}$
RPL17	-2.226845162	0.329575956	$1.46*10^{-4}$	*NRM*	1.710898743	0.050727817	$5.02*10^{-4}$
ISY1	-2.74487386	0.698388989	$1.55*10^{-4}$	*CEP295NL*	2.189792071	0.474598108	$5.02*10^{-4}$
ZNF641	-1.034460324	1.453444444	$2.58*10^{-4}$	*ACSM2A*	2.207444199	1.531809937	$5.32*10^{-4}$
COPB1	-1.693933632	0.103029465	$2.64*10^{-4}$	*RTL9*	3.759306708	3.473297986	$5.35*10^{-4}$
SF3A1	-2.225755015	0.412106586	$2.72*10^{-4}$	*KIAA1522*	3.362058267	3.27466253	$6.25*10^{-4}$
SNW1	-1.76539067	0.142531611	$2.76*10^{-4}$	*ZNF84*	2.204388764	0.765657329	$6.55*10^{-4}$
EIF2S1	-1.486721463	0.790741651	$3.45*10^{-4}$	*TFAP4*	3.032765033	3.340175625	$6.69*10^{-4}$
CCDC73	-1.041204586	1.27682775	$3.50*10^{-4}$	*TMEM182*	3.227517874	1.666669492	$7.29*10^{-4}$
RPL9	-1.715182797	0.249911985	$3.55*10^{-4}$	*WDR55*	1.967286849	1.365170916	$7.32*10^{-4}$
NXNL2	-1.199311468	1.135835784	$3.83*10^{-4}$	*DYNLL1*	2.268266743	0.467997927	$7.72*10^{-4}$
WBP11	-1.50591484	0.062444555	$4.03*10^{-4}$	*ADPRHL2*	2.078229013	0.322800093	$8.51*10^{-4}$
C2CD5	-1.097951788	1.954971449	$4.46*10^{-4}$	*ELAVL1*	1.945364959	0.968117905	$8.70*10^{-4}$
RPL21	-1.655773242	0.156797718	$4.72*10^{-4}$	*CFAP298*	1.883199038	0.378258022	$8.87*10^{-4}$
EPOP	-1.837314819	0.25795876	$4.80*10^{-4}$	*PMM1*	2.80926863	3.200260012	$8.92*10^{-4}$
RMND5B	-1.523957521	0.076773849	$5.07*10^{-4}$	*CASKIN2*	1.681223061	0.149986926	$9.07*10^{-4}$
TAPBPL	-1.52965773	0.154207886	$5.27*10^{-4}$	*CIZ1*	3.454694336	2.803876145	$9.37*10^{-4}$

STARD10	-1.527795273	0.115135889	$5.45*10^{-4}$	*BRICD5*	1.962503862	0.408074057	$9.41*10^{-4}$
PSMD1	-2.207116523	0.551747426	$5.63*10^{-4}$				
PFDN6	-0.881689024	1.740601376	$5.80*10^{-4}$				
PSMA1	-1.528976301	0.119805079	$5.85*10^{-4}$				
RTF2	-1.573924771	0.169765686	$6.14*10^{-4}$				
LSM2	-1.448888015	0.056454941	$6.40*10^{-4}$				
UBD	-1.171691024	1.530009178	$6.69*10^{-4}$				
LRRC14	-1.258311764	1.067910962	$6.84*10^{-4}$				
SUPT6H	-1.451332382	0.095214513	$7.27*10^{-4}$				
COPB2	-2.037140764	0.468882876	$7.34*10^{-4}$				
SF3A2	-1.347147433	0.758462926	$7.89*10^{-4}$				
ATP6V0C	-1.823918839	0.263639476	$7.90*10^{-4}$				
EMILIN3	-1.598631472	2.238859705	$8.03*10^{-4}$				
DMTN	-1.559252376	0.142024687	$8.20*10^{-4}$				
MRPL19	-0.755460842	1.688052373	$8.92*10^{-4}$				
MRO	-0.986783025	1.102624895	$9.14*10^{-4}$				
DDX59	-1.380513222	1.040634076	$9.25*10^{-4}$				
PSMD12	-1.761325035	0.367766123	$9.45*10^{-4}$				

[A]Assay score: normalized score for the median cytoplasm intensity assay feature. [B]Avg = average.

[C]SEM: standard error of the mean.

[D] p-values are not adjusted for multiple testing ($p < 3.5*10^{-6}$ after Bonferroni adjustment for 14’000 genes with expressed transcripts)

[E] The 15 hit genes involved in RNA splicing and validated in this study are highlighted in bold and underlined.

For Circulation Research Peer Review. Do not distribute. Destroy after use.

