

From the
Institute of Laboratory Medicine
of the Ludwig-Maximilians-Universität München
Director: Prof. Dr. med. Daniel Teupser

**Genetic causes of atherosclerosis and their molecular and cellular implications at the
susceptibility locus on chromosome 21**



A thesis submitted for the degree of
doctor medicinae
at the Faculty of Medicine of the
Ludwig-Maximilians-Universität München

submitted by
Thomas Kallweit

from
Oettingen i. Bay.

on
2022

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Supervisor: Prof. Dr. med. Dr. rer. nat. Lesca Holdt

Second evaluator: Prof. Dr. med. Andreas Schober
Prof. Dr. med. Jürgen Bernhagen

Dean: Prof. Dr. med. Thomas Gudermann

Date of oral defense: 20.10.2022

Affidavit

Kallweit, Thomas

Surname, first name

I hereby declare, that the submitted thesis entitled

Genetic causes of atherosclerosis and their molecular and cellular implications at the susceptibility locus on chromosome 21

is my own work. I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the submitted thesis or parts thereof have not been presented as part of an examination degree to any other university.

Zürich, 24.10.2022

Place, date

Thomas Kallweit

Signature doctoral candidate

For my wife, Mareike
I am truly grateful for having you in my life
für immer

Acknowledgements

First, I would like to express my sincere gratitude to Professor Daniel Teupser for the opportunity to do my doctoral thesis at the Institute of Laboratory Medicine of the University Hospital Munich and the excellent support and promotion through all the years.

Especially, I must thank my supervisor, Professor Lesca M. Holdt, for her tireless analysis, comments, and support over the last years, all of which have improved me both professionally and personally. I quite simply could not have wished for a better and more dedicated supervisor.

I would also like to pay tribute to all of the academic and administrative staff of the laboratory group of Professors Teupser and Holdt at the Institute of Laboratory Medicine. They are a credit to their university and their profession, and they are what makes the University of Munich such an extraordinary institution.

Specifically, I would like to thank Bernd Northoff for his extraordinary commitment in the help with the bioinformatical analyses and their visualisation. Furthermore, he was always a crucial reference for stimulating discussions. I would also like to thank Frederick Kleine, as with the help of his bioinformatical analyses the thesis was pushed even further.

I would like to thank Anika Stahringer, Maria Eberle and Yasemin Terzibas for introducing me to the world of experimental research and, of course, for their help in the planning, designing and executing of ambitious cell culture experiments. Without their cooperation many experiments could not have been realized. At the same time, I would like to thank Wolfgang Wilfert for his dedicated work and help at all times. He was able to fulfil any extra request a researcher could have and with his interesting discussions besides academia, he became an invaluable companion. I would also like to thank PD Andreas Herbst for his stimulating discussions and advice for my future career.

Last but not the least, I would like to thank Dr Alexander Kohlmaier for correcting the manuscript of my thesis, and for his constructive criticism and discussions. Without him, the present work would not have reached at its current state.

Finally, I would like to thank my family for their support over the years. My parents, Monika and Erich, for earning an honest living for me, for shaping the values of knowledge, education, and dedicated hard work from a young age on. Their good examples have taught me to work hard for the things I aspire to achieve – in this thesis, as in life.

Last of all, I must thank my wife, Mareike, for her patience, encouragement, faith, and love. Nothing of this would have been possible without her.

Abstract

Atherosclerosis is a chronic inflammatory disease of large and medium-sized vessels. It is induced and maintained by various cellular and molecular mechanisms. Besides environmental factors, the genetic background of an individual contributes to the predisposition for the development of atherosclerosis. With the help of genome-wide association studies (GWAS), many genetic variants associated with atherosclerosis have already been unraveled. However, for most of the variants, their causative role in atherogenesis is still widely unclear, not least because many GWAS hits lie in non-protein-coding parts of the genome which are, thus far, not so well understood.

It was the aim of this thesis to create a comprehensive overview of currently known atherosclerosis susceptibility loci from published GWAS and, subsequently, to focus on one non-protein-coding risk locus to explore its genetic, molecular and cellular implications for atherogenesis.

Therefore, in a first step, we compiled all GWAS ($n = 124$) associated with atherosclerosis and its risk factors published up to May 2016. The resulting polymorphisms (SNPs) ($n = 216$) were used to define haplotype blocks ($n = 120$), as a tool for examining candidate genetic effector mechanisms. 11 out of 120 atherosclerosis risk-associated haplotype blocks solely contained noncoding RNAs. Focusing on these, and after ranking by effect size, we identified a prominent (OR 1.14) and so far uncharacterized atherosclerosis risk locus on chromosome 21q22.11, and marked it for further examination. It contained two lead SNPs with robust atherosclerosis association, rs9982601 (p -value 1.33×10^{-13}) and rs28451064 (p -value 1.33×10^{-15}). Since both SNPs resided within a lncRNA (*AP000318.2*) but were in linkage disequilibrium with lower ranking atherosclerosis associations in a relatively large noncoding region (> 100 kilobases), we generated two knockouts of differing sizes in human induced pluripotent stem cells (iPSCs) by CRISPR/CAS9 technology, in order to address potential effector mechanisms. Knockout and control wild type hiPSCs were differentiated into cells relevant for atherosclerosis, namely vascular smooth muscle (SMCs) and endothelial cells (ECs). We detected enhanced expression of neighbouring genes when regions encompassing parts of the lncRNA-encoding core locus were deleted. This effect was most prominently seen in SMCs and pertained to genes 3' and 5' to the risk locus, including protein-coding genes, over a genomic distance of at least 700 kilobases. In terms of function, we observed altered apoptosis as well as decreased proliferation and migration, and increased adhesion for the knockout comprising the atherosclerosis risk interval. These are features observed in atherosclerotic lesions, hence underlining the atherogenic character of the knockout. Collectively, our data from knockout analyses narrowed down the effector element to reside between chr21:35 589 815-35 653 842 and gave rise to the hypothesis that the atherosclerosis-causing mutation on chr21q22.11 deactivates a noncoding, gene-regulatory element, thereby leading to proatherogenic gene expression nearby, and downstream altered proatherogenic mechanisms in cells that constitute the vascular wall.

