Translating disease to cardiovascular health
Contract number: FP7- HEALTH n°603091
Start Date: 01/09/2013 - Duration: 60 months
Coordinator: JA. Kuivenhoven, UMCG

PROJECT FINAL REPORT

Grant Agreement number: 603091

Project acronym: TRANSCARD

Project title: Translating disease to cardiovascular health

Funding Scheme: Collaborative project

Period covered: **from** 01/09/2013 **to** 31/08/2018

Name of the scientific representative of the project's co-ordinator[1], Title and Organisation:

Dr. Jan Albert Kuivenhoven, Academic Medical Center Groningen

Tel: +31 503632687

E-mail: j.a.kuivenhoven@umcg.nl

Project website address: www.transcard-research.org

[1] Usually the contact person of the coordinator as specified in Art. 8.1. of the Grant Agreement.

Table of contents

Final publishable summary report

1.1. Executive summary

TransCard concerned a 5-years project in which four academic and two industrial partners from Switzerland, Germany, Denmark and the Netherlands set out to find and study new targets for possible pharmaceutical intervention to treat atherosclerosis. This pathology underlies cardiovascular disease (CVD), the major cause of death in the European Union and worldwide. For recognition: a heart attack or myocardial infarction (MI) is a main outcome of CVD.

Atherosclerosis is a lipid-driven inflammatory disease of the blood vessel wall. TransCard has specifically addressed the lipid component of this chronic disorder with a main focus on cholesterol in particles that are known as low-density lipoproteins (LDL). This because LDL cholesterol is a major causal but modifiable risk factor for CVD. In other words, increased LDL cholesterol increases the risk of cardiovascular complications such as MI while reducing LDL cholesterol – through intervention can decrease this risk.

The expertise of the academic partners covers the etiology and treatment of human lipid disorders (AMC), cellular lipid metabolism and transport of lipid-containing particles (lipoproteins) (UZH), the genetic background of changes in plasma lipid levels and the risk of suffering from CVD in the general population (REGION-H), and knowledge of experimental animal studies in the lipid and cell biology research fields (UMCG). The industrial partners NEBION and POLYQANT have provided expertise on gene expression (mRNA) and the quantification of proteins in TransCard's research activities.

Understanding the genetic background of human lipid disorders has historically been instrumental to find novel targets for pharmaceutical intervention. The AMC and UMCG have a longstanding history in studying patients with unexplained very high cholesterol as well as individuals with very low LDL cholesterol. Studies in families of such index cases have so far not delivered what we had been hoping for. Parallel to this family-based high-risk part of the project, TransCard has however also employed alternative target identification strategies: ***1)*** In direct collaboration, NEBION and UMCG found several very interesting targets through *in silico* studies meaning here that computers were used to combine large publicly available mRNA datasets (curated by NEBION) with key knowledge of lipid metabolism. This has led to the identification of several intriguing targets that are studied in test tubes as well as experimental animals in ongoing experiments to verify their biological significance. ***2)*** A whole genome-wide based screen was conducted by UZH for factors that control the uptake of LDL cholesterol by the liver from blood. This screen has led to the identification of four clusters of targets that are largely involved in unanticipated biological pathways. In addition, other targets involved in the transport of lipoprotein over the inner lining of blood vessels, an important process in atherosclerosis, were identified. ***3)*** Studies in mice with elevated LDL cholesterol by the UMCG at the start of TransCard have furthermore proven extremely valuable: these studies have led to the discovery of numerous factors that regulate the capacity of the main receptor for LDL in the liver to take up cholesterol from blood. Multiple newly identified targets are/have been studied in experimental animals and tissue culture studies. These studies were supported by the technology that was provided by POLYQUANT. Throughout the project, REGION-H has helped TransCard with validating the role of (new) targets through studying the

association between variation in target genes with plasma lipids and with the risk of CVD. All targets identified in TransCard are currently studied in patients with unexplained extremely high LDL or low LDL cholesterol through dedicated DNA sequencing (AMC/UMCG) and through further experimental studies. Investigations into the feasibility of identified targets to be used for pharmaceutical intervention are ongoing as anticipated at the start of this project. Paragraph 1.3.6 (Dissemination) and 1.4.1 (Impact) provided the context of this important note and also summarize the major achievements of this collaborative effort.

1.2. Summary description of project context and objectives

1.2.1 Context (State of Art at the project start, Key challenges)

State of the Art at the project start

Five years ago, the worldwide impact of atherosclerosis, the pathology underlying cardiovascular diseases (CVD) was devastating and it still is: The World Health Organization currently indicates that CVD takes the lives of 17.9 billion people every year, 31% of all deaths (October 2018). Risk factors for developing CVD include tobacco smoking, increased blood pressure, sedentary lifestyle and male gender. From a blood lipid perspective, there is moreover unequivocal evidence that increases and decreases in LDL cholesterol cause increased and decreased risk of CVD, respectively. In this light, TransCard has primarily focused on the identification of factors that control the levels of LDL cholesterol as an important modifiable risk factor in blood. The evidence that an increased level of triglycerides, another lipid in blood, is also an independent risk factor for CVD has become more prominent over only the last few years due to large-scale genetic (Mendelian Randomization) studies. In these studies, triglycerides are now generally acknowledged a (bio)marker for remnant cholesterol. The same randomization method was more recently used to also provide evidence for a causal role of an LDL-like particle, known as lipoprotein (a), in atherogenesis. Until to date, there are no registered dugs that reduce risk of CVD through reducing triglycerides and or lipoprotein (a) levels but with the recently collected evidence this will certainly be addressed in the near future. At the start of TransCard, it was already known that pharmaceutical targeting of HDL metabolism to decrease CVD risk is very challenging and this has remained to be the case until to date. In 2013, it was finally not yet clear whether reducing LDL cholesterol with a new class of drugs, i.e. PCSK9 inhibitors, would protect against CVD without significant adverse effects. This background provides the strong rationale for TransCard to focus on the biology of LDL.

Key challenges

The patient suffering from high levels of LDL cholesterol of unknown etiology (or other extreme blood lipid phenotypes) has been central to TransCard. This because the understanding of the mechanism that cause high levels of LDL cholesterol can be instrumental to develop new therapeutics to reduce the burden of CVD. This philosophy actually gave rise to the acronym ‘TransCard’ of this collaborative effort: ‘*Translating Disease into Cardiovascular Health*’. The key challenge was thus to identify new origins of severe lipid disorders (‘targets’) such as very high LDL cholesterol (hypercholesterolemia) in patients. However, the ‘low hanging fruit’ (genetic causes of severe hypercholesterolemia) was already picked and exploited in the 25 years preceding the start of our effort: the identification of the LDL receptor (LDLR) for the hepatic uptake of circulating LDL

through studies in families with familial hypercholesterolemia (FH; Nobel Prize 1986) ultimately led to the development of statins. The subsequent identification of apolipoprotein (apo) B led to pharmaceutical targeting of *APOB* gene products (Mipomersen) while the identification of proprotein convertase subtilisin/kexin type 9 (*PCSK9*) in a single family with FH led to the development of an entire new class of drugs. Although approximately 50% of severe hypercholesterolemia cannot be explained with defects in any of the above genes, it was clear that it would be a challenge for TransCard to find novel origins of FH or of any other severe lipid disorder in HDL or triglyceride metabolism. To increase our chances of success, TransCard has strongly invested in the recruitment and genetic analyses of FH index cases with extreme lipid levels and their family members in the first 3 years.

In view of the ambitious objective to find novel genetic origins of human lipid disorders, the consortium organized three additional target identification strategies that ran in parallel. The partners conducted *1)* dedicated *in silico* analyses (NEBION/UMCG; combining the use of unbiased large mRNA datasets with key established insight in plasma lipid regulation), *2)* studies in cultured cells (UZH; identification of regulators of lipoprotein uptake by liver cells and processes controlling lipoprotein transport over the inner lining of the vessel wall), 3) and through studies in mice with abnormal lipid levels (UMCG).

With four lines of research to identify novel targets to treat CVD, the challenge for TransCard was finding the best route for follow-up research to validate newly identified targets. To take the proper and timely translational steps for each of the targets was anticipated to be challenging and this proved indeed to be true. In the following paragraphs, a few examples illustrate the complexity of conducting the best possible work with novel targets that could lead to the development of pharmaceutical drugs.

Prior to the start of TransCard, the AMC had already identified *STAP1* as a novel candidate gene through studies in families with FH. As this finding is close to the heart of TransCard, i.e. the origin of human lipid disorders, REGION-H immediately set out to study the relation between common and rare genetic variants in *STAP1* with plasma lipid levels and risk of CVD. However, the genetic tools at hand were too limited to directly include or exclude *STAP1* as target for further study. In this context, a mouse model was generated in the UMCG which *STAP1* was ablated. These analyses did not support *STAP1* as a FH candidate gene. Upon reanalysis of the families with FH that led to the discovery of *STAP1,* it became clear that *STAP1* has likely been a false lead.

UZH set out to conduct a genome-wide approach study to find targets that limit the uptake of lipoproteins (LDL and HDL) in liver cells. This means that each of the 21.000 (protein coding) genes were studied. This massive research effort took over 3 years, leaving less than 2 years to prioritize and validate the top targets in cell culture experiments, experimental animal models, and in general population cohorts. The remaining time excluded the possibility to execute validation studies prior to the ending of TransCard.

For targets identified through *in silico* analyses, the route to proof that such targets are relevant for human health and disease was expected to be long and bumpy but the eventual route cannot be predicted. For example, a gene that we identified at the start of TransCard was found to be associated with the concentration of cholesterol in blood but only after extensive analyses in general population

cohorts, it turned out that changes in plasma lipid levels did not affect the risk CVD. On the other hand, a mouse model clearly showed interesting effects on the major blood lipids that warrant further investigation. A second gene identified through *in silico* analyses could in contrast not be studied in humans due to the complete absence of genetic variation in public databases as well as databases that were generated by REGION-H and the UCMG. This severely hindered progress despite the notion that a dedicated mouse model showed the relevance of this gene to lipid metabolism.

The above examples illustrate that the key challenge of TransCard has been to validate the role of newly identified targets in human lipid metabolism and risk of CVD. With most novel targets identified half-way the program, this has been the main challenge.

1.2.2 Objectives of the project

The objectives of this project boiled down to :

1. the **identification** of targets that affect the concentration of the main lipids in blood. Here we can distinguish between cholesterol packaged in LDL, cholesterol in HDL and triglycerides that are packaged in other larger lipoproteins. TransCard has mainly focused on the genetics of LDL metabolism (see paragraph 1.1) meaning that we have been searching for new LDL genes as targets. To this purpose, we set out to
   - identify targets through contextual gene expression analyses. The hypothesis here is that (novel) genes that are strongly coregulated with genes with major established roles in lipid metabolism are also relevant. To make this work possible, large publicly available gene expression datasets were to be curated.
   - find novel LDL genes in families with very high or very low LDL cholesterol in which mutations in the known genes were to be excluded. The principle here is that very extreme levels of LDL cholesterol can only be driven through high impact mutations in the DNA and not through lifestyle. To this purpose, index cases and their families were to be recruited and studied.
   - identify genes that control the uptake of LDL (and HDL) in cultured liver cells. The principle here is that changes in the uptake of LDL by the liver will result in changes in the concentration of LDL cholesterol in blood. To this objective, all 21.000 protein coding genes of the human genome were downregulated and the consequence studied.
   - find factors that control the transport of LDL (and HDL) over the vessel wall. These processes may not only directly affect blood lipid levels but also the development of atherosclerosis in the vessel wall. To this purpose, several screens were to be employed.
   - study experimental animal models with increased LDL cholesterol levels. At the start of TransCard, it had become clear that intracellular trafficking of the hepatic LDL receptor was involved. This key insight was exploited in further animal studies and importantly also for study the relevance in human metabolism.

2. the **unravelling the molecular mechanisms** responsible for the association between (newly) identified factors and LDL cholesterol levels. Understanding the mechanism that underlies the observations in cells, animals and humans is key to advance our understanding of the regulation

of blood lipid levels and thereby the risk of CVD. Such basic scientific insight is needed to analyze the potential of newly identified targets to be exploited for the development of new pharmaceutical intervention. This challenging part of the project was to be tackled through experiments in the laboratory with cultured cells, dedicated mouse models to study effect on blood lipids and atherosclerosis, and where possible with patient-derived materials (blood products and cells).

3. the **validation of targets**, an important complex aspect of the work, was to be tackled through conducting a series of steps dependent on how targets were identified. Translation of intriguing possibly relevant findings in a test tube to relevance in humans is a challenge as will be further addressed below.

1.3. Description of the main S&T results/foregrounds

1.3.1 Target identification and prioritization in silico (WP2)

Task 2.1. Digital portal for TransCard members to use Nebion's databases (M1-M6)

NEBION rented, set up and installed a server to provide access to GENEVESTIGATOR for TransCard members. This server was used for initial training of consortium members and for subsequent data analysis.

Task 2.2. Training partner members to use Nebion's databases (M6-M24).

A first training for TransCard members was organized on June 11th, 2014, and took place in the realms of ETH Zurich and NEBION. 18 persons from the TransCard consortium participated. A personal follow-up training was offered on June 12th 2014. The objectives of the training were to introduce TransCard participants to transcriptomics and to the use of GENEVESTIGATOR for target and biomarker discovery.

Throughout the entire duration of the project, NEBION actively supported several participants in their use of the GENEVESTIGATOR platform. It was used primarily for research within TransCard, but also in other projects closely affiliated with TransCard.