Figure 1

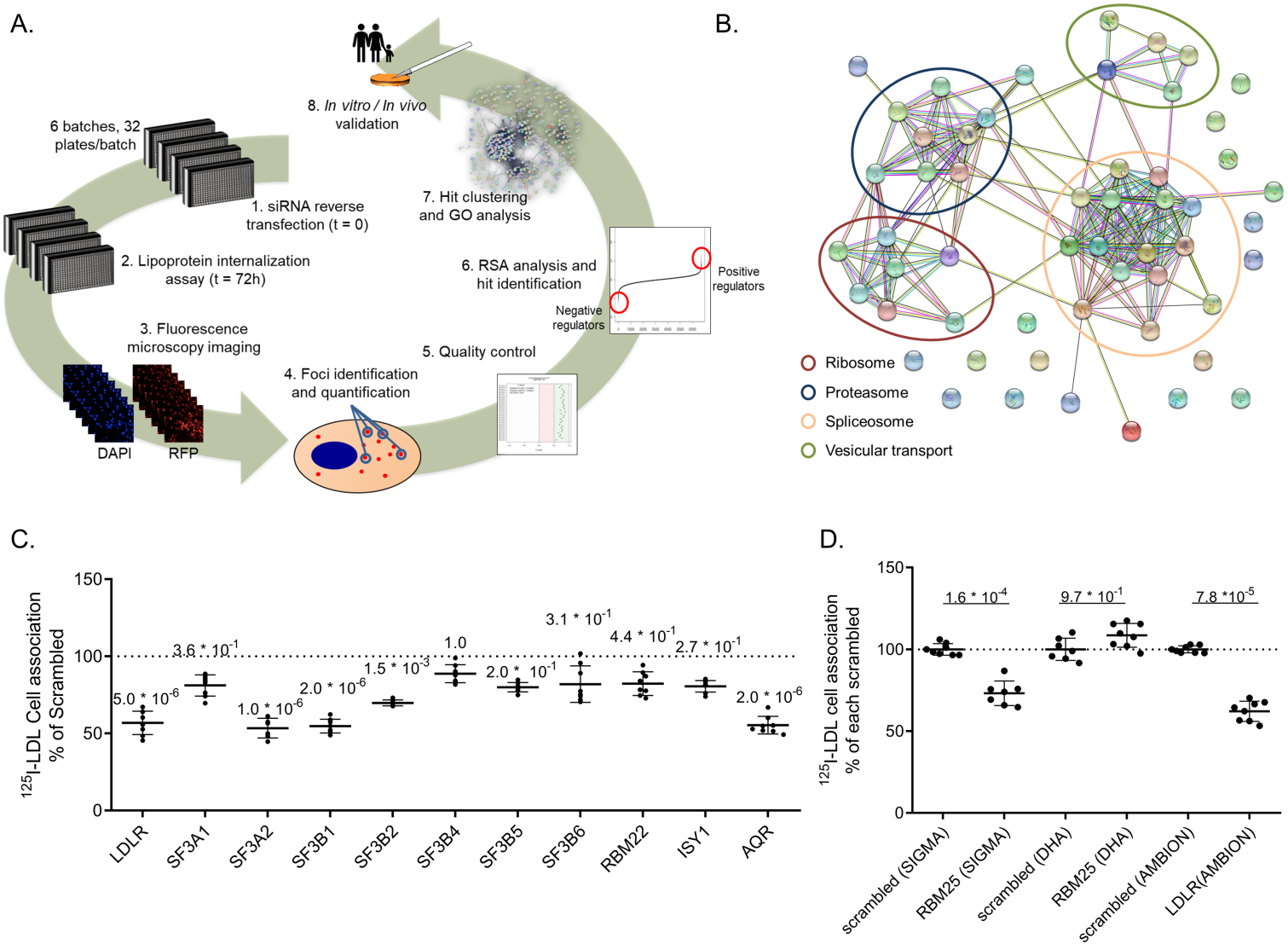
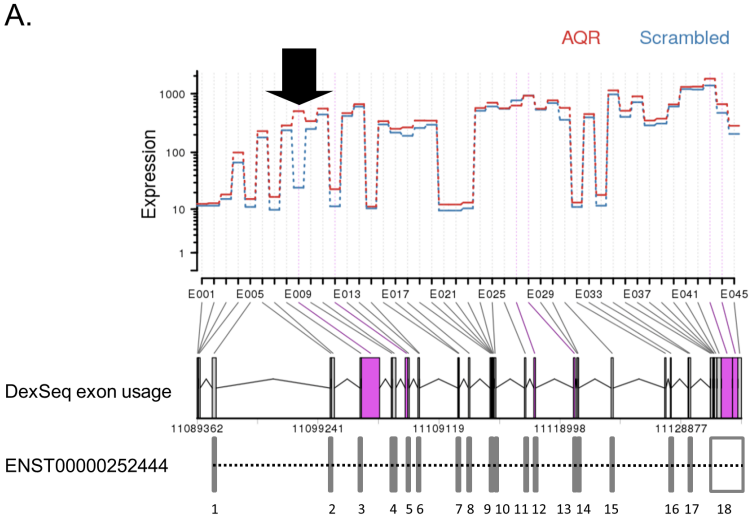
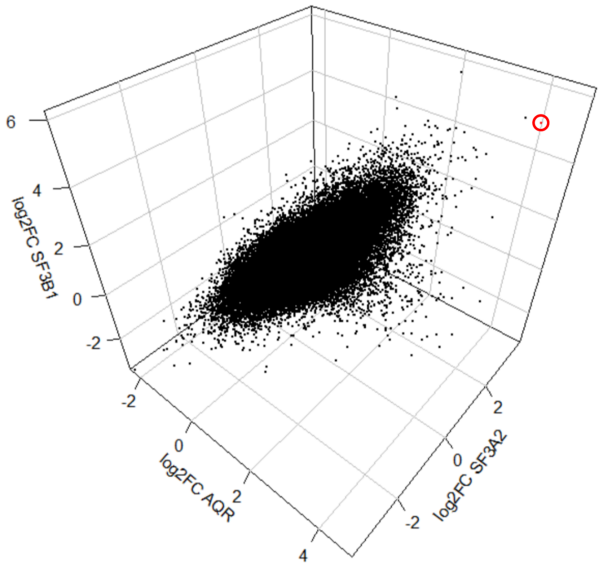