Abstract in German

Die Atherosklerose als chronisch entzündliche Erkrankung großer und mittlerer Gefäße, wird durch vielfältige zelluläre und molekulare Mechanismen ausgelöst und aufrechterhalten. Neben Umwelteinflüssen trägt auch der genetische Hintergrund eines Individuums zur Prädisposition und Entstehung der Atherosklerose bei. Mithilfe von genomweiten Assoziationsstudien (GWAS) konnten bereits viele genetische Varianten detektiert werden, die mit Atherosklerose assoziiert sind. Tatsächlich ist die kausale Rolle der meisten der Varianten weiterhin unklar, nicht zuletzt deshalb, weil viele in bisher nur rudimentär verstandenen nichtkodierenden Teilen des Genoms liegen. Das Ziel der vorliegenden Doktorarbeit ist eine umfassende Aufstellung bereits publizierter, durch genomweite Assoziationsstudien aufgedeckte Atherosklerose-Risikoloci zu geben, um anschließend einen nicht-proteincodierenden Locus zu priorisieren, und seine genetischen, molekularen und zellulären Auswirkungen auf die Atherogenese aufzuarbeiten.

Dazu haben wir in einem ersten Schritt alle GWAS ($n = 124$) erfasst, die mit Atherosklerose und deren Risikofaktoren assoziiert sind. Die resultierenden atheroskleroseassoziierten Polymorphismen (SNPs) ($n = 216$) wurden in Haplotypblöcken abgegrenzt ($n = 120$) und die Gene innerhalb der Haplotypblöcke klassifiziert. 11 der 120 Haplotypblöcke beherbergen ausschließlich Gene für nichtkodierende RNAs. Unter diesen, und aufgrund seiner großen Effektstärke (OR 1.14), wurde ein weitgehend unerforschter Risikolocus auf Chr21q22.11 weiter experimentell untersucht. Er beinhaltet zwei SNPs mit deutlicher Assoziation mit Atherosklerose, rs9982601 (p -Wert 1.33×10^{-13}) und rs28451064 (p -Wert 1.33×10^{-15}), welche – zusammen mit weiteren assoziierten DNA Varianten – in einem > 100 Kilobasen großen genomischen Bereich einer nichtcodierenden lncRNA (*AP000318.2*) kodiert sind. Um mögliche Effektorszenarien für das Atheroskleroserisiko an diesem Locus zu identifizieren, benutzten wir das Zellkultursystem humaner induzierter pluripotenter Stammzellen (iPSCs) und generierten mittels CRISPR/Cas9 Technologie zwei Knockouts unterschiedlicher Länge. Indem wir Knockout- und Wildtypstammzellen in Effektorzellen der Atherosklerose differenzierten, namentlich glatte Muskelzellen (SMCs) sowie Endothelzellen (ECs), entdeckten wir eine verstärkte Expression der 3' und 5' gelagerten benachbarten Gene auf beiden Seiten des Knockouts, und über eine genomische Distanz von circa 700 Kilobasen hinweg. Dieser Effekt wurde am deutlichsten in SMCs detektiert. Auf funktioneller Ebene beobachteten wir geänderte Apoptoseeigenschaften sowie verminderte Proliferation und Migration und erhöhte Adhäsionseigenschaften der Knockoutzelllinien. Dies sind Eigenschaften, die bekanntermaßen mit Atherosklerose assoziiert sind und daher den proatherogenen Charakter der Knockouts unterstreichen. Zusammenfassend grenzen diese Ergebnisse den Risikolocus auf den Bereich Chr21:35 589 815-35 653 842 ein, und weisen auf die Hypothese, dass eine atherosklerose-verursachende Mutation auf Chr21q22.11, ein nichtcodierendes, genregulatorisches, repressives Element deaktiviert, was wiederum zu einer verstärkten proatherogenen

Genexpression mit folgenden pathologisch veränderten, proatherogenen Funktionen in Zellen der Gefäßwand führt.

Contents

List of Figures	vii
List of Tables	viii
List of Abbreviations	x
1 Introduction	1
1.1 Atherosclerosis and cardiovascular disease	1
1.1.1 Epidemiology	1
1.1.2 Pathophysiology of atherosclerosis	1
1.1.3 Aetiology of atherosclerosis	3
1.2 Genome-wide association studies and their implications for studying atherosclerosis	4
1.2.1 Meiotic DNA recombination as force for shaping genetic variation and definition of haplotype blocks in the interpretation of GWAS.....	4
1.2.2 Quantifying genetic “linkage” for mapping SNP data from GWAS	5
1.2.3 The principle of “imputation” for identifying SNPs from microarray based GWAS.....	6
1.2.4 Statistical methods for ranking SNP-disease association from microarray-based GWAS ..	7
1.2.5 Importance of GWAS for studying atherosclerosis on the example of the chromosome 9p21 atherosclerosis risk locus	8
1.3 Functional elements in the human genome and the way to functional studies	9
1.3.1 Protein-coding and noncoding elements of the human genome.....	9
1.3.2 Functions of long non-coding RNAs.....	9
1.3.3 Functions of long non-coding RNAs in atherosclerosis and related mechanisms.....	10
1.3.4 From locus to biology: approaches and challenges of functional studies.....	12
2 Aims of the thesis	14
2.1 Compilation of atherosclerosis susceptibility loci.....	14
2.2 Experimental study of a selected atherosclerosis locus	14
3 Materials	16
3.1 Laboratory Equipment.....	16
3.2 Chemicals and Consumables	16
4 Methods	21
4.1 Molecular Biological Methods of DNA	21
4.1.1 Isolation and purification of DNA.....	21
4.1.2 Analysis of concentration and purity of nucleic acids.....	21
4.1.3 Polymerase chain reaction and gel electrophoresis	22
4.1.4 Genotyping.....	23
4.1.4.1 Genotyping using polymerase chain reaction.....	23