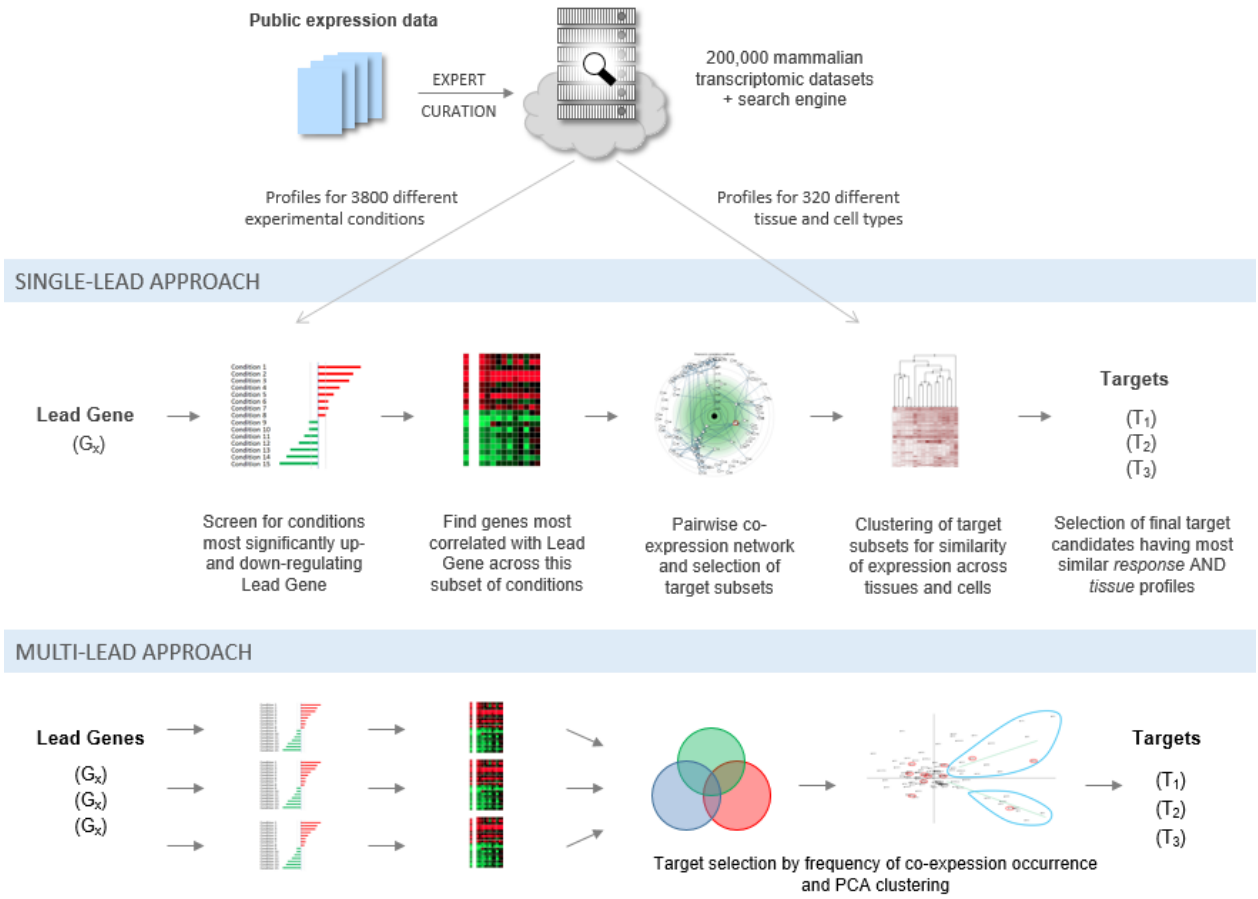
Task 2.3. Comprehensive literature screening of potential targets from existing GWAS studies (M1-M12).

TransCard screened and assembled lists of genome-wide association studies (GWAS) targets from various sources and matched them against other similar compendia, in particular with those of a major publication from the Global Lipids Genetics Consortium (Global Lipids Genetics Consortium, 2013. Discovery and refinement of loci associated with lipid levels. Nature genetics 2013; 45:11, 1274). Lists of relevant genetic associations from this publication were extracted for different scenarios of CVD (loci primarily associated with either HDL cholesterol (36 genes), LDL cholesterol (15 genes), total cholesterol (39 genes), triglycerides (28 genes).

Task 2.4. Identify novel targets by co-regulation network prediction (M1-M24)

Strategy 1

A first target discovery strategy consisted of identifying genes co-regulated with genes with function in human lipid and lipoprotein metabolism. In this approach, we first defined "lead genes" from molecular processes associated with CVD, i.e. triglyceride hydrolysis, anti-oxidation, uptake of oxidized LDL, cholesterol biosynthesis, cholesterol transport, lipid modification and atherosclerotic lesion development. In a second step, we screened for genes co-regulated with our "lead genes". Co-expression analysis was performed selectively, for each gene, across the subset of GENEVESTIGATOR *Perturbations* profiles in which they are significantly regulated. Approximately 100 significant perturbations per lead gene were chosen based on the curated compendium of Affymetrix Human 133 Plus 2 arrays. The approach was originally performed on a single gene level and extended to a multi-gene approach (shown in scheme below) by measuring the frequency of co-expression against a group of seed genes.



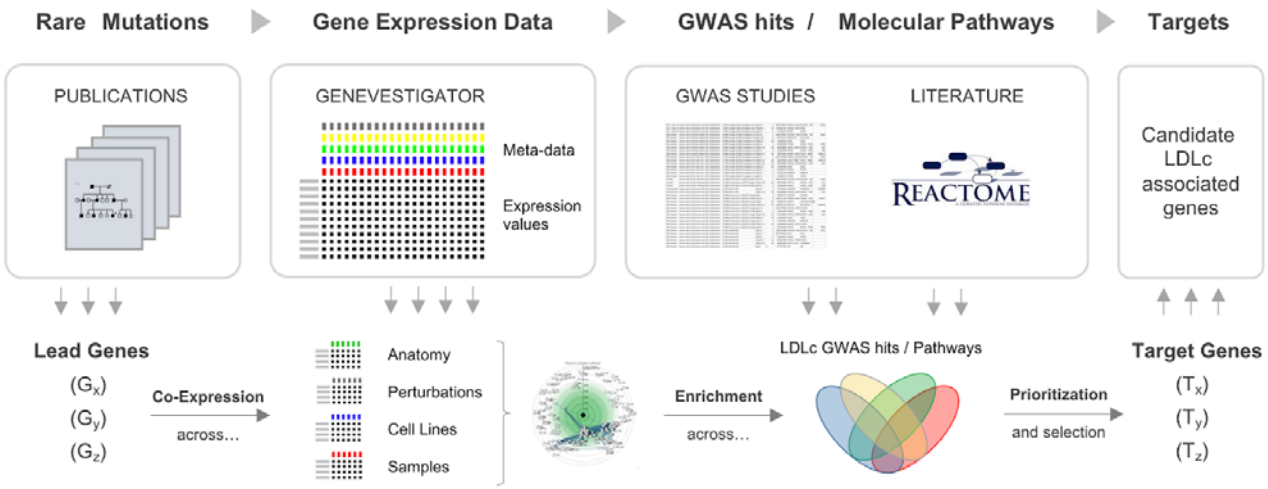
The single-lead target search strategy consists of identifying targets co-expressed with a single lead, while the multi-lead target search integrates results from multiple lead genes to identify targets that are correlated with multiple genes for a given molecular process.

Strategy 2

In this scenario, the search approach is based on molecular function and tissue-specific expression. The hypothesis was that age-accelerated atherosclerosis correlates with failure to upregulate antioxidant genes (Collins et al, 2009; PMID: 19265038). We selected all genes from three families of proteins having anti-oxidant properties - catalases, superoxide dismutases, and peroxidases - potentially associated with atherosclerosis. GENEVESTIGATOR *Anatomy* profiles (283 tissues and cell types) were used to identify genes expressed in vascular tissues. Two genes showed expression mainly in vascular endothelial cells: one of these genes encodes for a largely uncharacterized protein. GENVESTIGATOR results showed that this gene is over-expressed in thrombus-derived leukocytes as compared to peripheral blood leukocytes, as well as up-regulated in ruptured atherosclerotic plaques compared to stable plaques.

Strategy 3

The third strategy, developed in periods 3 and 4 of the project, consisted of
a) defining "lead genes" from literature mining of genetic studies of familial dyslipidemias,
b) running co-expression analysis on aggregated anatomical, perturbational or other data,
c) integrating the results from various profiles, and
d) performing enrichment against GWAS and pathway data for *in silico* validation.



Choice of "lead genes":
The goal of this analysis was to identify new genes co-expressed with "lead genes" known to be involved in monogenic forms of LDL dyslipidemias.

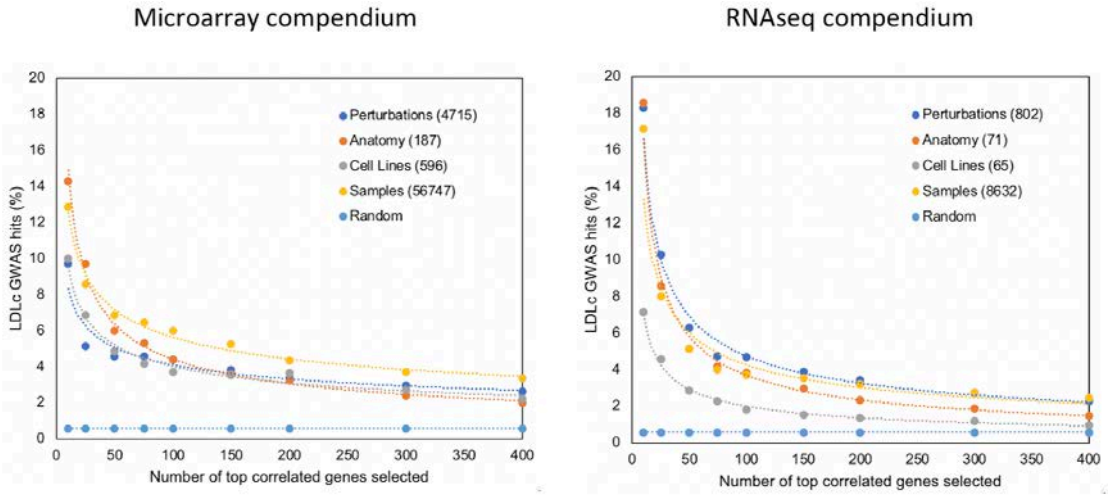
Datasets:
Co-expression analysis was carried out on expression data derived from the GENEVESTIGATOR curated compendium of microarray and RNA-seq data. The set of Affymetrix Human 133 Plus 2 arrays comprised 981 thoroughly curated studies with a total of 56747 samples. From this, aggregated expression profiles were generated containing a selection of 187 different tissue types ("Anatomy"), 596 cell lines ("Cell Lines"), and 4715 perturbational response profiles

("Perturbations"). The RNA-seq dataset comprised 192 studies with 8632 samples. From this, aggregated expression profiles were generated containing 71 different tissue types, 65 cell lines, and 802 perturbational response profiles (of which 115 were cancer profiles).
Analysis: We extended co-expression analysis from strategy 1 by separately performing co-expression by aggregated profile types of *Anatomy*, *Cell Lines*, *Perturbations* and *Samples* - denoted as *contextual co-expression* - and subsequently integrating the results.

Validation:
For the gene sets obtained, to assess the level of association with LDL cholesterol, we calculated the proportion of genes present in our assembly of LDL cholesterol associated GWAS hits. This was done separately for each profile types. Although the coverage and content of studies were different between the microarray and RNA-Seq compendia, no large discrepancies were found between these two technologies.

Results:
The number of LDL cholesterol GWAS genes in the top 50, 200 and 400 genes most correlated with ten "lead genes" across the profile types *Perturbations* and *Anatomy* was assessed. Several genes showed high enrichment of LDL cholesterol GWAS hits in their respective co-expressed genes primarily in the *Perturbations* dataset, but had low enrichments in the co-expression by *Anatomy*, indicating that their anatomical expression is not informative for finding disease association. Moreover, significant differences between lead genes were observed in terms of GWAS hit enrichment, both within a given profile type (*Samples*, *Anatomy*, *Cell Lines* or *Perturbations*), but also between the profile types. For example, genes co-expressed with e.g. our lead genes did not show higher proportions of LDL cholesterol GWAS genes than random gene sets of the same size, suggesting that their association with LDL cholesterol may not be visible at the transcriptional level. We therefore excluded them from further analysis.

The percentage of LDL cholesterol GWAS hits among genes most correlated with the chosen lead genes respectively, showed a higher enrichment for analyses done on the RNA-Seq than on microarray data in the top 10-150 range, and vice versa in the range 150 - 400. In both cases, the proportion of LDL cholesterol GWAS hits follows a power law, independently of the profile type and platform chosen. Up to 18% of co-expressed genes were GWAS hits related to LDL cholesterol.



We then set to derive the overlap of the enrichments found for each profile type (i.e. *Samples*, *Cell Lines*, *Perturbations* and *Anatomy*). Combined, the top 25 most correlated genes for each lead gene across each of the four profile types yielded 451 genes in total for the microarray compendium, and 474 genes for the RNA-Seq compendium. The overlap between each profile type was between 20% and 50%. Co-expressed genes identified in multiple profile types clearly showed higher enrichments in LDL cholesterol GWAS genes than those identified in only one profile type. An example can unfortunately not be shown in this public report.

Task 2.5 - Prioritization of targets by checking their context specificity of expression (M12-M36)

Candidate targets proposed by the consortium, as well as identified *in silico* under task 2.4., were prioritized using GENEVESTIGATOR. In summary, the prioritization consisted of combining gene expression regulation information by tissues and from perturbational data, combined with protein properties, genetic association, and drug-ability scores obtained from various online databases. The prioritization carried out on this list of genes resulted in the identification of 5 candidate genes. For each of these candidates, a detailed characterization of their expression in tissues, cell lines, cancers, drugs, diseases and correlation network, was carried out.

Task 2.6. Curation of additional data to improve prioritization and target identification (M1-M50)

Throughout the project, the curation of public expression datasets was crucial to improve and refine the methods. While at the beginning we used existing microarray compendia already available in GENEVESTIGATOR, our curation efforts helped us include studies performed on RNA-Seq platforms, but also to extend the diversity of experimental conditions to allow a more refined co-expression analysis. Including oncology data improved the enrichment scores in the context of LDL cholesterol. In summary, a large proportion of publicly available data for CVD were curated and integrated into GENEVESTIGATOR. In parallel, to extend the diversity of content, the profiles for drug response, other diseases and tissues were significantly enriched.

The following content was curated by NEBION curators during the project:

Therapeutic area	organism*	# experiments	# samples
Cardiology and angiology	HS, SS, MM, RN, CL	82	3567
Endocrinology and metabolism	HS, SS, MM, RN, CL	134	7944
Gastroenterology and hepathology	HS, SS, MM, RN	72	3505
Oncology and haematology	HS, MM, CL	222	12575
Pneumology	HS, MM, SS	124	7420
Inflammation	HS, MM, RN	85	3770
Toxicology and pharmacology	HS, SS, MM, RN	39	7761
Immunology	HS	36	1502
	TOTAL	**794**	**48‘044**

* HS: Homo sapiens, SS: Sus scrofa; MM: Mus musculus ; RN: Rattus norvegicus ; CL:Canis lupus

Through TransCard, NEBION and UMCG have been able combining the power of state-of-the-art transcriptome bioinformatics (using publicly available datasets that were curated by NEBION over the time of the project) with established knowledge of the biology of human lipid and lipoprotein metabolism. It allowed learning about new genes that are closely associated with genes with major roles in lipid metabolism. This has resulted in the identification of a druggable target which was later on picked up by GWAS and other experimental research groups with which we are currently collaborating. A second interesting target (another protein coding gene) was identified which apart from having a name is not studied by other investigators worldwide (based on all information in the public domain). Studies in experimental mice (WP6) by the UMCG have validated roles for these genes in the regulation of plasma lipid levels and these studies are ongoing. While a patent on applications based on the first gene was already filed at the time of our identification, NEBION and UMCG are currently working on filing a patent on applications that are based on the second gene.