A.
6 batches, 32 plates/batch
1. siRNA reverse transfection (t = 0)
2. Lipoprotein internalization assay (t = 72h)
3. Fluorescence microscopy imaging
DAPI
RFP
4. Foci identification and quantification
5. Quality control
6. RSA analysis and hit identification
Positive regulators
Negative regulators
7. Hit clustering and GO analysis
8. In vitro / In vivo validation
B.
Ribosome
Proteasome
Spliceosome
Vesicular transport
C.
125I-LDL Cell association % of Scrambled
150
100
50
0
5.0 * 10-6
3.6 * 10-1
1.0 * 10-6
2.0 * 10-6
1.5 * 10-3
1.0
2.0 * 10-1
3.1 * 10-1
4.4 * 10-1
2.7 * 10-1
2.0 * 10-6
LDLR
SF3A1
SF3A2
SF3B1
SF3B2
SF3B4
SF3B5
SF3B6
RBM22
ISY1
AQR
D.
125I-LDL cell association % of each scrambled
150
100
50
0
1.6 * 10-4
9.7 * 10-1
7.8 *10-5
scrambled (SIGMA)
RBM25 (SIGMA)
scrambled (DHA)
RBM25 (DHA)
scrambled (AMBION)
LDLR(AMBION)

Figure 2

A.


AQR
Scrambled
Expression
1000
100
10
1
E001 E005 E009 E013 E017 E021 E025 E029 E033 E037 E041 E045
DexSeq exon usage
11089362 11099241 11109119 11118998 11128877
ENST00000252444
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

B.


log2FC SF3B1
log2FC AQR
log2FC SF3A2

C.