4.1.4.2 Genotyping using melt curve-based analysis	24
4.1.5 Sanger sequencing.....	25
4.1.6 Quantitative Real Time PCR.....	26
4.1.6.1 Absolute quantification of gene expression using standard curve method	26
4.1.6.2 Relative quantification of DNA using internal standards	28
4.1.7 Production of plasmids for CRISPR/Cas9 systems.....	29
4.1.7.1 Design of guide RNA plasmids.....	30
4.1.7.2 Generation and transformation of guide RNA plasmids in <i>Escherichia coli</i>	30
4.1.7.3 Mini- and Maxipreparation of plasmids.....	32
4.1.8 Surveyor Nuclease Assay	33
4.2 Molecular Biological Methods of RNA	35
4.2.1 Isolation of RNA	35
4.2.2 Reverse transcription of RNA into complementary DNA (cDNA).....	35
4.3 Cell culture	36
4.3.1 Cultivation of cells	36
4.3.1.1 Cultivation of human embryonic kidney cells.....	36
4.3.1.2 Cultivation of human induced pluripotent stem cells	36
4.3.1.3 Cultivation of vascular endothelial cells.....	37
4.3.1.4 Cultivation of vascular smooth muscle cells.....	37
4.3.2 CRISPR/Cas9 methodology and transfection of cells	38
4.3.2.1 Test of CRISPR/Cas9 efficiency by transfection of human embryonic kidney cells	38
4.3.2.2 Electroporation of CRISPR/Cas9 and gRNA expression plasmids into human induced pluripotent stem cells	38
4.3.3 Fluorescence-activated cell sorting of human induced pluripotent stem cells.....	39
4.3.4 Manual picking and passaging of human induced pluripotent stem cells.....	40
4.3.5 Generation of vascular endothelial and smooth muscle cells from human induced pluripotent stem cells	40
4.3.6 Functional experiments in cell culture.....	42
4.3.6.1 Apoptosis	42
4.3.6.2 Proliferation.....	42
4.3.6.3 Adhesion	42
4.3.6.4 Migration.....	43
4.4 Bioinformatics analysis	44
4.4.1 Analysis of haplotype blocks	44
4.4.1.1 Definition of haplotype block margins	44
4.4.1.2 Determination of annotated genes within haplotype blocks.....	44
4.4.1.3 Gene expression analysis	45
4.5 Statistical analysis	45
4.5.1 Systematic review of published genome-wide association studies	45
4.5.2 Statistical analyses of functional assays in cell culture	45

5 Results	46
5.1 Survey of genetic loci linked to human atherosclerosis by genome-wide association studies in populations	46
5.1.1 Assembly of genetic variations associated with atherosclerosis risk from existing GWAS	46
5.1.2 Definition of haplotype block borders	48
5.2 Prioritization of haplotype blocks for functional studies	50
5.2.1 Colocalization of haplotype blocks and SNPs associated with risk factors	50
5.2.2 Genetic content of haplotype blocks	52
5.2.3 Candidate haplotype blocks for future experimental studies	53
5.3 Characterization of Chr21q22.11 locus	55
5.3.1 Overview of locus on chromosome 21	55
5.3.2 Expression of genes on chromosome 21 locus in tissues and cells	57
5.4 Design of experimental studies	61
5.4.1 Knockout of haplotype block and subblock	61
5.4.2 Differentiation of hiPSCs into cell types associated with atherosclerosis	66
5.5 Comparison of wild type versus knockout cells	68
5.5.1 Altered gene expression of knockout cells	68
5.5.2 Altered cellular functions of knockout cells	70
5.5.2.1 Altered apoptosis phenotype in SMCs and ECs	70
5.5.2.2 Decreased proliferation phenotype in SMCs and ECs	71
5.5.2.3 Decreased migration phenotype in SMCs	74
5.5.2.4 Increased adhesion phenotype in ECs	75
6 Discussion	77
6.1 Possible explanations for observed gene expression alteration	78
6.2 Interpretation of altered cellular functions	78
6.3 Background and relevance of examined genes	82
6.4 Outlook into <i>in vivo</i> relevance	83
6.5 Implications for future studies and clinical applications	84
Appendices	88
A Supplementary Tables	90
B References	120

List of Figures

Figure 1 Primary atherosclerotic lesion.....	2
Figure 2 Advanced atherosclerotic lesion	3
Figure 3 Functions of lncRNAs in atherosclerosis.....	11
Figure 4 Overview and allocation of all relevant 124 genome-wide association studies.....	47
Figure 5 Methodical work of definition of haplotype block borders	49
Figure 6 Degree of colocalization of atherosclerosis haplotype blocks and risk SNPs.....	50
Figure 7 Progression and methodology of work.....	51
Figure 8 Annotated genes in haplotype blocks	52
Figure 9 Overview of the studied chromosome 21 atherosclerosis risk locus.....	56
Figure 10 Primers for qPCR to probe long noncoding RNA expression from introns and exons located within the core chromosome 21 atherosclerosis risk locus.....	57
Figure 11 RNA expression levels of genes within chromosome 21 risk locus or nearby	60
Figure 12 Overview of CRISPR/Cas9-mediated knockout strategy on chromosome 21 locus.....	62
Figure 13 Knockout sequences and PCR-genotyping of knockout (agarose flash gel).....	65
Figure 14 Time course of differentiation markers	67
Figure 15 RNA expression of genes within and near the chromosome 21 atherosclerosis risk locus in iPSC-derived SMCs and ECs	69
Figure 16 Effect of chromosome 21 knockout on apoptosis levels in iPSC-derived SMCs and ECs..	71
Figure 17 Functional studies exploring proliferation rates in iPSC-derived SMCs and ECs	73
Figure 18 Functional studies exploring migration phenotype of iPSC-derived SMCs	74
Figure 19 Functional studies exploring adhesion phenotype of iPSC-derived ECs.....	76
Figure 20 Synopsis and impact of chromosome 21 locus in atherosclerosis.....	81