NEBION expects its integration of additional expression data, in particular RNA-sequencing data, will open improved and new opportunities for target discovery in the future. Not only the expression of more protein-coding genes is currently measured with the new technologies, but also non-coding transcripts are taken along which can be used as surrogate markers or constitute companion diagnostics signatures. The extension of NEBION’s tools and database content during TransCard significantly improved *in silico* target discovery and the predictive power of the analysis. Moreover, the development of user-friendly online tools now allows biologists with little bioinformatics training to perform complex and large-scale queries efficiently, without relying on external bioinformatics support. The productive collaboration between NEBION and UMCG has finally laid the basis for continued collaboration beyond TransCard.

1.3.2 Target identification in patients (WP3)

The main role of the AMC in TransCard was to identify patients with extreme levels of blood lipids with a focus on those with extremely high LDL cholesterol. A biobank was established and data has been collected from over 200 patients with different clinical phenotypes as further detailed below.

LDL cholesterol

The AMC mainly focused on collecting data and bio-materials from patients and families with extremely high LDL cholesterol levels to improve our understanding of the factors that control LDL cholesterol but with the ultimate wish to identify new Mendelian disorders of LDL metabolism.

The AMC has a long track record as a clinical center of expertise in FH and it has a large database of patients with extremely high LDL cholesterol levels in whom no mutations in the coding regions of established LDL cholesterol genes (*LDLR, APOB* and *PCSK9*) were found at the start of TransCard. This clinical entity is commonly referred to as FH4, i.e. FH caused by mutations in a fourth unknown gene. FH4 patients with LDL cholesterol concentrations above the 99th percentile for age and sex were recruited, blood samples, DNA, and clinical information was collected. If possible, also family members were included in these recruitment efforts. The AMC used whole genome and exome sequencing combined with classical linkage analysis to study LDL (candidate) genes. Altogether, these efforts did not lead to the identification of novel candidate genes and the research was focused on the effects of genetic variation in *STAP1*, *LIPA*, *ABCG5/G8* on LDL metabolism. In addition, we studied the functionality of an intronic *LDLR* variant in families with thus far unexplained severe hypercholesterolemia. The AMC efforts and those of REGION-H and the UMCG to validate *STAP1* as a candidate gene can be found under paragraph 1.3.5

Lessons learned

Regression to the mean. The LDL studies conducted as part of TransCard have first of all taught that it is crucial to thoroughly re-evaluate initial lipid phenotypes. In almost half of the patients who initially presented with a very prominent FH4 phenotype (defined as LDL cholesterol > 7 mmol/L, which is far above the 99th percentile for age and sex) LDL cholesterol levels were below 6.1 mmol/L (far above the 95th percentile) at the second measurement that were carried out as a specific study for TransCard.

Polygenic FH. Over the last few years, several studies provided evidence for a polygenic form of FH caused by the combined effect of a number of common - LDL cholesterol increasing - genetic variants quantified through a so-called gene risk score. In eight FH4 families, it was found that high LDL cholesterol in affected individuals could not be distinguished from unaffected family members (with normal LDL cholesterol) with such scores. In addition, established polygenic forms of FH were not present in all FH4 patients. These findings suggest that FH4 does not appear to have a polygenic origin in the AMC patient cohort.

Roles for additional non-canonical lipid genes. The AMC studies also showed that a thorough evaluation of established lipid genes is warranted. Mutations in *LIPA* have previously been described to affect HDL cholesterol and triglyceride levels. However, in one of the promising FH4 families, the AMC showed that homozygosity for a previously described mutation in *LIPA* resulted in extreme LDL cholesterol levels with surprisingly no effect on HDL cholesterol and TG levels.

More recently, the AMC also explored *ABCG5* and *ABCG8* as LDL candidate genes after several publications showed that carriers of mutations in these genes are characterized by high LDL cholesterol levels. These studies were, however, carried out in small numbers of patients. The AMC studied whether mutations that were predicted to be deleterious co-segregated with increased levels

of LDL cholesterol in multiple large families. The results do not support a role for *ABCG5* or *ABCG8* in clinical FH.

An unanticipated role for canonical gene loci. Another striking insight was obtained through whole genome sequencing studies in two FH4 families of which the pedigrees are shown in **Figure 1**.

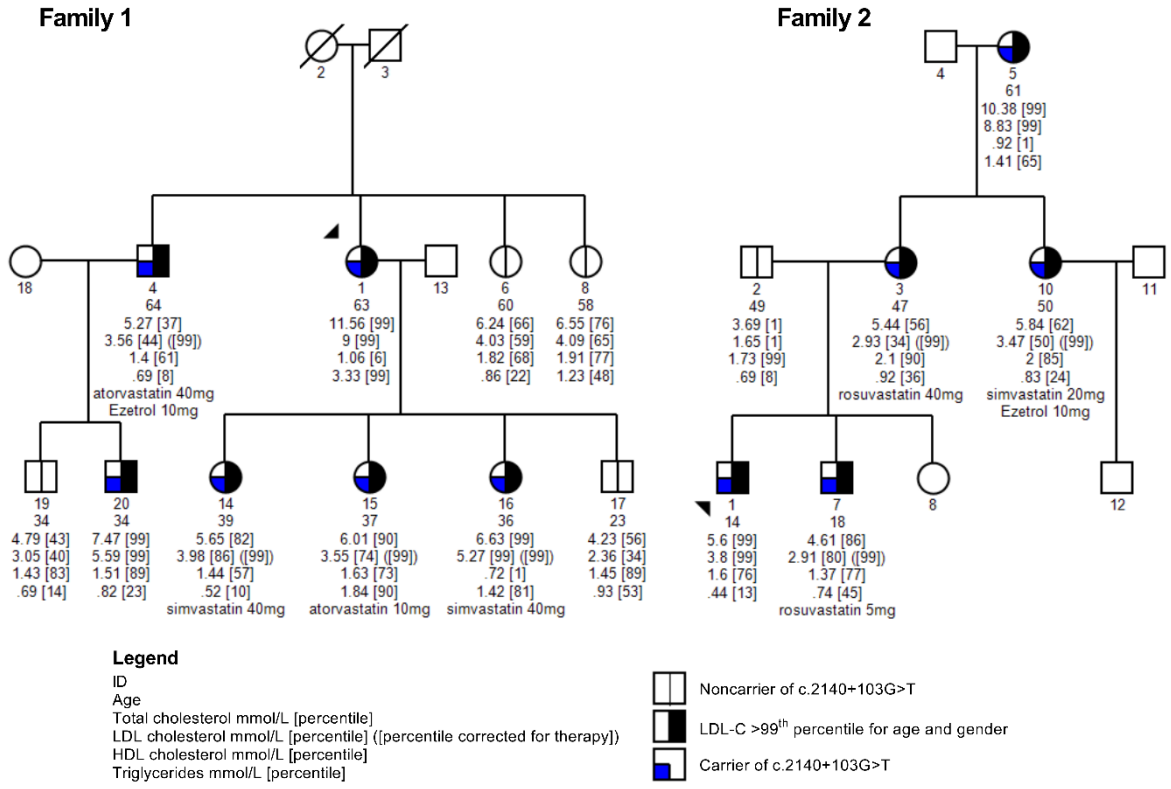


***Figure 1**. Two FH4 families turning out suffer from FH1, i.e. a functional mutation in the LDLR gene.*

In the first family, we identified a deep intronic variant in LDLR in 5 affected, but not in 4 unaffected individuals. This genetic variant thus does thus not directly change the encoded LDLR protein. The same variant was identified in the second FH4 family. The variant was shown to cause on aberrant mRNA splicing which resulted in a 97 base pairs intronic insertion leading to a frameshift and premature stop codon in exon 15. Further studies revealed a prevalence of this variant of 0.24% (8/3299) in patients referred for FH mutation screening. Co-segregation analysis in the second family showed complete penetrance of this variant with the FH phenotype over 3 generations.

These results suggest a need to analyze the intronic regions in addition to the protein coding regions of the *LDLR* locus in patients with FH. Following this intriguing finding, the AMC embarked on finding additional intronic variants in their FH4 cohort. In these ongoing studies, there will be a focus on the impact of combinations of intronic variants on plasma LDL cholesterol concentrations.

Premature atherosclerosis

The AMC also studied a highly interesting population of patients without plasma lipid disorders who nevertheless suffered from premature atherosclerosis. This was defined as having a cardiovascular

endpoint before the age of 65 for women and 55 for men without classical risk factors. In families with an autosomal dominant phenotype for premature atherosclerosis, the AMC analyzed exome sequencing data. In one of the pedigrees, the exome data were combined with classical linkage analysis and this led to the identification of a unique variant in *SUSD2* resulting in a premature stop which fully segregated with the phenotype in the family as can be seen in **Figure 2**.

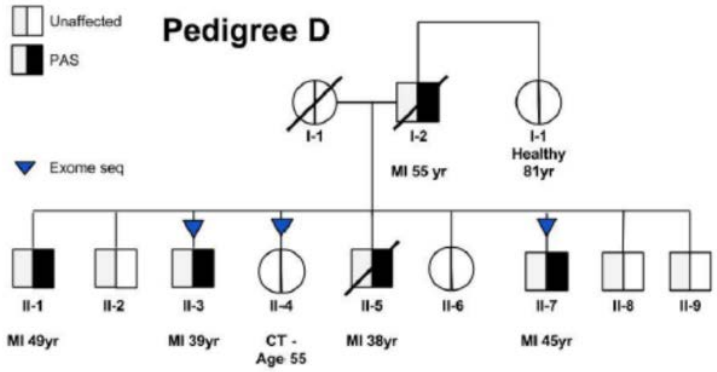


***Figure 2**. PAS pedigree, where heterozygosity for SUSD2 mutation co-segregates with severe premature atherosclerosis. MI, myocardial infection.*

Additional studies indicated that *SUSD2* is expressed in healthy human aortic specimens while cell tissue culture experiment suggested that it is highly expressed in pericytes. In collaboration, REGION-H explored the association of two common variants in the *SUSD2* locus with CVD outcome. It was found that the rare allele of rs8141797 was significantly associated with a lower risk for CVD which support the idea that *SUSD2* is a candidate gene for premature atherosclerosis but the mechanism remains to be resolved.

HDL cholesterol

Because HDL cholesterol levels are inversely associated with CVD risk, the AMC initially also set out to also identify novel genes in HDL metabolism. This effort was, however, halted prematurely in the light of further failure of HDL cholesterol raising drugs combined with the outcome of Mendelian Randomization showing that increased HDL cholesterol is not causally related to CVD risk. The AMC stopped further collection of data after the finalization of studies on several HDL candidate genes, i.e. *OSBPL1a* and *LRP1* (in collaboration with UMCG).

Technical advances

The AMC furthermore assessed the lipidome using in house developed methods in 16 FH4 patients and 5 unaffected controls from the same families. Lipidome analyses entails detailed analysis of lipids using a high-performance liquid chromatography-mass spectrometry. We identified 46 lipids (most notably phosphatidylcholine, lysophosphatidylcholine and sphingomyelin) that were differentially abundant between FH4 and normolipidemic subjects. Further studies are carried out to evaluate the differences in lipidome between normolipidemic controls, patients with mutations in

true FH genes and FH4 patients. These studies are anticipated to help better differentiate affected and unaffected family members preceding subsequent genetic analysis, while it may also help with pathways that are involved in changes in the lipidome.

The AMC finally optimized cholesterol uptake studies in CHO and HEPG2 cell culture models which are now sensitive enough to discriminate between heterozygotes for LDLR mutations and unaffected controls.

In conclusion, thorough reassessment of LDL cholesterol measurements, extended next generation sequencing approaches (covering non-coding DNA) to exclude causal defects in the canonical genes, exclusion of other causal defects in non-canonical lipid genes, and technical advances will be needed to find novel genes in families with extreme hypercholesterolemia.

1.3.3 Target identification and prioritization in vitro (WP4)

RNAi screening for genes limiting hepatic lipoprotein uptake

Plasma lipoprotein levels are strongly associated with the risk of developing atherosclerotic cardiovascular disease (ASCVD). The uptake of circulating lipoproteins by the liver is the most important step in regulating their plasma levels and in mediating some of their biological effects. Thus, identifying novel mediators of HDL and LDL endocytosis in the liver would allow to potentially discover new drug targets for the prevention and the treatment of atherosclerosis. To discover novel genes involved in holoparticle uptake of LDL and HDL, UZH performed a microscopy-based genome-wide siRNA screen. UZH tested the effect of approximately 68000 different siRNAs (3 siRNAs per gene) on the uptake of LDL and HDL fluorescently labelled in their protein moiety (fl-HDL) by the human hepatocarcinoma cell line Huh-7. Briefly, 72 hours after siRNA transfection, the cells were exposed to fl-HDL for 4 hours, followed by fixation in paraformaldehyde. The cells were then imaged with an automated wide field fluorescence microscope. The images were segmented to identify the nuclei and the LDL or HDL-containing vesicles, and features related to the vesicles (e.g. vesicle intensity and cytoplasm granularity) were measured automatically. This image analysis phase was followed by a data analysis phase aimed to identify the genes that, when knocked down, displayed the largest effect on LDL or HDL holoparticle uptake. **Figure 1** summarizes the screening process for HDL.