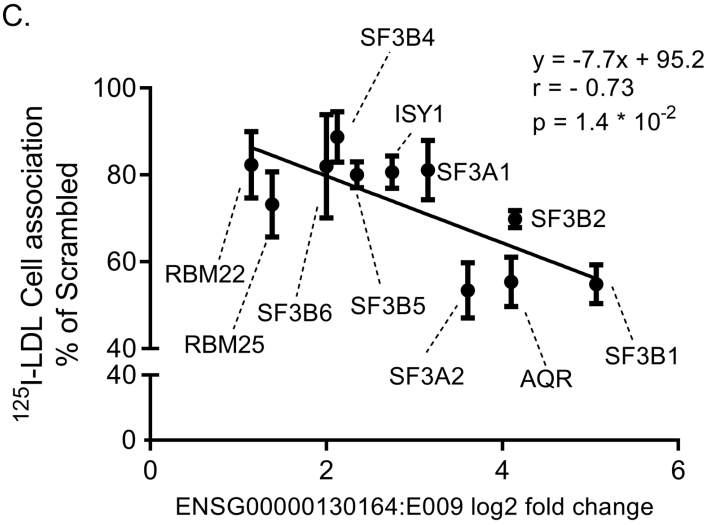
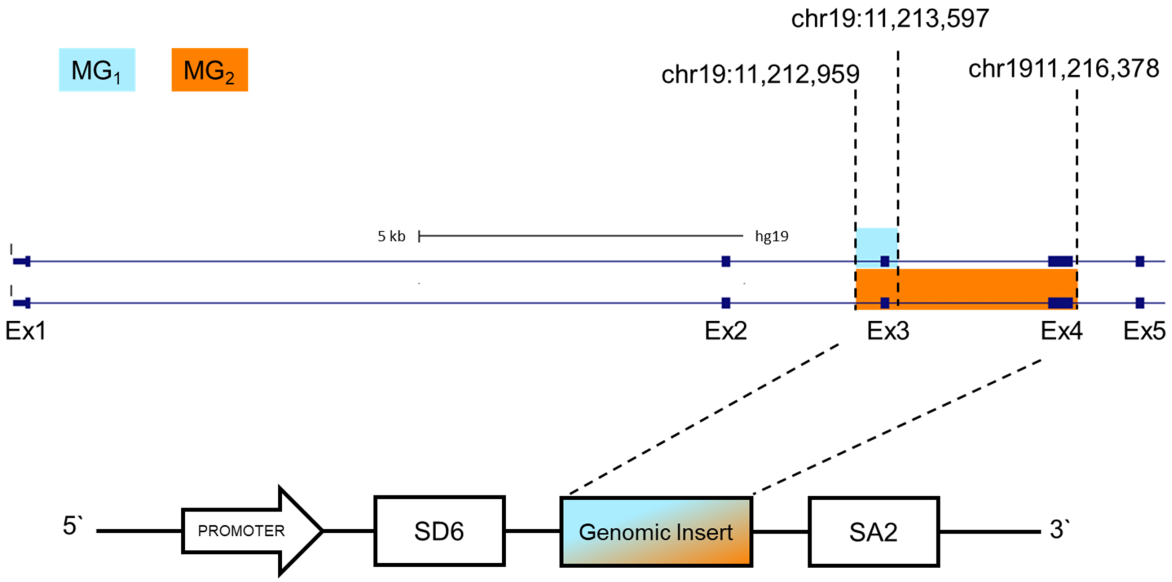
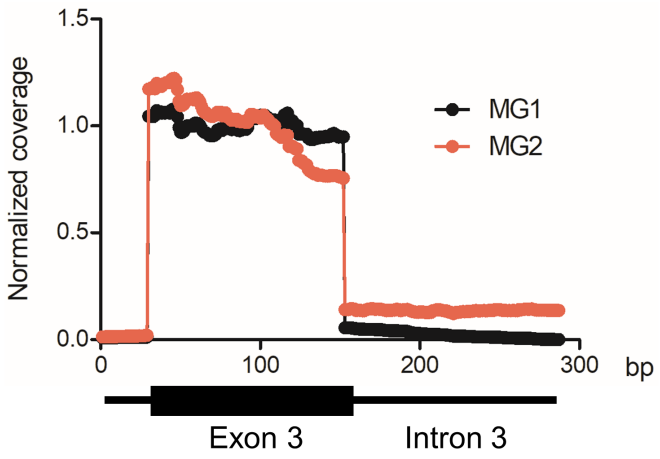

SF3B4
y = -7.7x + 95.2
r = - 0.73
p = 1.4 * 10-2
ISY1
SF3A1
SF3B2
RBM22
SF3B6
SF3B5
RBM25
SF3A2
AQR
SF3B1
125I-LDL Cell association
% of Scrambled
ENSG00000130164:E009 log2 fold change

Figure 3

A.


chr19:11,213,597
MG1
MG2
chr19:11,212,959
chr1911,216,378
5 kb
hg19
Ex1
Ex2
Ex3
Ex4
Ex5
5`
PROMOTER
SD6
Genomic Insert
SA2
3`

B.


Normalized coverage
1.5
1.0
0.5
0.0
MG1
MG2
0
100
200
300
bp
Exon 3
Intron 3

C.


BP score: 0.68379
Homo sapiens 5'-tcacacggtgatggtggtctcggcccatccatccctgcag-
Exon 4
BP score: 1.35885
Mus musculus 5'-tagcctgacacggccctctcctccctgtgcacccctgcag-
Exon 4