List of Tables

Table 1 PCR reaction mix	22
Table 2 PCR thermal cycling conditions.....	22
Table 3 PCR reaction mix	23
Table 4 PCR thermal cycling conditions.....	23
Table 5 PCR reaction mix	25
Table 6 PCR and melt curve thermal cycling conditions	25
Table 7 PCR reaction mix	26
Table 8 PCR thermal cycling conditions.....	26
Table 9 qRT-PCR reaction mix	27
Table 10 qRT-PCR thermal cycling conditions.....	28
Table 11 qRT-PCR reaction mix	28
Table 12 qRT-PCR reaction mix	29
Table 13 Reaction mix for digestion with BbsI	30
Table 14 Composition of Annealing Buffer	31
Table 15 Reaction mix for kinase reaction.....	31
Table 16 Reaction mix for ligation	31
Table 17 PCR reaction mix	33
Table 18 PCR thermal cycling conditions.....	33
Table 19 Hybridization protocol for thermal cycler.....	34
Table 20 Reaction mix for digestion.....	34
Table 21 RT reaction mix.....	36
Table 22 Reaction mix for lipofection	38
Table 23 Reaction mix for electroporation.....	39
Table 24 List of non-protein-coding candidate haplotype blocks, used for selecting a locus for experimental studies in the thesis.....	54

Supplementary Table 1 Oligonucleotides for genotyping of SNP rs9982601 by melt curve	90
Supplementary Table 2 Oligonucleotides for PCR of knockout positive clones	90
Supplementary Table 3 Primer and Probes for qRT-PCR.....	90
Supplementary Table 4 Oligonucleotides for PCR of transfected <i>E. coli</i>	92
Supplementary Table 5 Oligonucleotides for Surveyor Assay	93
Supplementary Table 6 Expected sizes of undigested PCR product and digested fragments	93
Supplementary Table 7 Oligonucleotides for generation of protospacers for gRNAs of CRISPR/Cas9 system.....	93
Supplementary Table 8 Default traits for query in GWAS Catalog.....	93
Supplementary Table 9 Assessed genome-wide association studies	95
Supplementary Table 10 Atherosclerosis SNPs.....	101
Supplementary Table 11 Lipid SNPs	104
Supplementary Table 12 Blood pressure SNPs	108
Supplementary Table 13 Diabetes mellitus SNPs.....	111
Supplementary Table 14 Smoking SNPs.....	113
Supplementary Table 15 Atherosclerosis haplotype blocks	114

List of Abbreviations

AMP	Adenosine monophosphate
ANRIL	Antisense noncoding RNA in the INK4 locus
ASO	Antisense oligonucleotide
ATP	Adenosine triphosphate
ATP5O	ATP synthase subunit O
BACE1-AS	Beta secretase 1 antisense RNA
BMP4	Bone morphogenetic protein 4
bp	Basepair
BSA	Bovine serum albumin
CAD	Coronary artery disease
Cas9	CRISPR associated protein 9
Caspase	Cysteine aspartic acid-specific protease
CD	Cluster of differentiation
CDK4	Cyclin-dependent kinase 4
cDNA	Complementary deoxyribonucleic acid
CO ₂	Carbon dioxide
CoA	Coenzyme A
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
CT	Cycle threshold
CVD	Cardiovascular disease
DALY	Disability-adjusted life year
DEPC	Diethyl pyrocarbonate
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
DTT	Dithiothreitol
E. coli	Escherichia coli
EC	Endothelial cell
EDTA	Ethylenediaminetetraacetic acid
EHS tumor	Engelbreth-Holm-Swarm tumor
eQTL	Expression quantitative trait loci
FACS	Fluorescence-activated cell sorting

FAM.....	5,6-carboxyfluorescein
FB	Fibroblast
FBS	Fetal bovine serum
FDR.....	False discovery rate
FRET.....	Förster resonance energy transfer phenomenon
FSC	Forward scatter
GAS5	Growth arrest specific 5
GFP	Green fluorescent protein
gRNA.....	Guide RNA
GSK3 β	Glycogen synthase kinase 3 β
GWAS.....	Genome-wide association study
H ₂ O	Water
HDL	High-density lipoprotein
HDR.....	Homology-directed repair
HEK-293 cells	Human embryonic kidney cells 293
hESC	Human embryonic stem cell
hg19	Human genome version 19
hiPSC	Human induced pluripotent stem cells
HMG-CoA.....	3-Hydroxy-3-methyl-glutaryl-coenzyme A
HOTAIR.....	HOX transcript antisense RNA
HTT-AS	Huntingtin antisense RNA
indels.....	Insertion/deletion mutations
kb	Kilobase
KCNE1.....	Potassium voltage-gated channel subfamily E member 1
KCNE2.....	Potassium voltage-gated channel subfamily E member 2
LB	Lysogeny broth
LD.....	Linkage disequilibrium
LINC00310.....	Long intergenic non-protein coding RNA 310
lncRNA	Long non-coding RNA
MAF.....	Minor allele frequency
MALAT1	Metastasis associated lung adenocarcinoma transcript 1
mb	Megabase
MgCl ₂	Magnesium chloride
MIAT	Myocardial infarction associated transcript
miRNA.....	Micro RNA
M-MLV RT.....	Moloney murine leukemia virus reverse transcriptase
mRNA	Messenger RNA
MRPS6.....	Mitochondrial ribosomal protein S6