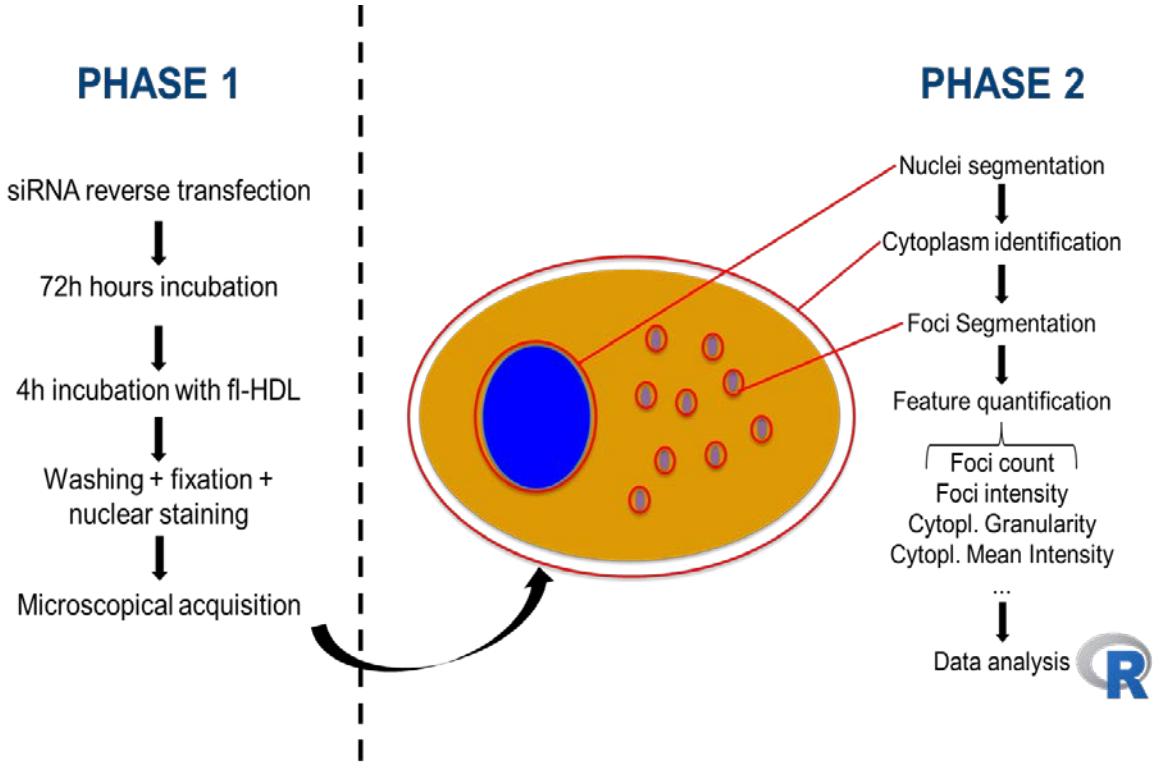


Figure 1. siRNA screening process*. Briefly, the cells were transfected with 3 siRNAs against each gene in the human genome. 72 hours after transfection, the cells were exposed to fluorescent HDL for 4 hours, followed by washing, fixation, and staining of the nuclei with Hoechst33258. The wells were then imaged with two automated wide field microscopes. The nuclei, the cytoplasms and the vesicles were segmented by applying an automated algorithm and features related to the vesicles (named 'foci' after segmentation) and to the cytoplasms. Data analysis was performed with the R statistical software.*

Transfection efficiency and assay windows were measured and used in the quality control phase. To determine transfection efficiency, four wells in each of the 192 assay plates were spotted with an siRNA against PLK1, a housekeeping gene which is necessary for cell survival. The loss of cells in PLK1-transfected wells was then determined by counting the nuclei and comparing the nuclei count to control wells. The assay window was estimated by measuring the signal intensity from the HDL-containing vesicles in wells that received fl-LDL or fl-HDL and in wells that did not. For both transfection efficiency and assay signal intensity, the quality of the assay was expressed as Z'-factor. Both the transfection efficiency throughout the screening and the assay windows between negative and positive controls were excellent (yielding Z'-factors constantly above 0.5), giving confidence in the quality of the hits that were generated.

After the quality control step, the data for each feature were normalized by batch, microscope and plate, followed by robust Z-score normalization. The redundant siRNA activity test (RSA) was then applied to rank the siRNA molecules based on their experimental effect and to determine whether the three siRNAs against each target gene were distributed significantly higher in the ranking than would be expected by chance. As the RSA analysis is directional, this generated four hit lists for each assay feature, namely for decreasing or increasing the uptake of HDL or LDL. The screen was optimized towards the identification of genes whose knockdown decreases lipoprotein uptake. Median cytoplasm intensity was the feature with the highest SNR for both the HDL and LDL dataset and

thus, not surprisingly, generated the highest number of highly significant gene hits. Although median cytoplasm intensity appeared clearly to be the best assay feature, additional gene hits were searched by applying different dimensionality reduction strategies to the whole dataset. A high degree of correlation was present across all features and between the HDL and LDL datasets, both when assay scores and RSA p-values were taken into consideration. Considering the aforementioned high degree of correlation between the HDL and LDL datasets, not surprisingly these genes were ranking very high in the LDL dataset too (**Figure 2A**).

To appreciate the pathways and the multiprotein complexes that this kind of screenings are able to unveil, UZH performed functional clustering of the top gene hits using the STRING online tool. As shown in **Figure 2B**, four major hit clusters were identified in both the HDL and the LDL datasets. Similar information was obtained by applying a gene ontology analysis to the top hits (data not shown). **Figures 2C and 2D** show the distribution of the assay scores for the top hit genes within the HDL and LDL dataset (RSA p-value cut off of 0.001).

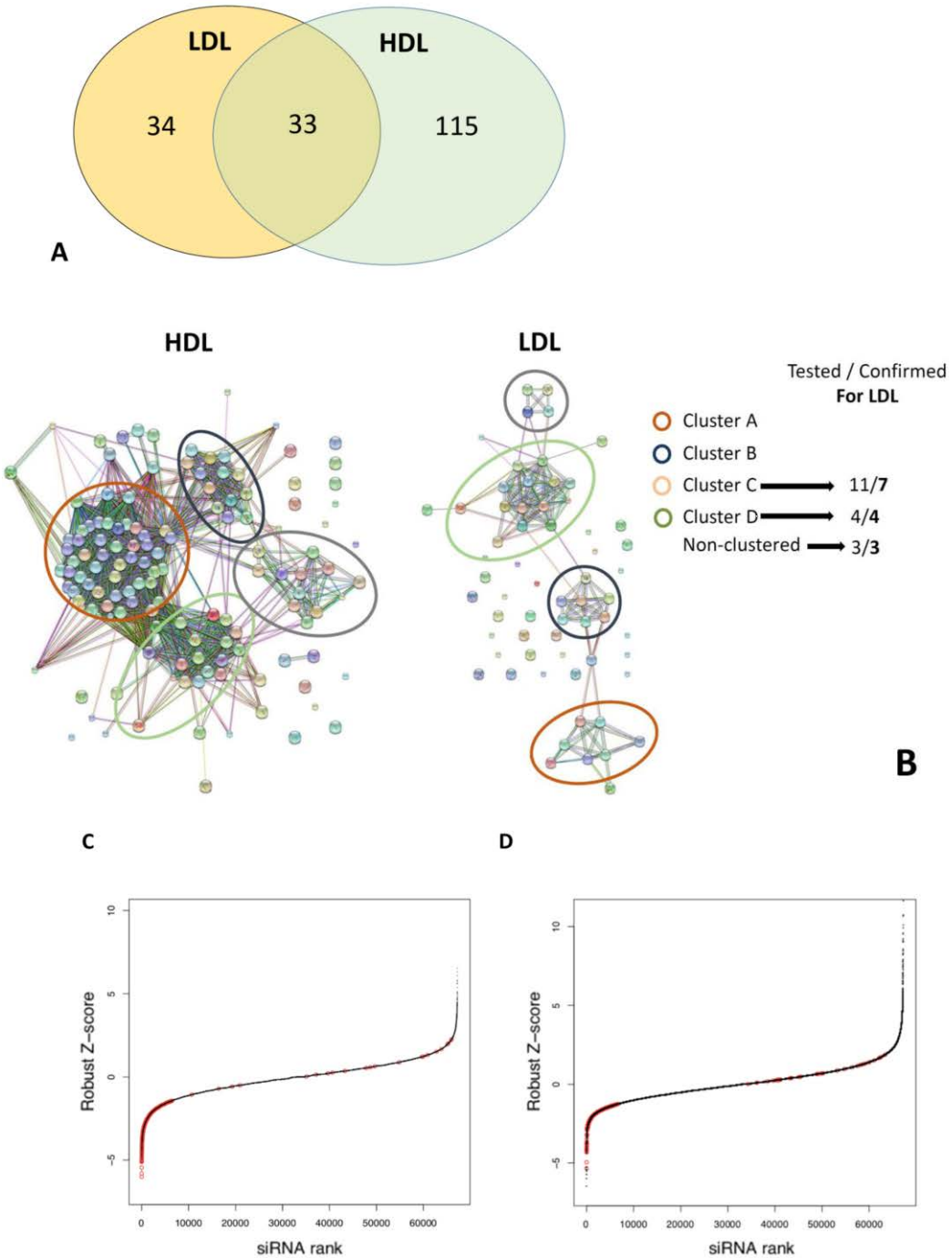


Figure 2. Functional clustering of the main screening hits*. A. Numbers of hit genes for LDL and HDL for the Median Cytoplasm Median Intensity feature, p-value cut off of 0.001, no further filtering done. B. Clustering of the top HDL and LDL hits for the median cytoplasm intensity feature according to the STRING online database. For both figures, a p-value cut off of 0.001 was applied to the RSA analysis for the median cytoplasm intensity feature to select top hits. Each oval represents a functional cluster of genes. The number of hits that were tested for each cluster and the confirmation rate are reported on the side of each graph (see 1.3.5). C and D. Distribution of the screening assay scores. Each dot represents the robust Z-score for the median cytoplasm intensity feature for HDL (C) and LDL (D) for one siRNA against a target gene. siRNAs targeting genes with an RSA p-value lower than 0.001 displayed as red circles. Control siRNAs were not included in the graph. A few outliers that induced extremely high scores were excluded to allow for better scaling.*

The raw siRNA scores of each feature showed a quite skewed and non-Gaussian frequency distribution (**Figure 3A**). This implies that regular principal component analysis (PCA) is not the best strategy to perform dimensionality reduction and that non-linear approaches have to be preferred. After comparing multiple alternatives, the data were first $\log_2$-normalized (**Figure 3B**) for subsequent dimensionality reduction using the Locally-Linear-Embedding method. The strong correlation between the two datasets (**Figure 2c**) was maintained also after dimensionality reduction, indicating the independence from the assay feature selected. In fact, the top HDL hits selected for validation, clustered on a protrusion located on the C1 axis of both the HDL and the LDL dataset and clearly separated from the rest of the datasets. On the one hand, this indicates that we indeed singled out for validation from the bulk of the dataset the genes that display the highest difference in lipoprotein uptake. On the other hand, when considering the results of the validation experiments (see 1.3.5) it may indicate that these hits are mostly LDL-related genes that somehow have an effect on the uptake of fluorescent HDL too. Interestingly enough though, another group of genes on the C2 axis of the HDL dataset was identified through this strategy. They represent a group of promising HDL-specific genes that were not discernible from the bulk of the dataset prior to the application of a dimensionality reduction.

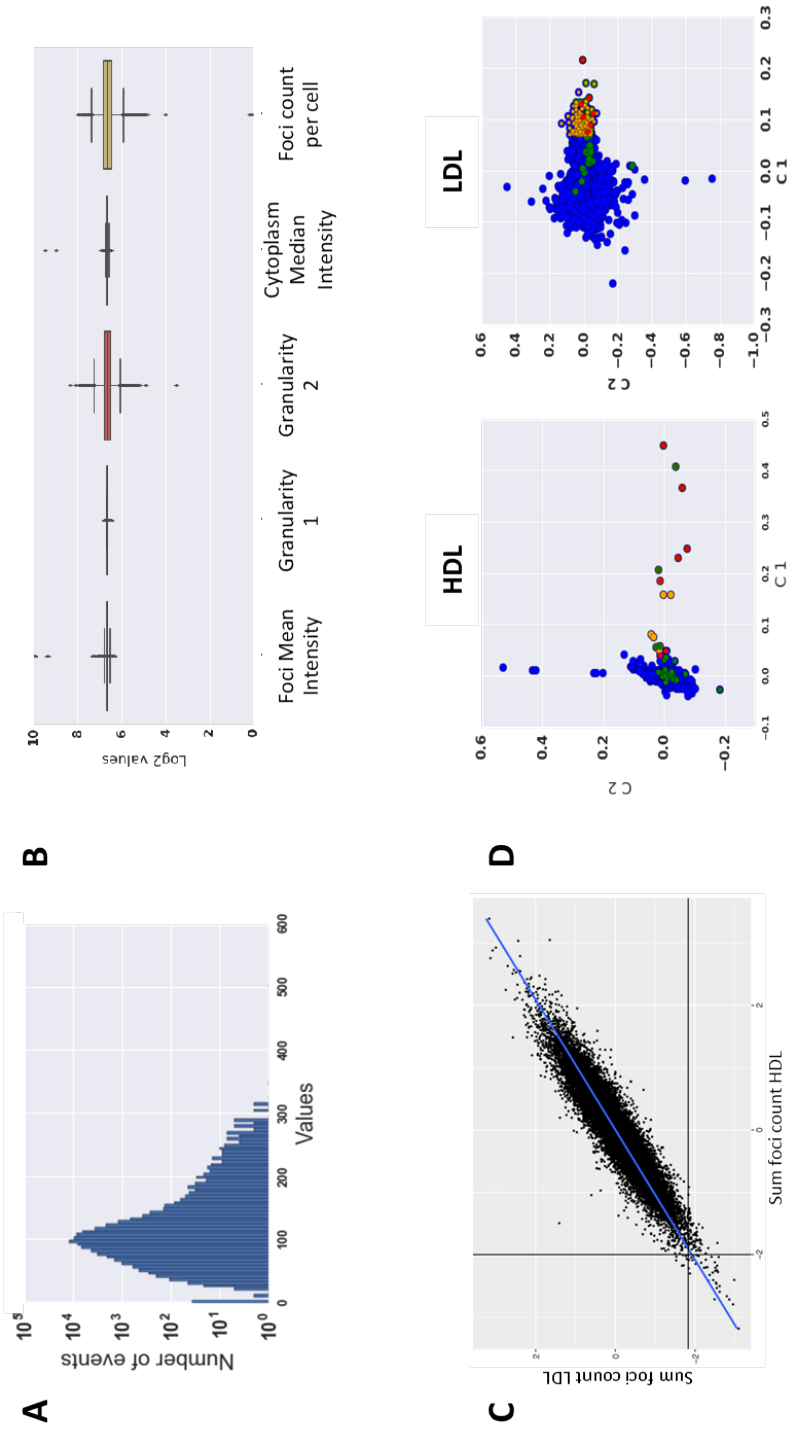


Figure 3. Dimensionality reduction approaches. A. Normality of the dataset. *The graph shows the frequency distribution of the median foci count data points for the HDL dataset according to their value. For multiple assay features, the distribution of the data was skewed. B. Log2 normalized data. The boxplot displays the data for 5 assay features (including the one shown in panel A) in the HDL dataset after log2-normalization. Despite the presence of outliers, the data are now normally distributed and can be used for dimensionality reduction. C. Correlation between the HDL and LDL datasets. Each dot in the scatterplot represents the median assay score for the sum foci count feature in the HDL (x axis) and LDL (y axis) channels. The regression line for a linear model is shown in blue. The black lines are at x = -2 for HDL and at the intersection of the regression line with x = -2 for LDL, in an attempt to identify genes that would affect only one lipoprotein subclass. Similar results were obtained with the other assay features taken into consideration. D. Location of the HDL top hits after non linear PCA. The graphs show the first two components of the HDL (right) and LDL (left) datasets after Local-linear-Embedding dimensionality reduction was applied. Each dot represents one gene. Red dots represents the top HDL hits that were selected for in vitro validation. The top genes in each C1 component were selected by choosing an arbitrary cutoff of 0.5 and are depicted in yellow. Note how the red dots all map at the extreme of C1 in both datasets. The HDL dataset presents also a group of outlier genes on the C2 axis that may represent HDL-specific hits that did not affect LDL uptake. Genes that were not expressed in Huh-7 cells according to RNA sequencing data were excluded from this figure.*

Drug inhibitor screening of kinases limiting uptake and transport of lipoproteins by endothelial cells

The accumulation of LDL in the subendothelial matrix and macrophages contributes to the pathogenesis of atherosclerosis. Conversely, removal of cholesterol from the subendothelial space by cholesterol efflux and the subsequent initiation of reverse cholesterol transport (RCT) have been postulated to confer protection against atherosclerosis. To reach the subendothelial space, both LDL and HDL have to cross the intact endothelial barrier. It is yet controversial whether transendothelial HDL transport is mediated by specific mechanisms or the result of passive filtration. Previous *in vitro* data from UZH and other investigators showed that endothelial cells internalize lipoproteins via a non-classical endocytic route involving dynamin and cytoskeletal networks. Many studies have shown that endocytosis is subject to regulatory control. Large groups of kinases were in fact identified that couple the endocytic transport with signal transduction. Since the initial observations that kinases regulate endocytosis, their role in lipoprotein endocytosis has in fact been lacking. Therefore, to identify signaling cascades regulating uptake of LDL and HDL, UZH performed a microscopy-based high-content screening on human aortic endothelial cells using a kinase inhibitor drug library.