Figure 4

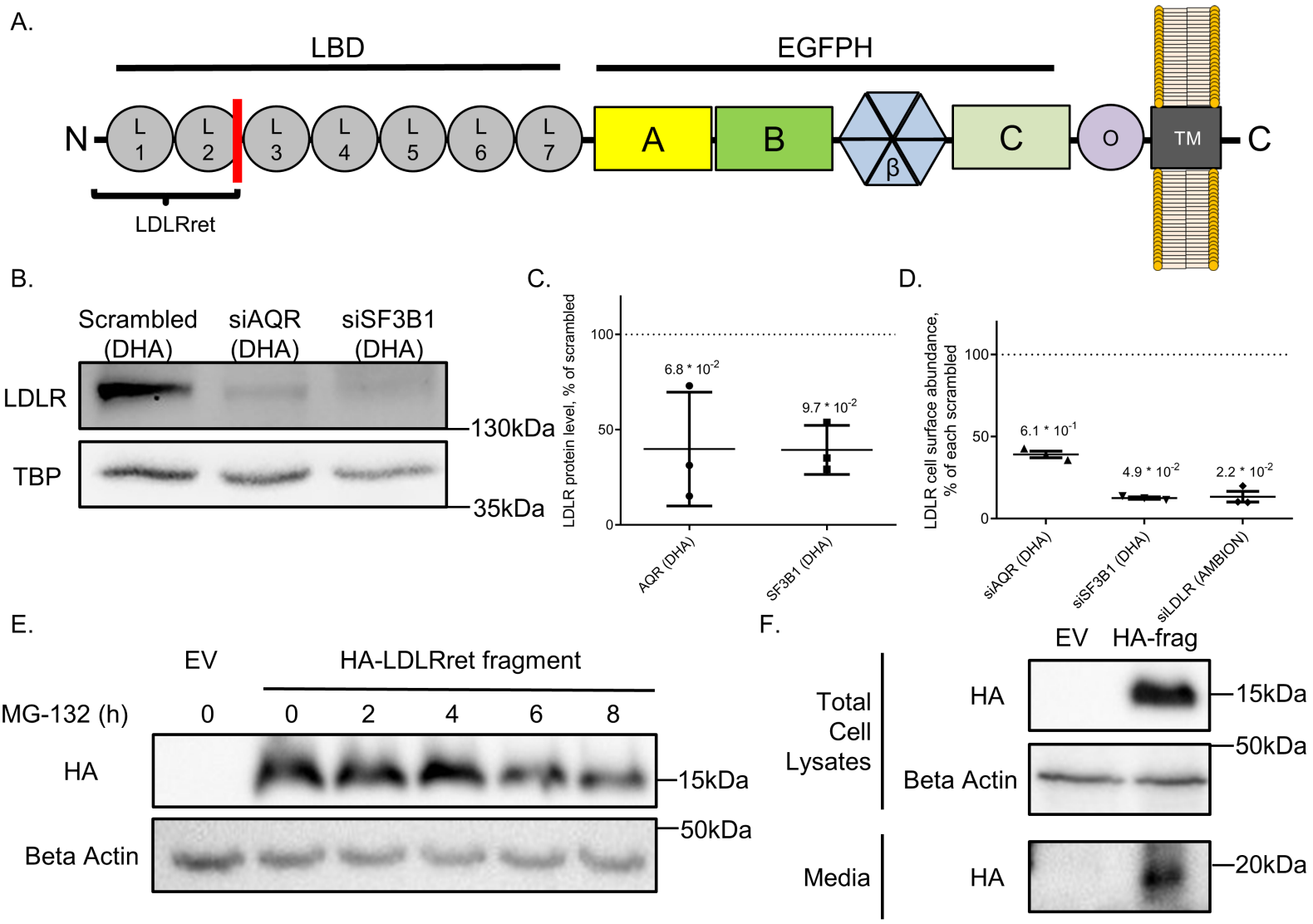

A.
LBD
EGFPH
N
L 1
L 2
L 3
L 4
L 5
L 6
L 7
A
B
β
C
O
TM
C
LDLRret
B.
Scrambled (DHA)
siAQR (DHA)
siSF3B1 (DHA)
LDLR
130kDa
TBP
35kDa
C.
LDLR protein level, % of scrambled
100
50
0
6.8 * 10-2
9.7 * 10-2
AQR (DHA)
SF3B1 (DHA)
D.
LDLR cell surface abundance, % of each scrambled
100
50
0
6.1 * 10-1
4.9 * 10-2
2.2 * 10-2
siAQR (DHA)
siSF3B1 (DHA)
siLDLR (AMBION)
E.
EV
HA-LDLRret fragment
MG-132 (h)
0
0
2
4
6
8
HA
15kDa
Beta Actin
50kDa
F.
EV
HA-frag
Total Cell Lysates
HA
15kDa
Beta Actin
50kDa
Media
HA
20kDa