NaCl.....	Sodium chloride
ncRNA	Non-coding RNA
NEAT1_2	Nuclear paraspeckle assembly transcript 1_2
NHEJ.....	Non-homologous end joining
NHGRI-EBI	National Human Genome Research Institute - European Bioinformatics Institute
nt.....	Nucleotide
ORF.....	Open reading frame
PAM.....	Protospacer adjacent motif
PANDA.....	P21 associated non coding RNA DNA damage activated
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate-buffered saline
PCR.....	Polymerase chain reaction
PDGF-BB.....	Platelet-derived growth factor BB
PEG.....	Polyethylene glycol
PenStep	Penicillin/Streptomycin
PPAR γ	Peroxisome proliferator-activated receptor gamma
PRC.....	Polycomb repressive complex
qPCR.....	Quantitative polymerase chain reaction
qRT-PCR.....	Quantitative real time polymerase chain reaction
RCAN1	Regulator of calcineurin 1
rh.....	Recombinant human
rh bFGF	Recombinant human basic fibroblast growth factor
rh BMP4.....	R&D Systems
rh EGF.....	Recombinant human epidermal growth factor
rh FGF-B	Recombinant human fibroblast growth factor-B
rh TGF β	Recombinant human transforming growth factor β
rh VEGF	Recombinant human vascular endothelial cell growth factor
RNA.....	Ribonucleic acid
ROCK	Rho-associated, coiled-coil containing protein kinase
RPKM.....	Reads per kilobase per million mapped reads
RPMI medium	Roswell Park Memorial Institute medium
rRNA.....	Ribosomal RNA
RT	Reverse transcription
RWD	Relative wound density
SEM	Standard error of the mean
siRNA	Small interfering RNA
SLC5A3	Solute carrier family 5 member 3
SmBM.....	Smooth Muscle Cell Basal Medium

SMC	Smooth muscle cell
SMD	Staufen mediated decay
SmGM-2	Smooth Muscle Cell Growth Medium-2
SMIM11	Small integral membrane protein 11
snoRNAs	Small nucleolar RNA
SNP	Single nucleotide polymorphism
snRNA	Small nuclear RNA
SSC	Side Scatter
TAE	Tris-Acetate-Ethylenediaminetetraacetic acid
TAMRA	Tetramethylrhodamine
TE buffer	Tris-EDTA buffer
tracrRNA	Transactivating CRISPR RNA
tRNA	Transfer RNA
UCHL1-AS	Ubiquitin carboxyl terminal hydrolase L1 antisense RNA
UCSC	University of California, Santa Cruz
UTR	Untranslated region
UV	Ultraviolet
VE-cadherin	Vascular endothelial cadherin
VEGF-A	Vascular endothelial growth factor A
VSMC	Vascular smooth muscle cell
Y-27632 dihydrochloride	Rho-associated, coiled-coil containing protein kinase
YLL	Years of life lost

1

Introduction

1.1 Atherosclerosis and cardiovascular disease

1.1.1 Epidemiology

Atherosclerosis and its clinical manifestations (e.g., myocardial infarction or stroke) are the leading causes of death and disability worldwide. With an overall prevalence of 6.6%, atherosclerosis and its effects on the human body affect nearly half a billion people.¹ The disability-adjusted life years (DALYs) are a measurement of disease burden expressing the number of years lost to ill health, disability or premature mortality. Atherosclerotic diseases form the largest part contributing 14.6% of all global DALYs.² In terms of years of life lost, as a measurement of premature death, there was an increase of 14.7% between 2007 and 2017 with ischemic heart disease ranked first and stroke ranked third.³ Of all 55.9 million global deaths in 2017, 17.8 million (32%) were due to atherosclerotic diseases, being the largest entity.³ Deaths due to atherosclerosis related diseases showed an increase of 21.1% between 2007 and 2017, revealing a clear tendency and underlining the global importance of research in this field.³

1.1.2 Pathophysiology of atherosclerosis

Atherosclerosis is a chronic inflammatory disease, induced and maintained by various cellular and molecular mechanisms. Briefly summarizing, at a cellular level a number of key alterations have been linked to the onset of atherosclerosis, which can be a decades-long process: As a first step during atherogenesis, aberrantly stimulated immune cells are captured by pathologically activated endothelial cells (ECs), often at specifically disease-prone regions in the vasculature. This causes focal changes in the vessel wall and triggers a cascade of non-resolving inflammatory processes contributing to growth of the atherosclerotic lesion (Figure 1).⁴ As next step, vascular smooth muscle cells (VSMCs) aberrantly invade inner regions of the vessel wall in the nascent lesion, exit their normal quiescence and start to divide, and at the same time change their cell fate. These, together with diverse activated immune cell types recruited from blood, influence each other to further aberrantly stimulate lesion growth.⁴ Last, cells in the lesion also acquire a certain resistance to apoptosis and to proper clearance of dead cell bodies when dying. The latter leads to a persistent, inflammatory necrotic core in the lesion.⁴ In summary, many cellular changes culminate in a situation such that the atherosclerotic lesion continues

to grow, and progressively narrows the blood flow by constricting the lumen of the affected blood vessel. In this thesis we focus on assays in ECs and VSMCs.

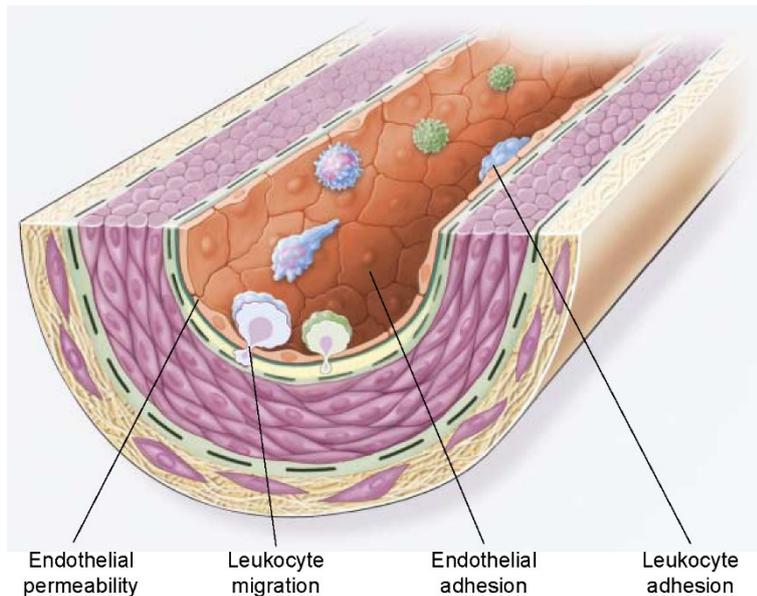


Figure 1 | Primary atherosclerotic lesion. Early changes of atherosclerosis including increased endothelial permeability, up-regulation of adhesion molecules in leukocytes and ECs and the migration of leukocytes into the artery wall. Reproduced with permission from Ross⁴, Copyright Massachusetts Medical Society.