141 kinase inhibiting drugs were screened at seven different concentrations for their effects on the cellular uptake of fluorescent-labeled Atto 594-HDL or Atto 594-LDL by HAECs. Statistical evaluation was done by Fisher exact test to analyze the probability of finding targeting-drugs that actively decrease the uptake of fluorescent HDL or LDL in the presence of maximal concentration versus non-targeting drugs. After correction of p values for multiple statistical testing, VEGFR was identified as the only target that decreased HDL uptake compared to all other targets. No drug treatment affected LDL uptake, indicating a specific role of VEGFR in HDL uptake. The validation of this finding in endothelial cells and clear cell renal carcinoma cells proved the findings (see 1.3.5).

1.3.4 Validation of targets in human populations (WP5)

REGION-H's main role in TransCard was to validate the role of carefully selected targets of importance for the regulation of blood lipids and thereby the development of CVD. This was achieved through performing DNA sequencing and Mendelian randomization studies in large-scale population cohorts.

Validation of the first candidate genes

In the starting phase, REGIONH has established a platform for in-house next generation sequencing (NGS) based DNA diagnostics similar to the platform used at the Department of Genetics of the UMCG. In parallel, they developed a gene panel for targeted sequencing of genes with established roles in lipid metabolism (monogenic disorders & GWAS hits) and 5 genes that were studied by TransCard partners prior and at the initiation of the project.

The first set of TransCard candidate genes included an interesting candidate that was identified through a first series of *in silico* analyses (NEBION/UMCG; WP2), a gene identified in a family with

FH (AMC; WP3), a gene identified through literature studies (UMCG) and a gene known to affect lipid levels in mice (UMCG; WP6).

A total of 64 candidate genes (including the set of 5 TransCard candidate genes) with anticipated effect on

- *(premature) atherosclerosis*
  - *identified through exome/cytoSNP array family studies (AMC)*
  - *identified through bioinformatics analysis of 79 index patients with PAS (AMC-UMCG)*
  - *in vitro and animal studies – genes with roles in inflammation/bile acid metabolism*
- *or total cholesterol and/or LDL cholesterol phenotype*
  - *plasma cholesterol levels as readout*
  - *intracellular cholesterol metabolism as readout*
- *or triglycerides and/or HDLc phenotype*

was used for sequencing the DNA of participants of the CCHS cohort with the top 2% highest (n=200) and lowest of LDL cholesterol (n=200).

The initial assessment of the 5 target genes in the human population samples has been conducted using the following strategy:

- search for non-synonymous, missense and frameshift variants in publicly available databases such as the Exome Variant Server;
- use of *in silico* prediction tools to predict the potential impact of these variants. Only variants predicted to be deleterious with all prediction tools were studied further;
- selected variants should have a minor allele frequency > 0.5% in order to ensure sufficient statistical power for genotyping in the CCHS;
- evaluate the associations between the genetic variants of interest with plasma lipid and lipoprotein levels.

In three out of the first five candidate genes, the initial data were promising and a second validation step was taken to study the association of genetic variation with clinical endpoints of ischemic heart disease and/or myocardial infarction.

<u>Validation of a newly identified candidate genes</u>

From the third year onwards, REGION-H started with the validation of novel candidate genes that were identified through *in silico* studies (WP2), studies in patients (WP3), *in vitro* experiments (WP4) and through additional studies in mice (WP6).

These validation studies were based on natural genetic variation that could be identified in the respective candidate genes in large public databases for which the above described assessment criteria were used. Variants of interest were first genotyped in CCHS and in some cases also in the Copenhagen General Population Study (CGPS). In addition, REGION-H could in some cases make use of their exome variant data for part of their cohort studies.

Actual resequencing of all the newly identified candidate genes has been organized by the coordinator using a second dedicated UMCG gene panel. For these studies, TransCard had access to participants of the prospective Lifelines study (www.lifelines.nl) as well as patients suffering from FH of unknown etiology (AMC). This validation work was started at the end of the third year and included e.g. the top hits of the *in vitro* studies of UZH (WP4). In these studies, over 700 individuals (primarily suffering from FH) have by now been analyzed. This effort will continue beyond TransCard.

In the last phase of the project, REGION-H focused on the most promising candidate genes. These studies e.g. encompassed a new gene identified by the UMCG that holds great promise toward the translation of initial findings in experimental mice to relevance to human lipid metabolism and risk of CVD. In several instances, meta-analyses with data from publicly available databases were carried out. Here, REGION-H estimated the relative effect of the genes on lipid phenotypes and on ischemic endpoints.

REGION-H has over the course of the project assessed and validated a total of 20 genes for their effects on plasma lipid phenotypes and disease endpoints relevant to the project. These human validation studies have helped TransCard to prioritize further experimental studies. Over the last 2 years, REGIONH has furthermore developed a method to help the prioritization of candidate genes. This method can be used to quickly generate a complete overview of the impact of all variants that were identified in all 64 genes that were sequenced in the second year (similar to Manhattan plots that are generally used in GWAS). As indicated, the UMCG has with their gene panel (n=135) generated data of over 700 individuals. The same method is expected to help the prioritization of subsequent studies in human population samples.

Data inferred from the studied genetic variants and associations with phenotypes and disease endpoints in our human populations is expected to be of a high value, and is a robust source of information for the pharmaceutical companies in drug target studies.

1.3.5 Validation & characterization of existing targets and those identified in WP2, 3 and 4 (WP6)

Validation of genes identified through computational (*in silico*) analyses (WP2)

In the first year, NEBION and UMCG identified an interesting target gene with a potential role in triglyceride metabolism. Whole body knockout mice characterized in the UMCG were found to exhibit lower triglycerides but also reductions in plasma cholesterol. The early identification of this gene allowed for a timely involvement of REGION-H to study the relevance of this gene for the human lipid metabolism and risk of CVD. Relatively common variation in this gene was found to affect total cholesterol and apoB in Copenhagen City Heart Study (CCHS). Extending the study to the Copenhagen General Population Study (CGPS) however showed only mild effects on plasma lipids and no effect on cardiovascular outcome. On the other hand, experimental mouse studies showed promising results and basic studies are ongoing to understand the mechanisms that underlie the changes in plasma lipids.

In the third year a newly developed strategy (section 1.3.1) led to the identification of another gene on which there is no data available in the public domain. Unfortunately, the gene locus is not polymorphic meaning that are no direct tools to study the association between genetic variants and blood lipids in human populations (REGION-H). In this light, whole body knockout mice have been generated by the UMCG and the first characterizations show an interesting plasma lipid phenotype. The gene is part of the gene panel that has been developed by the UMCG (see section 1.3.4) and a several very rare variants were identified that are being followed up in families as well as experimental studies in the laboratory. This intriguing topic will be further studied beyond TransCard.

Validation of genes identified through studies in families (WP3)

Prior to the start of TransCard, the AMC reported that variants in *STAP1* are associated with hypercholesterolemia. As such, *STAP1* was considered a genuine FH4 gene following LDLR (FH1), APOB (FH2), and PCSK9 (FH3). The mechanisms by which a mutation in STAP1 might affect LDL cholesterol has been a topic of study in TransCard in which three of its partners collaborated (AMC, UMCG, REGION-H).

This gene was studied by REGION-H through DNA sequencing of individuals with very low or very high LDL CHOLESTEROL (in their CCHS cohort). The genetic tools to study this gene for a role in lipid metabolism were limited but the results did not look promising.

In parallel, the UMCG generated whole body knockout mice to understand how this gene may affect concentrations of lipids in blood. In agreement with the human data of REGION-H, the results were negative. In addition, one of the *STAP1* variants identified in the AMC was also found in a cohort with very low LDL cholesterol.

The AMC rebled the five families in which *STAP1* variants were initially detected through re-bleeds. In 39 STAP1 variant carriers and 71 non-carriers, the AMC could not validate the previously reported difference in LDL cholesterol concentrations (see **Table 1**). This unanticipated finding clearly underlined the absolute need of repetitive lipid measurements, as spurious associations may arise based on a single measurement as was the case here.

Table 1. *Plasma lipids in STAP1 variant carriers and their family members without a STAP1 variant. Values are averages ± SD. TC, total cholesterol; TG, triglycerides;* [1]*, LDL CHOLESTEROL corrected for statin use.*

	STAP1 Non carriers	STAP1 variant carriers			
		All	p.Glu97Asp	p.Ile71Thr	p.Leu69Ser
No. of subjects	71	39	18	7	14
Male (%)	46	49	56	43	43
TC (mM)	5.5 ± 0.9	5.7 ± 1.3	5.5 ± 1.5	5.5 ± 0.5	6.0 ± 1.4
LDL CHOLESTEROL (mM)	3.5 ± 0.8	3.6 ± 1.1	3.5 ± 1.3	3.5 ± 0.5	3.8 ± 1.1
LDL CHOLESTEROL corr. (mM) [1]	3.9 ± 1.2	3.6 ± 1.7	3.9 ± 1.6	3.7 ± 0.5	3.3 ± 2.2
HDL-C (mM)	1.3 ± 0.3	1.3 ± 0.2	1.3 ± 0.2	1.3 ± 0.2	1.3 ± 0.3
TG (mM)	1.5 ± 0.6	1.7 ± 1.0	1.7 ± 1.1	1.5 ± 0.8	1.8 ± 1.0
Lp(a) (mg/dl)	25 ± 35	37 ± 42	26 ± 291	101 ± 51	19 ± 12

In parallel with remeasuring lipids, the AMC also conducted mechanistic and cell-type specific studies. Since *STAP1* is predominantly expressed in B-lymphocytes, the AMC determined B-lymphocyte subpopulations, *i.e.* the relative amount of the CD19-positive cells by FACS analysis of 10 individuals with STAP1 variants and 10 family controls. These data showed no marked differences between groups. In addition, we also determined the effect of isolated peripheral blood mononuclear cells of STAP1 p.Leu69Ser and p.Glu97Asp variants on LDL metabolism in the human hepatoma cell-line HepG2. The cells of the carriers of the *STAP1* variants did not affect LDL uptake or LDLR membrane protein in these cell culture experiments. In conclusion, mutations in *STAP1* are not associated with hypercholesterolemia or changes in LDL related process in cultured cells. The outcome of these negative studies are being prepared for publication.

Validation of novel genes identified through *in vitro* (cultured cells) studies (WP4)

Genes identified by RNAi screening to regulate HDL or LDL uptake by hepatocytes

One of the major achievements of UZH was the validation of the top hits of an image-based RNA interference screening for genes that limit the uptake of HDL and LDL by hepatocytes. The RNAi-screen of HDL and LDL uptake by hepatocytes revealed 152 genes regulating HDL uptake and 83 genes regulating LDL uptake (see 1.3.3). Replication experiments of 20 top hits confirmed their limiting effect on LDL uptake but not on HDL uptake. Therefore, the subsequent functional validation in vitro, mice and humans has been focused on LDL uptake.

Most of the genes limiting LDL uptake were clustered into four functional groups of which two appeared the most interesting. Through combining knock-down with RNA sequencing, a first group of genes was identified that affect the production of LDL receptors. This was validated in human liver and cells circulating in blood. The expression of one of the genes correlates with LDL cholesterol and age in blood cells (UMCG). Knock-down of second set of genes limits the abundance of LDLR on the cell surface and LDL-uptake by hepatocytes. Two of the second set of genes were ablated in the liver of mice using somatic CRISPR/Cas9-mediated gene editing by the UMCG. However, the efficiency of gene ablation in the mouse livers was unfortunately too low to observe a phenotype. Thus, follow up studies are needed to assess the role of these genes in regulating the functioning of LDLR *in vivo*. In the population, some common genetic variants were found associated with LDL cholesterol (REGIONH). Using a gene panel for targeted sequencing of novel candidate genes that has been developed by the UMCG, several patients with severe (familial) hypercholesterolemia carry mutations in the respective genes and family studies are ongoing to show genotype-phenotype segregation (AMC/UMCG).

Regulation of transendothelial lipoprotein transport by VEGF and sphingosine-1-phosphate: Implications for atherosclerosis and kidney cancer

According to the “response-to-injury” theory of atherosclerosis, the accumulation of LDL in the arterial wall plays a pivotal role in the pathogenesis of atherosclerosis. HDL enter and leave the arterial wall to induce cholesterol efflux from lipid-laden macrophages and mediate the first step of the reverse transport of cholesterol pathway to the liver. The accumulation of both LDL and HDL observed in atherosclerotic arteries can thus be seen as the result of increased influx relative to

efflux. Apart from plasma concentration, the determinants of lipoprotein accumulation in the arterial wall are little understood. Especially the mechanism by which LDL and HDL transverse into and out of the artery wall has thus far received little attention. In this light, UZH set out to identify and validate the molecular mechanisms that control the trafficking of LDL and HDL through the intact endothelium. By hypothesis-based and candidate-driven research as well as a drug-inhibitor screen (section 1.3.3), UZH has identified signaling molecules that control the uptake and transcytosis of LDL or HDL by endothelial cells as illustrated in **Figure 4**.