Figure 5

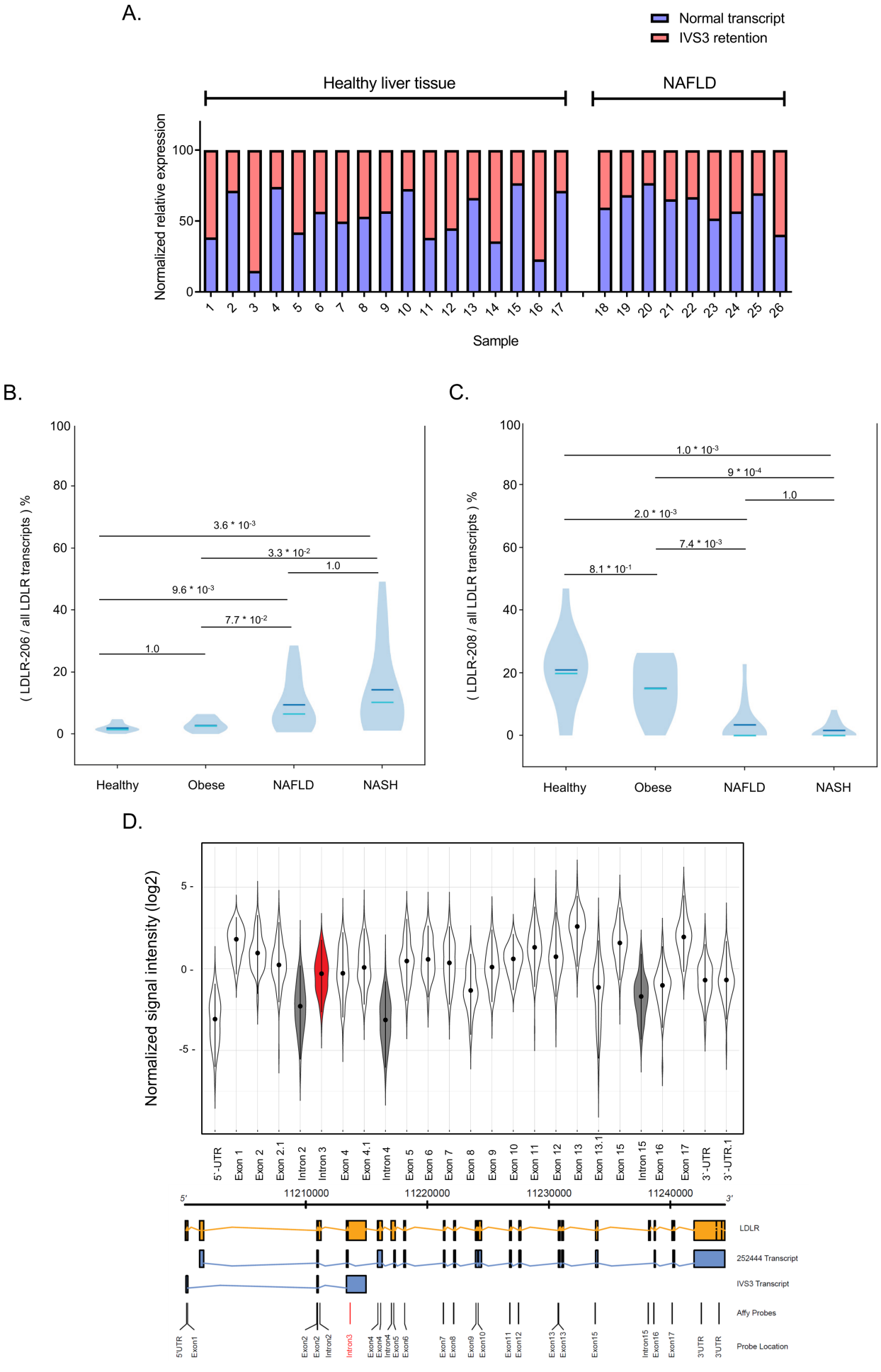

A.
Normal transcript
IVS3 retention
Healthy liver tissue
NAFLD
Normalized relative expression
100
50
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Sample
B.
(LDLR-206 / all LDLR transcripts) %
100
80
60
40
20
0
3.6 * 10-3
3.3 * 10-2
1.0
9.6 * 10-3
7.7 * 10-2
1.0
Healthy
Obese
NAFLD
NASH
C.
(LDLR-208 / all LDLR transcripts) %
100
80
60
40
20
0
1.0 * 10-3
9 * 10-4
1.0
2.0 * 10-3
7.4 * 10-3
8.1 * 10-1
Healthy
Obese
NAFLD
NASH
D.
Normalized signal intensity (log2)
5
0
-5
5'-UTR Exon 1 Exon 2 Exon 2.1 Intron 2 Intron 3 Exon 4 Exon 4.1 Intron 4 Exon 5 Exon 6 Exon 7 Exon 8 Exon 9 Exon 10 Exon 11 Exon 12 Exon 13 Exon 13.1 Exon 15 Intron 15 Exon 16 Exon 17 3'-UTR 3'-UTR.1
5'
11210000
11220000
11230000
11240000
3'
LDLR
252444 Transcript
IVS3 Transcript
Affy Probes
Probe Location
5'UTR Exon1
Exon2 Exon2 Intron2
Intron3
Exon4 Exon4 Intron4 Exon5 Exon6
Exon7 Exon8
Exon9 Exon10
Exon11 Exon12
Exon13 Exon13
Exon15
Intron15 Exon16 Exon17
3'UTR
3'UTR