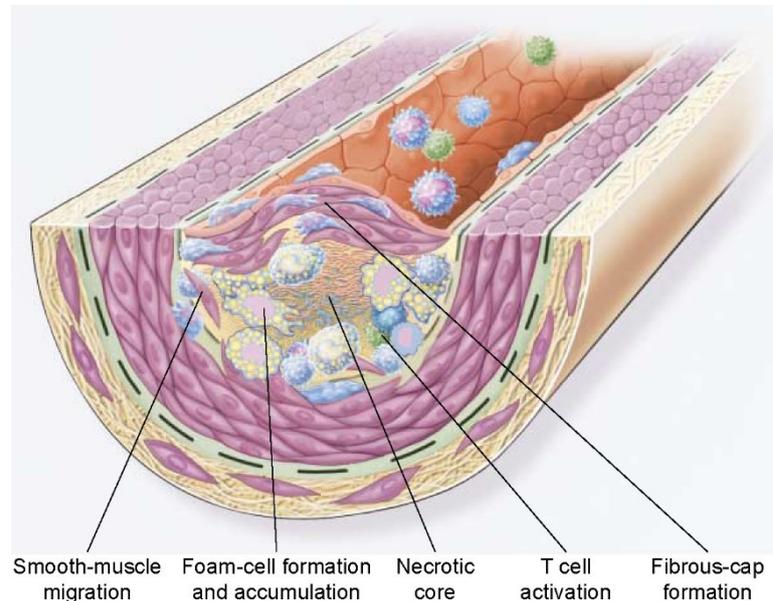
Describing the atherogenic process in more detail and at the molecular level, endothelial cells have been identified as cells hit first during atherogenesis: endothelial dysfunction, with subsequent accumulation of LDL and other lipoproteins in the subendothelial matrix, act as first step of atherogenesis. In one prominent model of atherogenesis, it is thought that LDL particles undergo oxidative modification in the intima.⁵ These oxidized LDL particles promote inflammation by stimulating overlying ECs to produce cytokines, adhesion molecules and chemotactic factors for both monocytes and T cells. The released cytokines and other proinflammatory molecules stimulate monocyte migration, proliferation, and differentiation into macrophages.^{5,6} However, physical forces caused by blood flow can also act as atherosclerosis-triggers: In regions with laminar flow, ellipsoid-shaped cells aligned in the direction of flow are observed. In regions with turbulent flow such as vascular branches, curvatures or bifurcations, polygonal-shaped cells with no specific orientation are observed.⁷⁻¹⁰ This leads to increased permeability of the endothelial lining for molecules such as LDL cholesterol. Therefore, these are sites of predilection for the formation of atherosclerosis.^{9,10}

With the already growing atherosclerotic lesion, highly oxidized LDL particles bind to scavenger receptors of macrophages mediating their uptake. This finally results in formation of foam cells that are very typical of atherosclerotic lesions and have been related to hypercholesterinaemia as one central atherosclerosis risk factor.¹¹ Furthermore, activated macrophages present processed antigens to T cells. These activated T cells subsequently amplify the inflammatory response by secretion of cytokines responsible for migration and proliferation of SMCs from the media into the intima. On a macroscopic scale, the fatty streak is now formed (Figure 2).^{4,9}

Activated macrophages and foam cells experience apoptosis due to stimulation of specific cytokines. This forms the lipid rich necrotic core of the lesion. The covering fibrous cap is formed by

the accumulation of invading SMCs and their secreted extracellular matrix products, mainly collagen. This highlights the primarily secretory phenotype of SMCs within an atherosclerotic lesion, compared to the mainly contractile phenotype of medial SMCs. The necrotic core with overlying fibrous cap is now called atheromatous plaque (Figure 2).^{4,9,12}

Figure 2 | Advanced atherosclerotic lesion. The formation of fatty streaks is brought about by the migration of smooth-muscle cells, T-cell activation, foam-cell formation, and continuous migration of leukocytes. This is the basis for the development of the necrotic core, a result of apoptosis, necrosis, and lipid accumulation. Reproduced with permission from Ross⁴, Copyright Massachusetts Medical Society.



The atheromatous plaque grows by further accumulation of cholesterol, invasion of inflammatory cells, and expansion of the necrotic core. This protrusion of the plaque into the lumen of the vessel leads to stenosis. While stable plaques are characterized by a uniformly dense fibrous cap, unstable plaques show thinning of the fibrous cap. The degradation of the fibrous cap is driven by the increased production and activity of matrix metalloproteinases in the lesion. This results in an ultimately unstable plaque, vulnerable to rupture, leading to arterial occlusion when the plaque or parts of it are taken up in the blood stream.¹²

1.1.3 Aetiology of atherosclerosis

There is not one single origin, but rather various causes that all contribute to the development of atherosclerosis in different ways and to different degrees. They can be roughly divided into biological, psychological, and social factors. While the latter two often act as trigger in terms of an individual's behaviour and lifestyle, the biological factor, drawn from a person's genetic background, is certainly a basic predisposition for atherosclerosis.

Concerning predisposition, genetic epidemiological studies of family history and studies of twins have estimated the heritability of atherosclerosis to be between 40% and 60%. In a follow-up study of over 20 000 twins in the Swedish Twin Registry, heritability of death from CAD was 57% for men and 38% for women, respectively.¹³ Equally, in a study using data from the Danish Twin Registry, heritability was found to be 53% for men and 58% for women, respectively.¹⁴ In the offspring cohort of

the Framingham Heart Study, participants with a family history of premature atherosclerotic diseases had a greater risk for atherosclerotic events with odds ratios of 2.0 for men and 1.7 for women, after multivariable adjustment.¹⁵ A main objective of current research is to identify the specific gene variants and mutations causing atherosclerosis; in other words, to correlate genotypic variation in the population with phenotype.