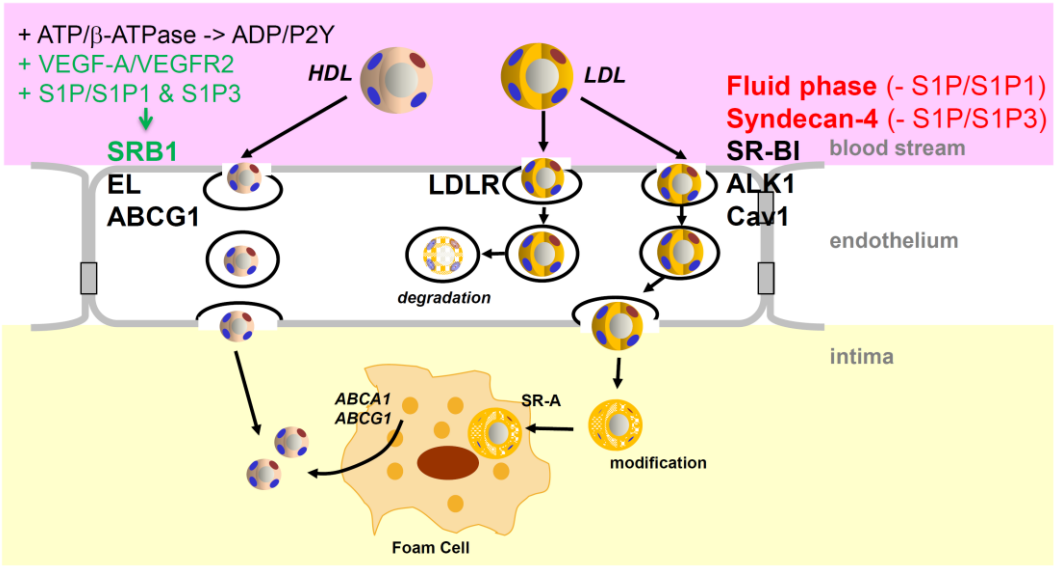


Figure 4. *Working model on the regulation of transendothelial lipoprotein transport. ABCA1 and ABCG1 = ATP binding cassette transporters A1 and G1, respectively; ADP = Adenosine-diphosphate, ALK1 = activin like knase 1; ATP =Adenosine triphosphate; cav1 = caveolin 1; EL = endothelial lipase; LDLR = LDL receptor; P2Y = purnergic receptor; S1P = sphingosine-1-phosphate; S1P1 = S1P receptor 1; S1P3 = S1P receptor 3; SR-A = scavenger receptor A; SR-BI = scavenger receptor B1; VEGF-A = vascular endothelial growth factor A; VEGFR2 = VEGF receptor 2*

Vascular Endothelial Growth Factor (VEGF)-A was identified by a microscopy-based high-content screening of HAECs incubated with 141 kinase-inhibiting drugs and fluorescently labeled LDL or HDL (section 1.3.3). Inhibitors of VEGF receptors significantly decreased the uptake of HDL but not LDL. Silencing of VEGF receptor 2 significantly decreased cellular binding, association, and transendothelial transport of radio-iodinated (125I) HDL but not of 125I-LDL. RNA interference with VEGF receptor 1 or VEGF receptor 3 had, however, no effect. Binding, uptake, and transport of HDL but not LDL were strongly reduced in the absence of VEGF-A from the cell culture medium and were restored by the addition of VEGF-A. The restoring effect of VEGF-A on endothelial binding, uptake, and transport of HDL was abrogated by pharmacological inhibition of phosphatidyl-inositol 3 kinase (PI3K)/protein kinase B (Akt) or p38 mitogen-activated protein kinase (p38MAPK), as well as silencing of SR-BI. Moreover, the presence of VEGF-A was found to be a prerequisite for the localization of SR-BI in the plasma membrane of endothelial cells.

Interestingly, UZH found the same mechanism to be operative in clear cell renal carcinomas (ccRCC). ccRCCs are characterized by inactivation of von Hippel-Lindau (*VHL*) gene and

intracellular lipid accumulation by unknown pathomechanism. The immunochemical analysis of 356 renal cell carcinomas revealed high abundance of apolipoproteins apoA-I and apoB as well as SR-BI in the clear cell RCC subtype. Given the characteristic loss of VHL function in ccRCC, we used VHL-defective and VHL-proficient cells to study the potential influence of VHL on lipoprotein uptake. VHL-defective patient-derived ccRCC cells and cell lines (786O and RCC4) showed enhanced uptake as well as less re-secretion and degradation of 125I high and low-density lipoproteins compared to the VHL-proficient cells. The ccRCC cells showed enhanced VEGF and SR-BI expression compared to normal kidney epithelial cells. Uptake of 125I-HDL and 125I-LDL by patient-derived normal kidney epithelial cells as well as VHL-re-expressing ccRCC cell lines 786-O-VHL and RCC4-O-VHL cells was strongly enhanced by VEGF treatment. The knock-down of VEGF co-receptor neuropilin (NRP1) as well as blocking of SR-BI significantly reduced the uptake of lipoproteins into ccRCC cells *in vitro*. LDL stimulated proliferation of 786-O cells more potently than 786-O-VHL cells in a NRP1- and SR-BI- dependent manner. In conclusion, enhanced lipoprotein uptake due to increased activities of VEGF/NRP1 and SR-BI promotes lipid accumulation and proliferation of VHL-defective ccRCC cells (**Figure 5**).

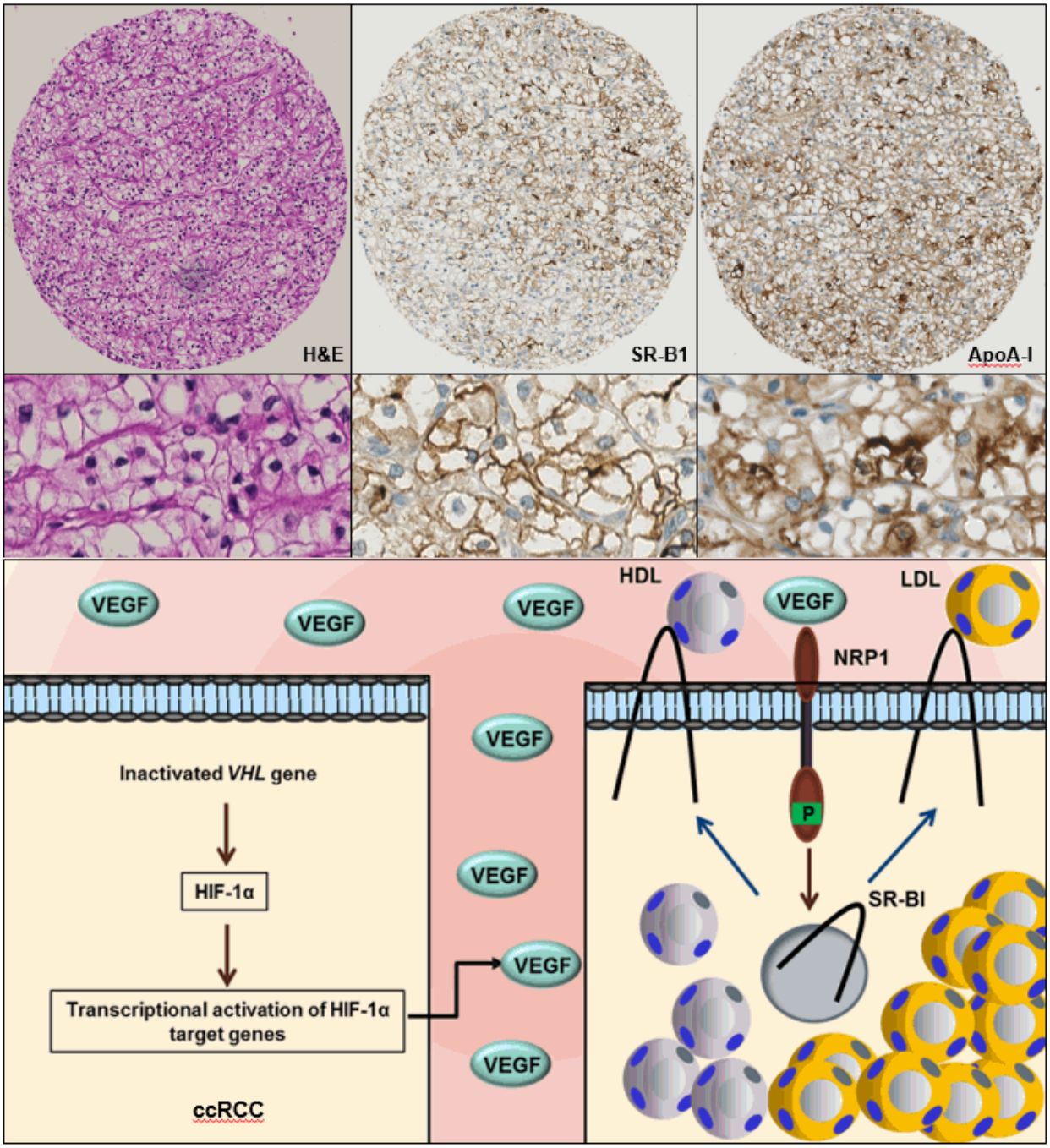


Figure 5. *Cytoplasmic lipoprotein accumulation in clear-cell renal cell carcinoma (ccRCC). The characteristic clear cytoplasm of clear-cell renal cell carcinoma (ccRCC) is caused by the intracellular accumulation of cholesterol. As the likely origin, the interaction of vascular endothelial growth factor VEGF with neuropilin (NRP1) was found to stimulate the cell surface translocation of scavenger receptor SR-B1, which promotes the uptake of both HDL and LDL into a non-degrading intracellular compartment (Velagapudi et al J Lipid Res. 2018 in press. doi: 10.1194/jlr.M083311)*

Sphingosine-1-phosphate (S1P) is an endogenous lipid agonist of five G-protein coupled receptors termed S1P1, S1P2, S1P3, S1P4 and S1P5 27. In the endothelium, binding of S1P to the S1P1 or S1P3 receptors promotes the closure of intercellular junctions and hence the maintenance of the endothelial barrier. Thereby S1P controls the trafficking of solutes, proteins, and cells between intra- and extravascular compartments. UZH found that agonists of S1P1 and S1P3 increase the apical-to-basolateral transport of 125I-HDL but decreased the transcytosis of 125I-LDL through HAECs cultivated in a trans-well system. Conversely, pre-treatment with inhibitors or siRNAs against S1P1 and S1P3 decreased the transport of 125I-HDL but increased the transport of 125I-LDL. Silencing of SR-BI abrogated the stimulation of 125I-HDL transport by both S1P1 and S1P3 agonists. The stimulatory effects of the S1P1 and the S1P3 inhibitor on endothelial transport of 125I-LDL were abolished by treatment with the fluid-phase inhibitor amiloride and heparinase, respectively. Supporting UZH's *in vitro* findings of differential effects of S1P1 and S1P3, their collaborators Dr. J.-R. Nofer (University of Münster, Germany) and Dr. F. Poti (University of Modena, Italy) found, that the endothelial knock-in of S1P1 into apoE-haploinsufficient mice decreases the transport of both LDL and Evan's Blue into the peritoneal cave whereas the endothelial knock-in of S1P3 decreased the transport of LDL but not Evan's Blue. Moreover, either knock-in decreased atherosclerosis of apoE-haploinsufficient mice. This differential and antagonistic regulation of transendothelial HDL and LDL transport by S1P and its cognate receptors S1P1 and S1P3 observed both *in vitro* and *in vivo* provide the as yet strongest evidence that transendothelial transport of HDL and LDL through the endothelium occurs by specific mechanisms rather than passive filtration

The liver plays a crucial role in cholesterol homeostasis through controlling lipid uptake and synthesis. Clearance of plasma lipids via the liver is mediated by the low-density lipoprotein receptor (LDLR) and LDLR-related protein 1 (LRP1). Both receptors are member of the evolutionarily conserved LDLR family, which comprises of seven members. The role of LDLR in clearing circulating atherogenic lipoprotein particles, such as VLDL and LDL, has been well established. Familial hypercholesterolemia (FH), an autosomal-dominant inherited disorder characterized by high plasma LDL cholesterol levels and accelerated atherosclerosis, is caused by a plethora of mutations in LDLR. Genetic ablation of the *Ldlr* in mice also leads to elevation in plasma LDL cholesterol and makes these animals susceptible to atherosclerosis upon high cholesterol feeding.

Validation of genes regulating the trafficking of the LDL receptor (WP6)

Prior to the start of TransCard, the UCMG observed increased plasma cholesterol in a mouse model with hepatic *Commd1* deficiency. Mechanistic evidence pointed at a role of the respective protein in the intracellular trafficking of the hepatic LDL receptor (LDLR). TransCard enabled expansion of these intriguing findings which turned out to be very successful: a large series of associated genes that all impact plasma lipid levels have been studied over the course of the entire project. An interesting aspect to this specific topic is that the expression of the genes involved is not regulated. The regulation occurs at the protein level. This fits with the initial findings of REGION-H in which mutations in the COMMD1 gene in individuals with extreme LDL cholesterol levels, did not increase our understanding of the role of the encoded protein in human lipid metabolism. Rather these findings prompted the question whether these studies should be continued. This because the pharmaceutical industry currently tends to only invest in targets for which human genetic evidence for a role in lipid metabolism and CVD can be provided (preferably through Mendelian Randomization studies). TransCard nevertheless decided to continue these fundamental studies to

improve our understanding how the intracellular trafficking of membrane receptors such as the LDLR in the cells of the liver is regulated. These studies have ultimately led to the insight that the pathways investigated in mice may be equally important to humans with important effects on both plasma lipid levels and risk of CVD. The section below summarizes this work which shows that the use of genetic epidemiological studies to select targets of study misses out on important opportunities.

At the cell surface, LDL binds through ApoB100 to LDLR. LDL and LDLR are endocytosed together (**Figure 5**). Cholesterol is taken up by the cells and LDLR can be reused by the cell through transporting the receptor from the endosomes back the cell surface, where the next cargo can be taken up by the receptor. LDLR can be recycled ±100 times but the mechanism by which LDLR is recognized at the endosomes and transported back to the cell surface has been a mystery. The UMCG identified novel players that are required for the trafficking of LDLR from the endosomes back to the cell surface. It was first of all shown that the mechanism by which these players facilitate endosomal LDLR trafficking is highly conserved between mice and human.