Figure 6

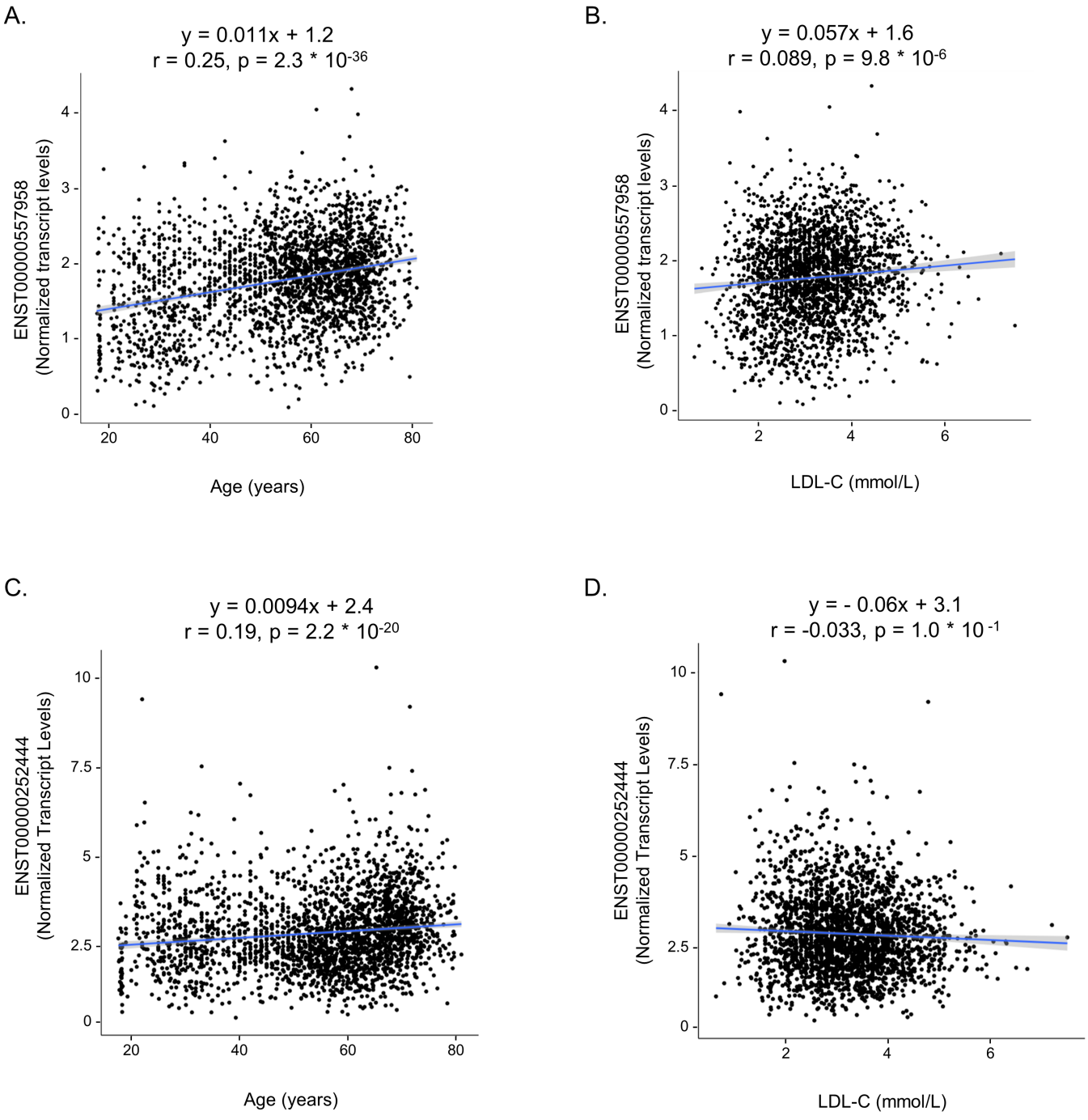
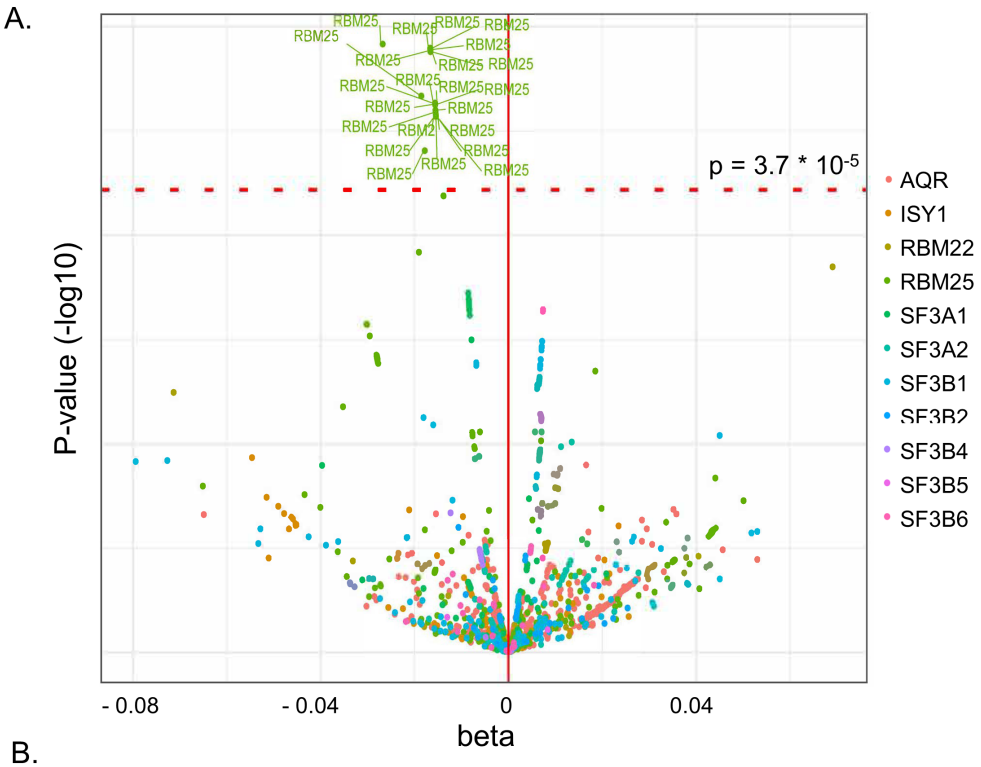
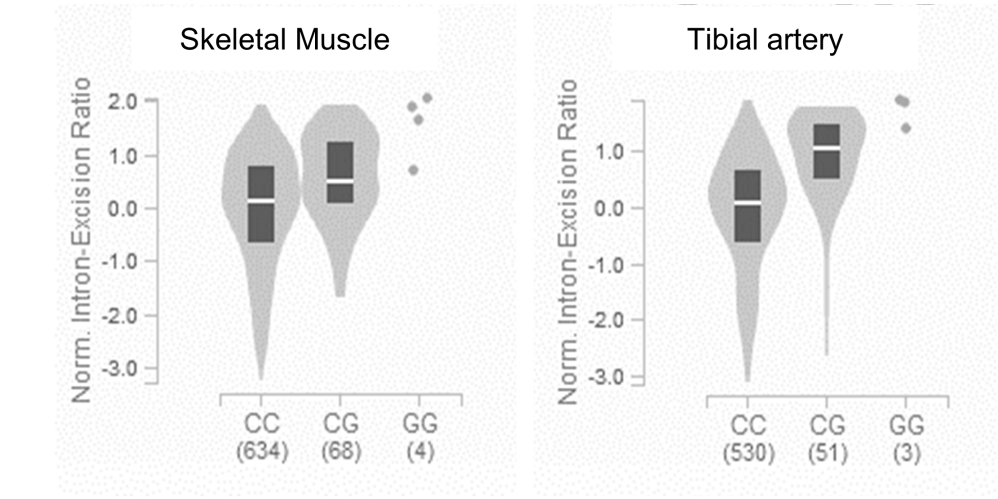

A.
y = 0.011x + 1.2
r = 0.25, p = 2.3 * 10-36
ENST00000557958
(Normalized transcript levels)
Age (years)
B.
y = 0.057x + 1.6
r = 0.089, p = 9.8 * 10-6
ENST00000557958
(Normalized transcript levels)
LDL-C (mmol/L)
C.
y = 0.0094x + 2.4
r = 0.19, p = 2.2 * 10-20
ENST00000252444
(Normalized Transcript Levels)
Age (years)
D.
y = - 0.06x + 3.1
r = -0.033, p = 1.0 * 10-1
ENST00000252444
(Normalized Transcript Levels)
LDL-C (mmol/L)

Figure 7


A.
P-value (-log10)
p = 3.7 * 10-5
RBM25
- 0.08
- 0.04
0
0.04
beta
AQR
ISY1
RBM22
RBM25
SF3A1
SF3A2
SF3B1
SF3B2
SF3B4
SF3B5
SF3B6


B.
Skeletal Muscle
Tibial artery
Norm. Intron-Excision Ratio
2.0
1.0
0.0
-1.0
-2.0
-3.0
CC (634)
CG (68)
GG (4)
CC (530)
CG (51)
GG (3)