A small minority of atherosclerotic diseases can be traced back to monogenic causes.¹⁶ In this case, one mutation in a single gene is able to cause a noticeable disease phenotype. In cases of homozygous familial hypercholesterolemia, mutations in the LDL receptor significantly increase plasma LDL cholesterol, causing hypercholesterolemia and subsequently atherosclerosis.¹⁷⁻¹⁹ Irrespective of their high penetrance, Mendelian inheritance of atherosclerosis is rare and only explains a small portion of trait variance.¹⁶

The large majority of atherosclerotic diseases, as is the case also for many other human diseases and complex traits, show a polygenic inheritance pattern. That is also reflected in the multiple risk factors associated with atherosclerosis, such as hypertension, dyslipidaemia, diabetes mellitus, and smoking.²⁰ As proposed by the “common disease – common variant” hypothesis, common disorders are likely due to genetic variants that are also common in the population.¹⁶ This leads to two deductions: First, if the affecting genetic variant is common, its penetrance has to be relatively low. And second, if these common variants have low penetrance, they must appear at multiple loci (in higher numbers) throughout the genome, to explain the total heritability. Up until today, we are able to explain only about 20-30% of the heritability of atherosclerotic diseases.²¹ The main tool in identifying variants contributing to atherosclerosis so far have been the genome-wide association studies (GWAS).²² Therefore, in part one of this thesis we explore existing GWAS with relevance for atherosclerosis.

1.2 Genome-wide association studies and their implications for studying atherosclerosis

1.2.1 Meiotic DNA recombination as force for shaping genetic variation and definition of haplotype blocks in the interpretation of GWAS

The human genome naturally varies in its detailed DNA sequence from individual to individual. In total, over 88 million variants were detected so far by genome-wide sequencing approaches.²³ The most common form of genetic variation is due to single nucleotide polymorphisms (SNPs), contributing 84,7 million (96,3%) hits.²³ The frequency of these polymorphisms is expressed by minor allele frequency (MAF), indicating the abundance of the less common allele in the group of alleles in a reference population. If mutations occur with $MAF > 1\%$, they are referred to as SNPs.²³

Usually, SNPs that are located at nearby sites (close together *in cis* on the same chromosomal region) have a higher chance to be inherited together, which adds further difficulty in attributing

individual SNPs to a single disease or trait. Whether sites are more often inherited together, rather than being split, depends on the rate of their separation in germ cell development. Thereby, during meiosis in germ cell formation, recombination of homologous chromosomes ensures breakage and rearrangement of DNA. Recombination events are more common when two investigated loci (or SNPs) are far apart on a chromosome, and less common when two loci are close together.²⁴ This effect is further exacerbated because recombination is not totally random along a chromosome, but rather tends to occur at narrow hotspots, thus creating genomic sections that are inherited together, called haplotype blocks.²⁴ The pattern of these blocks varies between different populations. For example, as the African population has gone through the most recombination events (as it has the longest known common genetic history), its haplotype blocks are smaller than these from people of Caucasian descent.²⁴ In this thesis we studied GWAS data from the 1000 genomes project from the Northern European CEU population with all associated implications (see below).^{23,25}

1.2.2 Quantifying genetic “linkage” for mapping SNP data from GWAS

The degree to which two SNPs are inherited together within a population can be quantified and is referred to as “linkage”, a concept which has also been instrumental for this thesis in terms of distinguishing the location and size of different atherosclerosis-relevant and risk SNP-containing stretches in the human genome. Two genetic loci are said to be in high linkage disequilibrium (LD) when their respective alleles are inherited together in the same gamete more often than expected by chance alone given their corresponding allele frequencies.²⁶ The level of LD between two SNPs can be assessed by two different measures D' and r^2 . The coefficient of linkage disequilibrium (D) between two alleles A and B is defined as

$$D = P_{AB} - (P_A \times P_B)$$

with P_A and P_B being the frequency of allele A and B, respectively, and P_{AB} being the frequency of both alleles occurring together. The values of D vary with allele frequencies, which can make it difficult to interpret. Therefore, D' is more commonly used. It is defined as the absolute value of D divided by the maximum value that D could take given the allele frequencies.^{26,27} D' ranges between 0 and 1, with $D' = 0$ if both SNPs are inherited independently of each other (complete linkage equilibrium) and $D' = 1$ if there is no evidence of recombination between both SNPs, meaning they are inherited together (complete linkage disequilibrium).^{26,27} An alternative way of measuring LD is by using the squared correlation coefficient r^2 , which is expressed as

$$r^2 = \frac{D^2}{P_A(1 - P_A) P_B(1 - P_B)}$$

It ranges between 0 and 1, with higher values indicating that both SNPs are highly correlated, thus conveying the same genetic variance.²⁶ We used both approaches for this thesis. This redundancy allows one to break down the genome into pieces with high LD, the aforementioned haplotype blocks that are not broken up by recombination and tend to transmit dozens or hundreds of SNPs on one stretch of DNA, so that it is impossible to directly determine which of the contained SNPs were causative for an associating trait or disease. By attaching a selected set of SNPs to each haplotype block, sufficient to explain the variance of the whole block, one can however capture the genetic variation of the whole genome, using only a limited number of SNPs. These SNPs are called tag or lead SNPs (and do not necessarily represent the linked causative SNPs nearby).²⁴

The International HapMap Project aimed to create a genome-wide coverage of LD and haplotype blocks of different ethnic groups. This was achieved by whole genome sequencing techniques that identified common SNPs and assigned their LD patterns. Their findings revealed that 80% of the common variation in the human genome can be captured by analysing only a subset of approximately 500 000 - 1 000 000 SNPs.²⁶ Later projects such as the 1000 Genomes project, encompassing larger sample sizes and more populations, are the current standard for LD and haplotype maps of the human genome.²³