The CCC complex

As already indicated, we identified COMMD1 as a novel player in the regulation of plasma cholesterol levels. We showed that hepatic *Commd1* deficiency in mice and dogs results in hypercholesterolemia (**Table 1**). COMMD1 forms together with three other proteins the so called CCC complex. A recent study has found that the CCC complex is physically associated with another multiprotein complex composed of five proteins. One of these proteins (from now on denoted as W) facilitates the endosomal trafficking of an array of transmembrane proteins. The physical interaction between CCC and W led us to hypothesize that the CCC complex mediates the trafficking of LDLR from the endosomes back to the cell surface. Indeed, we found that loss of COMMD1 impairs the endosomal trafficking of LDLR, leading to decreased levels of LDLR at the cell surface and impaired LDL uptake. Reduced cellular LDL uptake due to compromised endosomal LDLR trafficking likely causes hypercholesterolemia in mice and dogs lacking hepatic expression of COMMD1. In addition to LDLR, we uncovered that COMMD1 is also required to facilitate the endosomal trafficking of another lipoprotein receptor. Using a mouse model with a human-like lipoprotein profile, we assessed the contribution of hepatic COMMD1 deficiency

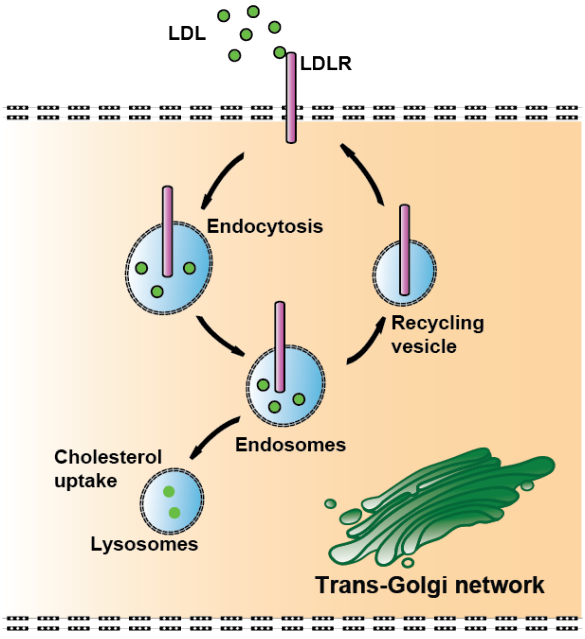


Figure 5. *Simplified model of the trafficking route of LDLR. Cholesterol-rich ApoE and ApoB- containing lipoprotein particles bind to LDLR and are together with LDLR internalized and accordingly directed to the endosomes. From the endosomes LDLR can be transported back to the cell surface for reuse. The mechanism by which LDLR is transported from the endosomes back to the cell surface is investigated within Transcard.*

on atherosclerosis. We found that ablation of hepatic COMMD1 aggravates dyslipidemia and accelerates atherosclerosis.

We found showed that ablation of COMMD1 blunted the expression of all other CCC components. A similar reduction of the all proteins of the CCC complex was also observed when another key component of the CCC complex was genetically ablated in mouse hepatocytes using somatic CRISPR/Cas9-mediated genome editing. Again, hepatic deficiency of this other CCC component also caused hypercholesterolemia in mice.

COMMD family of proteins: effects on plasma lipids and atherosclerosis in mice

COMMD1 forms together with nine other COMMD proteins, the COMMD family of proteins. In previous studies, we and others have shown that the COMMD proteins can interact which each other. In addition, it has been found that all COMMD proteins can interact with one of the CCC components, but the biological role of the interaction between COMMD proteins and CCC was still unclear. In this subproject, we showed that ablation of either COMMD1 or the CCC component resulted in a strong reduction of the other COMMD. This observation made us decide to investigate the effect of hepatic ablation of three other COMMD proteins. Unexpectedly, loss of either of the three led to a marked decrease in the protein expression of all COMMD proteins and the CCC core components. This reduction in the CCC complex was accompanied with hypercholesterolemia in mice. Altogether, in this study found that all COMMD proteins participate in the CCC complex and form together with CCC core components a stable multiprotein complex required for the endosomal trafficking of the lipoprotein receptors LDLR (and an additional lipoprotein receptor). Inactivation of this multiprotein complex impairs the trafficking of these receptors from the endosomes back to the cell surface, resulting in reduced LDL cholesterol uptake and accelerated atherosclerosis in mice.

A second multiprotein complex – W - involved in trafficking of the LDLR

Another protein complex denoted as W (see section on CCC complex above) was also found to impair the endosomal trafficking of the low-density lipoprotein receptor (LDLR). This attenuated endosomal LDLR trafficking leads to LDLR accumulation in endosomes, reduced LDLR surface levels and subsequently impaired LDL uptake. In the present study, we generated liver specific W deficient mice and found that disturbance of hepatic W leads to increased plasma LDL cholesterol as well as and HDL cholesterol levels in mice. We observed that W is not only required for the cell surface expression of LDLR family members but also other main lipoprotein receptors. Metabolic studies moreover showed that selective HDL cholesteryl ester uptake was impaired in hepatic W deficient mice. Finally, we genetically confirmed that W acts in concert with the CCC complex to facilitate the endosomal trafficking of LDLR.

A third complex – V - involved in the trafficking of the LDLR

The first step in cargo recycling is the retrieval of integral membrane proteins (also referred to as cargos) from lysosomal degradation. Cargos that are destined for degradation will undergo ubiquitination and are subsequently degraded by the lysosomes. Cargos that are not marked for degradation can be recognized for retrieval from the degradative fate for recycling back to the cell surface. A central player in the process of cargo recognition to prevent lysosomal degradation of endosomal cargo is the V complex which consists of three proteins and is recruited to the endosomes

through an interaction with another protein. In concert with adaptor proteins, it will recognize and sort specific cargos to their subcellular destination. It has been shown that W (see above) acts in concert with V. A very recent study has indicated that the localization of the W complex to the endosomes does not completely rely on a specific V protein. To investigate the contribution of this V protein on W-mediated lipoprotein metabolism *in vivo*, we ablated its expression in the liver of mice which increased plasma LDL cholesterol levels compared to wild-type mice. Remarkably, in contrast to W deficiency, the inactivation of the gene encoding of the V protein resulted in significant reduction in plasma triglyceride. Interestingly, we also noticed markedly attenuated diet-induced body weight gain and liver weight. Follow up studies are ongoing to assess the relation of this V protein in the pathway in which CCC and W control plasma LDL cholesterol and triglyceride levels.

A fourth multiprotein complex – R - involved in the trafficking of the LDLR

We identified that the W complex acts in concert with the CCC complex to facilitate the endosomal trafficking of several lipoprotein receptors. Recently, another multiprotein complex has been identified, that is involved in the regulation of endosomal recycling of the $\alpha_1\beta_5$ integrin, independent of the V protein described above. This so-called R complex is formed by three proteins and has structural similarities to the V complex. The R complex is thought to associate with both CCC and W complexes and to localize to the endosomes. There, the R complex coupled to the cargo adaptor protein Z, which had been previously shown to be involved in endocytosis of LDLR, likely by accelerating the rate of LDLR cycling through the early endocytic compartment. Z seems to facilitate endosomal trafficking by binding the NPxY motif on cargoes including $\alpha_1\beta_5$ integrin and the LDLR, indicating that the Z protein and R both participate in the CCC-W axis facilitating the endosomal trafficking of LDLR. We elucidated the contribution of protein of the R complex in LDLR recycling in the liver by using somatic CRISPR/Cas9 genome editing approach. We successfully ablated the expression of a subunit of R complex in mouse livers (±80% reduction). Preliminary analyses show that cholesterol levels in plasma were significantly increased in these mice. A slight increase in plasma cholesterol was also shown in mice lacking hepatic Z, however, this increase was not statistically significant compared to control mice. Triglyceride levels in plasma were unaffected in both groups. The increase in total cholesterol levels in hepatic R deficient mice was caused by elevated levels of LDL cholesterol.

The R component has been identified to facilitate the retrieval and sorting of transmembrane proteins (cargos) in the endosomes in a retromer-independent manner. However, our data show that both components of the retriever and V are required to maintain normal plasma LDL cholesterol, likely through coordinating the recycling of LDLR from the endosome back to the plasma membrane. Thus, our data suggest that R and V are not forming different sorting axes at the endosomes, but are likely both involved in the W/CCC-mediated LDLR recycling pathway. The mechanism by which these protein complexes act together in the W-CCC axis has to be determined but our data exclude the possibility that these protein complexes are required for the integrity of the CCC complex. Currently, followed up *in vitro* studies are needed to elucidate this mechanism, and whether they directly facilitate the trafficking of LDLR.

Translating basic science to relevance for human lipid metabolism and risk of CVD

As already indicated, the initial studies by REGION-H did not support our studies into 'LDLR trafficking genes'. However, very rare mutations did elude to a possible role for these genes in human lipid metabolism. These were initially found in the gene encoding for one of the CCC complex components which cause a severe developmental and neurological disorder. We found that these patients are hypercholesterolemic, with elevations in both plasma LDL cholesterol and HDL cholesterol. Mutations in several components of the W complex have been identified and are causing different neurological and developmental disorders. In one of these patients, we found that a mutation in one the candidate genes is correlated with hypercholesterolemia, however, whether the other mutations in related proteins are also causing hypercholesterolemia is unclear.

The identification of a frequent (protein coding) variant one of in the CCC components by the UMCG, however, has greatly helped to study the role of LDLR trafficking in humans. Together with REGION-H, we revealed that this variant is correlated with decreased plasma LDL cholesterol and lower risk for myocardial infarction. Hepatic overexpression of human the human gene in mice was furthermore shown to reduce plasma LDL cholesterol. Based on these data, we speculate that this common variant improves the recycling of LDLR, enhances hepatic LDL uptake, and subsequently reduces the risk for MI. These results imply that improving the functioning of the endosomal sorting machinery of LDLR might lower plasma LDL cholesterol and reduce CVD risk and thus assign this pathway as a potential new therapeutic target to lower plasma cholesterol.

As outlined under 1.3.4 (Validation of a newly identified candidate genes), the UMCG has developed an NGS-based gene panel to sequence all candidate genes in patients with unresolved FH (AMC) as well as individuals with by contrast extremely low LDL cholesterol. This validation effort is ongoing but multiple rare variants in the LDLR trafficking genes have already been identified which are currently followed up in families (for segregation analyses) as well as in functional studies.

1.3.6 Dissemination, exploitation & valorisation (WP7)

The dissemination strategy of the basic scientific knowledge that has been generated in TransCard is mostly addressing the scientific community. The TransCard consortium has been strongly driven and highly committed in presenting the outcomes of their research at high quality conferences and in publishing work related to the project in top tier journals. The communication strategy also aimed at providing an interface for the various public to be kept informed about the impact of TransCard research activities which has been mainly ensured via the TransCard website (www.transcard-research.org). The TransCard partners have published multiple papers but the main outcomes will be published in the near future as basic science simply takes long. Different lectures have been presented by the partners to European and other International scientific congresses that were gathering experts in the fields of cardiology, genetics and atherosclerosis. The lists of publications and participation to conferences and scientific meetings (external to the project) are available under section 2.

Showcase reports

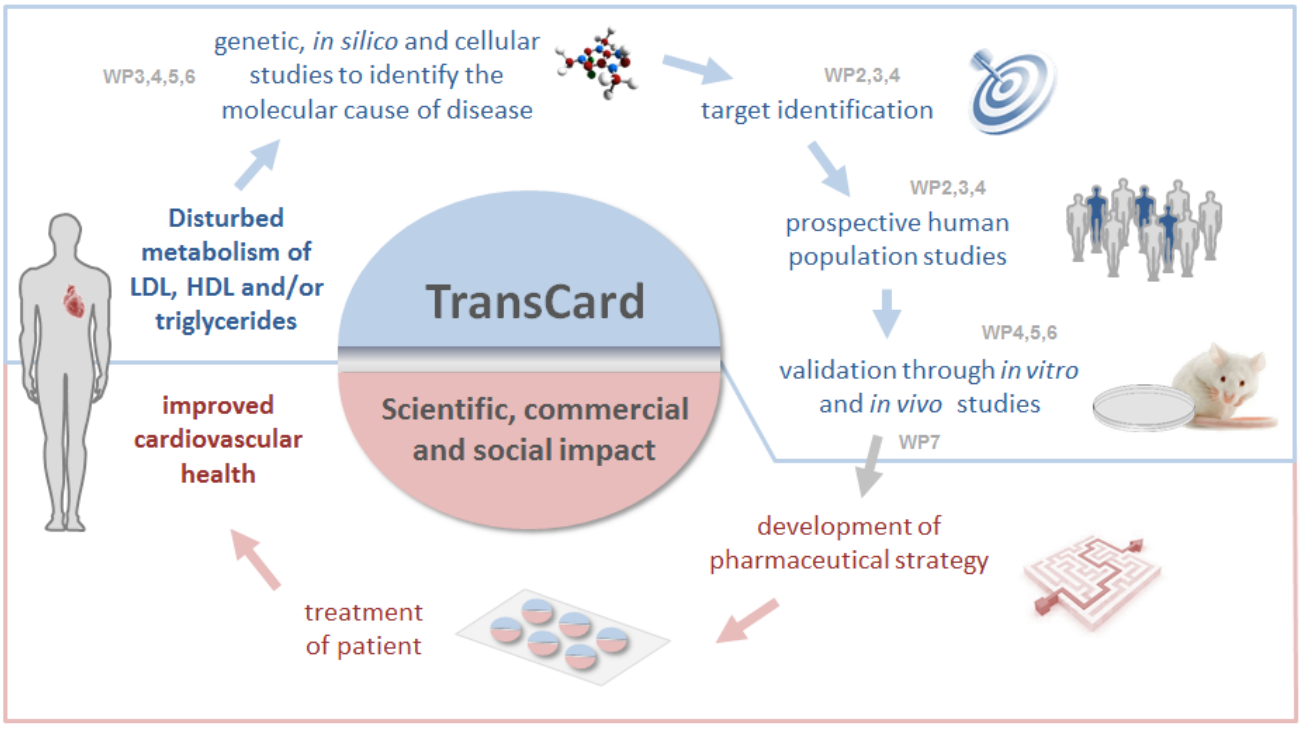
- UZH (manuscript close to submission) and UMCG (manuscript under review) have finalized summarizing key results for publication in top-tier journals. It concerns a main outcome of the genome-wide RNAi screen, and the finding that the trafficking the LDLR in humans affect risk of MI, respectively. These studies are true showcases of the work that has been carried out in TransCard.
- In 2018, UZH has already published a detailed review on the outcome of all the work conducted in TransCard on targets that have been identified to regulate the transport of lipoproteins over the blood vessel wall. This review is timely as this process is currently receiving increased attention in the scientific community as well as the pharmaceutical industry. This because their work and that of others has proven the dogma wrong that the transport of lipoproteins over the endothelium is a mere passive process. In other words, it can be targeted with pharmaceutical compounds.
- UMCG is currently summarizing all the work on the factors that regulate the trafficking of the LDLR in a detailed review. This work is unique in that basic cellular biology is combined with work in experimental animals as well as the relevance to atherosclerosis in humans. As for the transport of lipoprotein over the vascular wall, this work shows that the recycling of the LDLR is a potential target for pharmaceutical intervention to decrease the risk of CVD.
- The QconCAT products – made by POLYQUANT - that have been used in TransCard for absolute quantification of multiple proteins in complex samples by mass-spectrometry can be of specific interest to those involved in translational research. This because these products have been designed to be applicable for both murine and human proteins based on shared peptides. The adaptability of findings in murine samples to human tissue and *vice versa* is a common problem in basic research. In addition, QconCATs can be designed to distinguish between several protein isoforms. A general problem of antibodies is cross-reactivity with protein isoforms of the targeted protein. This problem can perturb results of basic research, as the exact epitope of antibodies and therefore the exact nature of cross-reactivity is often unknown. By using dedicated QconCATs, peptides can be chosen to match only a certain protein isoform. The development of a method for optimal peptide selection for several organisms and specific isoforms of proteins have improved the customer service and the product portfolio of POLYQUANT. A showcase report will be generated to demonstrate the superiority of the QconCAT technique over standard antibody-based procedures.

Immediately marketable products

TransCard did not aim to produce immediately marketable products. Their basic science research efforts have, however, helped to illustrate the advantages of the use of QconCAT technology of POLYQUANT in pure basic science as well as more translational science (in multiple scientific publications). A last series of QconCATs is currently tested for possible use in a clinical setting in the UMCG as it allows to simultaneously assess multiple biomarkers at once in only small samples. Through TransCard, NEBION has finally strengthened its position in the cardiovascular research arena which is likely to render more customers for GENEVESTIGOR, their main product.

Outlicensing of validated targets

Taken the five-year time course of TransCard, out licensing of validated targets was not feasible as illustrated in the key figure of the TransCard application:



This is probably most clearly illustrated by discussing PCSK9 as the foremost example in this field of research. Its identification in a single family with FH in France led to a marvelous breakthrough but despite the notion that all the odds were in favor of this target (detailed below), it took 12 years until approval of pharmaceutical PCSK9 inhibition in the clinic. After its identification, ample genetic variation in the *PCKS9* gene locus allowed to directly validate its role in controlling LDL cholesterol in patients and in the general population and thereby the risk of CVD (allowing for positive Mendelian Randomization studies). Secondly, there was already a wealth of scientific knowledge on the of proprotein convertase family of proteins including PCSK9 when it was identified as an FH gene. This naturally helped elucidating the actual molecular mechanism through which PCSK9 affected LDL metabolism in experimental studies in cells and mice. Last but not least, inhibition of PCSK9 would render a favorable outcome and the pharmaceutical industry is best equipped to produce inhibitors. Finally, the PCSK9 protein is secreted into blood and targeting it right there affects its intracellular function in the liver which made it a perfect target for intervention with pharmaceutical strategies.

1.4. Potential impact, main dissemination activities and exploitation of results

1.4.1 Potential impact

TransCard has been a basic science project which means that it is difficult if not impossible to estimate its future socio-economic impact and the wider societal implications of its activities. As already discussed under paragraph 1.3.6, basic science is mandatory to identify and understand the mechanisms that could help developing appropriate pharmaceutical means to effectively treat

diseases like CVD. Without fundamental knowledge of e.g. the LDLR recycling pathway (Nobel Prize 1986) and *a priori* basic knowledge of the proprotein convertase family of proteins, the path from the discovery of PSCK9 as a target to the use of a new drug that inhibits PCSK9 would have been much, much longer than 12 years.

The discovery of four classes of factors (many of which are novel) that control the cellular uptake of LDL (UZH), and improved insight in the machinery that controls the trafficking of the LDLR to the cell membrane (UMCG) as well as factors that control the transport of lipoproteins of the vascular (UZH) have already contributed and will strongly contribute to existing knowledge which can be used to invent new/better pharmaceutical interventions. Importantly, the UCMG has by now screened the genes of 700 individuals (primarily suffering from hypercholesterolemia) for rare variants in the genes that have been studied and identified in TransCard. The outcome of these studies will be instrumental to translate our basic findings to relevance to human lipid metabolism and risk of atherosclerosis. REGION-H has moreover collected a huge amount of data on established and less established lipid genes that will be exploited for their associations with not only LDL cholesterol but also triglycerides in blood as well as their association with risk of CVD. Trough TransCard, NEBION and POLYQUANT, have exposed their advanced technology to the cardiovascular research community. Finally, the efforts of the AMC have led to a better stratification of their FH4 patient cohort which will improve future chances to find the origins of unresolved hypercholesterolemia. In the same light, the AMC has shown that the molecular diagnosis as well as the clinical diagnosis of FH needs further improvement which may impact patient care.

Last but not least, several targets that have been identified in TransCard which are potentially targetable with drugs but further research is needed.

1.4.2 Main dissemination activities

Main dissemination activities

Type of activities	Main leader	Title	Date	Place	Type of audience	Size of audience	Countries addressed
Posters	Nebion Ag	A high-throughput screening of kinase inhibitors to identify rate limiting steps in the transendothelial lipoprotein transport	21/08/15	Zurich/ Switzerland	Scientific community (higher education, Research)	20	Switzerland
Oral presentation to a scientific event	Universitaet Zuerich	International Atherosclerosis Reseach Summer School : HDL	08/08/18	Hamburg, Germany	Scientific community (higher education, Research)	25	Europe

Type of activities	Main leader	Title	Date	Place	Type of audience	Size of audience	Countries addressed
Oral presentation to a scientific event	Universitaet Zuerich	HDL ? Where are we now?	24/08/15	Prague	Scientific community (higher education, Research)	25	international
Oral presentation to a scientific event	Universitaet Zuerich	9. DeGAG-Kongress	23/09/17	Bad Oeynhausen (DE)	Scientific community (higher education, Research)	50	Germany
Oral presentation to a scientific event	Academisch Ziekenhuis Groningen	Dutch Liver retreat	14/10/16	Spier (NL)	Scientific community (higher education, Research)	50	The Netherlands
Posters	Academisch Medisch Centrum Bij De Universiteit Van Amsterdam	ISA	23/05/15	Amsterdam	Scientific community (higher education, Research)	1000	Europe
Oral presentation to a scientific event	Academisch Medisch Centrum Bij De Universiteit Van Amsterdam	International Atherosclerosis Society (IAS)	09/06/18	Toronto, Canada	Scientific community (higher education, Research)	1500	International
Oral presentation to a scientific event	University Medical Center Groningen	International Atherosclerosis Society (IAS)	09/06/18	Toronto, Canada	Scientific community (higher education, Research)	1500	International
Oral presentation to a scientific event	Universitaet Zuerich	86th European Atherosclerosis Society Congress: Differerential and antagonistic regulation of transendothelial transport of HDL and LDL by sphingosine-1-phosphate receptors 1 and 3	06/05/18	Lisbon Portugal	Scientific community (higher education, Research)	2000	International
Oral presentation to a scientific event	Academisch Medisch Centrum bij de Universiteit	American Heart Association (AHA)	11/11/17	New Orleans, Louisiana, USA	Scientific community (higher education, Research)	10000	International

Type of activities	Main leader	Title	Date	Place	Type of audience	Size of audience	Countries addressed
	van Amsterdam						
Oral presentation to a scientific event	Academisch Medisch Centrum bij de Universiteit van Amsterdam	American Heart Association (AHA)	10/11/18	Anaheim, California, USA	Scientific community (higher education, Research)	10000	International
Oral presentation to a scientific event	Academisch Medisch Centrum bij de Universiteit van Amsterdam	European Society of Cardiology (ESC)	30/08/17	Barcelona, Spain	Scientific community (higher education, Research)	30000	International
Oral presentation to a scientific event	Academisch Medisch Centrum bij de Universiteit van Amsterdam	European Society of Cardiology (ESC)	29/08/18	Munich, Germany	Scientific community (higher education, Research)	30000	International

Main publications

D.O.I.	Title	Author(s)	Title of the periodical or the series	N°	Publisher	Place of publication	Date of publication	Relevant pages
10.1194/jlr.M050021	Apolipoprotein M modulates erythrocyte efflux and tubular reabsorption of sphingosine-1-phosphate	I. Sutter , R. Park , A. Othman , L. Rohrer , T. Hornemann , M. Stoffel , O. Devuyst , A. von Eckardstein	Journal of Lipid Research	Vol. 55/Issue 8	American Society for Biochemistry and Molecular Biology Inc.	United States	01/08/14	1730-1737
10.1161/ATVBAHA.117.309284	VEGF-A Regulates Cellular Localization of SR-BI as Well as Transendothelial Transport of HDL but Not	Srividya Velagapudi , Mustafa Yalcinkaya , Antonio Piemontese , Roger Meier ,	Arteriosclerosis, Thrombosis, and Vascular Biology	Vol. 37/Issue 5	Lippincott Williams and Wilkins	United States	01/05/17	794-803

D.O.I.	Title	Author(s)	Title of the periodical or the series	N°	Publisher	Place of publication	Date of publication	Relevant pages
	LDLHighlights	Simon Flyvbjerg Nørrelykke , Damir Perisa , Arnold von Eckardstein						
10.1161/CIRCRESAHA.117.312004	COMMD Family Regulates Plasma LDL Levels and Attenuates Atherosclerosis Through Stabilizing the CCC Complex in Endosomal LDLR Trafficking	Alina Fedoseienko , Melinde Wijers , Justina C Wolters , Daphne Dekker , Marieke Smit , , Jan Albert Kuivenhoven , Bart van de Sluis	Circulation Research	-	Lippincott Williams and Wilkins	United States	01/01/18	CIRCRESAHA.117.312004
10.1097/MOL.0000000000000411	News on the molecular regulation and function of hepatic low-density lipoprotein receptor and LDLR-related protein 1	Bart van de Sluis , Melinde Wijers , Joachim Herz	Current Opinion in Lipidology	Vol. 28/Issue 3	Lippincott Williams and Wilkins	United States	01/01/17	241-247
10.1038/ncomms10961	CCC- and WASH-mediated endosomal sorting of LDLR is required for normal clearance of circulating LDL	Paulina Bartuzi , Daniel D. Billadeau , Robert Favier , , , Albert K. Groen , Alison M. Elliott , Jan Albert Kuivenhoven , , , Marten H. Hofker , Bart van de Sluis	Nature Communications	Vol. 7	Nature Publishing Group	United Kingdom	11/03/16	10961
10.1194/jlr.M083311	Scavenger receptor BI promotes cytoplasmic accumulation of lipoproteins in clear-cell renal cell carcinoma	Srividya Velagapudi , Peter Schraml , Mustafa Yalcinkaya , Hella Anna Bolck , Lucia Rohrer , Holger Moch , Arnold von Eckardstein	Journal of Lipid Research	FASEB	American Society for Biochemistry and Molecular Biology Inc.	United States	06/03/18	jlr.M083311
10.1016/j.atherosclerosis.2018.06.881	Endocytosis of lipoproteins	Paolo Zanoni , Srividya Velagapudi , Mustafa Yalcinkaya , Lucia Rohrer , Arnold von Eckardstein	Atherosclerosis	Vol. 275	Elsevier Ireland Ltd	Ireland	01/08/18	273-295

D.O.I.	Title	Author(s)	Title of the periodical or the series	N°	Publisher	Place of publication	Date of publication	Relevant pages
10.1093/eurheartj/ehy068	Genetic variants in CYP7A1 and risk of myocardial infarction and symptomatic gallstone disease	Faiza Qayyum , Bo K Lauridsen , Ruth Frikke-Schmidt , Klaus F Kofoed , Børge G Nordestgaard , Anne Tybjærg-Hansen	European Heart Journal	Vol. 39/Issue 22	Oxford University Press	United Kingdom	07/06/18	2106-2116
10.1016/j.jacc.2018.05.044	Clinical Genetic Testing for Familial Hypercholesterolemia	Amy C. Sturm , Joshua W. Knowles , Samuel S. Gidding , Zahid S. Ahmad , E. Hershberger , G. Kees Hovingh , Lala Karayan , , William A. Neal , Børge	Journal of the American College of Cardiology	Vol. 72/Issue 6	Elsevier USA	United States	01/08/18	662-680
10.1161/CIRCULATIONAHA.118.034706	Complete and Partial Lecithin:Cholesterol Acyltransferase Deficiency Is Differentially Associated With Atherosclerosis	Federico Oldoni , Damiano Baldassarre , , G. Kees Hovingh , , Guido Franceschini , Jan Albert Kuivenhoven , Adriaan G. Holleboom , Laura Calabresi	Circulation	Vol. 138/Issue 10	Lippincott Williams and Wilkins	United States	04/09/18	1000-1007
10.1016/S0140-6736(17)32292-4	Very low LDL-cholesterol concentrations achieved: which target is next?	G Kees Hovingh , S Matthijs Boekholdt , Erik S Stroes	Lancet, The	Vol. 390/Issue 10106	Elsevier Limited	United Kingdom	01/10/17	1930-1931
10.1161/CIRCULATIONAHA.116.023942	Diagnosis and Management of Individuals With Heterozygous Familial Hypercholesterolemia	G. Kees Hovingh , John J.P. Kastelein	Circulation	Vol. 134/Issue 10	Lippincott Williams and Wilkins	United States	06/09/16	710-712
10.1126/science.aad3517	Rare variant in scavenger receptor BI raises HDL cholesterol and increases risk of coronary heart disease	Paolo Zanoni , Sumeet A. Khetarpal , Daniel B. Larach , William F. Hancock-Cerutti , John S. Millar , Marina Cuchel , Stephanie	Science	Vol. 351/Issue 6278	American Association for the Advancement of Science	United States	11/03/16	1166-1171

D.O.I.	Title	Author(s)	Title of the periodical or the series	N°	Publisher	Place of publication	Date of publication	Relevant pages
		DerOhannessian , Anatol Kontush , Praveen Surendran , Danish Saleheen , Stella Trompet , J. Wouter Jukema , Anton De Craen ,						
10.1016/j.atherosclerosis.2015.09.021	Familial hypercholesterolaemia: A global call to arms	Antonio J. Vallejo-Vaz , Sreenivasa Rao Kondapally Seshasai , Della Cole , G. Kees Hovingh , John J.P. Kastelein ,	Atherosclerosis	Vol. 243 /Issue 1	Elsevier Ireland Ltd	Ireland	01/11/15	257-259

1.4.3 Exploitation of results (if any)

NEBION and UMCG are in the process of filing a patent.