1.2.3 The principle of “imputation” for identifying SNPs from microarray based GWAS

SNPs contribute to the genetic foundation of heritability of common complex diseases such as atherosclerosis. Therefore, they are the object of investigation in GWAS. The hypothesis-free approach of GWAS aims to detect SNPs that occur at different frequency in cases and controls, thereby associating genotype with phenotype. Complicating to pinpoint which SNPs cause atherosclerosis right away (also in this thesis), not all SNPs can technically be measured in each assay, and certain statistics-based inferences have to be made²⁴: As it would still be too expensive and time-consuming to sequence the whole genome of thousands of individuals, GWAS today still rely on the groundwork of the HapMap Project and the 1000 Genomes Project, which screened over 88 million genetic variants and provided data of LD for haplotype block lead SNPs and follow a principle termed “imputation”^{23,24}:

The development of microarray technology enabled research groups to use GWAS in a cost efficient and scalable way. The two primarily-used platforms are manufactured by Affymetrix (Santa Clara, CA) and Illumina (San Diego, CA). Affymetrix operates on solid-phase arrays where 25-mer oligonucleotides are attached to a surface, whereas Illumina uses 50-mer dyed oligonucleotides attached to polystyrene beads.²⁸ Differential hybridization results in detection of specific alleles on both platforms. Apart from the underlying technology, the selection of tagging SNPs is a crucial factor. Ideally, tagging SNPs on the microarray should reflect the LD structure of the study population in order to achieve the best possible genome coverage. Affymetrix chose to use tagging SNPs that are evenly spaced throughout the genome, not paying attention to population-related differences in LD structure.²⁸

This resulted in average genome coverage throughout different populations. Illumina specialized using selected tagging SNPs for the Caucasian population, based on LD data of the HapMap and the 1000 Genome Project, therefore achieving better genomic coverage in these individuals.²⁸ Up to the present date, both platforms offer arrays containing up to 5 million SNPs.²⁸

The principle of “imputation”: The low penetrance of common variants in complex diseases requires large sample sizes in order to detect statistically significant genetic effects. Therefore, research groups started to form consortia in order to combine their data. The use of different platforms using different tagging SNPs proved hard to combine. This was solved by a process called imputation. Hereby, alleles for untyped SNPs were estimated based on genotyped variants using LD data from the HapMap or 1000 Genomes Project as reference panels.²⁴ In the first step, redundant variants within shared haplotypes of the study sample are linked together (phasing). In the second step, resulting haplotypes are compared to a reference panel and missing variants in the study population are filled in.²⁴ Therefore, the reference panel and study sample should have a similar ethnic background, assuring similar LD patterns.²⁴

1.2.4 Statistical methods for ranking SNP-disease association from microarray-based GWAS

The need of large sample sizes with millions of analysed SNPs require appropriate statistical tests. In standard epidemiological statistical tests, a result is considered significant if the p-value is smaller than the predefined alpha value of typically 0.05. With millions of performed tests in GWAS, the likelihood of false positive results increases. Therefore, adjustments for multiple testing is performed. One commonly applied method is the Bonferroni correction where the significance threshold ($\alpha = 0.05$) is divided by the number of tests conducted ($\alpha = 0.05/k$).^{24,29} An alternative approach is to determine the false discovery rate FDR. The FDR estimates the proportion of false positive tests. By setting a threshold for FDR, the proportion of false positives are kept under a certain limit. Usually, FDR is set to 0.05.^{24,29}

The tagging SNPs used in GWAS indicate a correlation between genotype and phenotype. In fact, all SNPs in LD with the tagging SNP are correlated with the phenotype. The causal variant leading to the phenotype is called functional SNP but remains unspecified in the group of associated SNPs without further experiments. If, by coincidence, the functional SNP is also the genotyped SNP, it is referred to as direct association.³⁰ If the functional SNP is any SNP in high LD with the tagging SNP, it is referred to as indirect association.³⁰ The identified SNPs are, in most cases, not the main functional variant and additional studies are necessary to identify the causal variant.

There are thousands of GWAS available and in order to facilitate the access to them, the NHGRI-EBI Catalog of published GWAS (<https://www.ebi.ac.uk/gwas/>) was established. It systematically catalogues available GWAS and captured SNPs with their association, providing a uniform, searchable database for researchers worldwide. As of May 2019, the GWAS Catalog contains 3 989 publications which yield 138 312 SNP associations.

1.2.5 Importance of GWAS for studying atherosclerosis on the example of the chromosome 9p21 atherosclerosis risk locus

GWAS findings, in the best case, link biological pathways to diseases that have not been considered before in this way, thereby sparking hypothesis-driven studies to identify the underlying molecular and cellular mechanisms of disease. This improves the understanding of the pathophysiology and aetiology of the disease, which in turn provides new opportunities for prediction, prevention, and treatment, especially in the field of personalized medicine.¹⁶ We followed this principle in this thesis to focus on a novel risk locus, where no effector mechanism has been discovered so far.

To set this into context, in 2007, the first GWAS in the field of atherosclerosis were conducted. In the beginning, modest sample sizes of approximately 5 000 cases and controls revealed only nine loci associated with atherosclerosis related traits.³¹ In the following years, more GWAS with larger sample sizes added additional loci. At present, the formation of large international consortia, as well as denser genotyping added more statistical power to unravel previously unknown atherosclerosis loci. The alliance of GWAS, the CARDIoGRAMplusC4D consortium, allowed the analysis of over 180 000 individuals, unravelling 56 loci associated with atherosclerosis related traits.²² This number is ever growing and the number of known atherosclerosis loci exceeds 200 as of May 2019.³²⁻³⁵

Probably the most prominent and prototypical finding is the chromosome 9p21 locus, which was already identified in the first GWAS and thereafter was replicated in almost every other study. It still comprises one of the strongest genetic factors of atherosclerosis known. What makes the locus particularly interesting, is the fact that risk variants do not lie in protein-coding genes. In fact, the risk locus contains the lncRNA *ANRIL* and recent studies showed its role in modulating atherosclerosis susceptibility, e.g. influencing expression of inflammation-related genes elsewhere in the genome.³⁶ Further, it has been shown that besides regular linear *ANRIL* RNA, circular RNAs (circRNAs) are transcribed from the locus as well.³⁷ In the case of the 9p21 locus, there is a coexistence of linear *ANRIL* and circular *ANRIL*. The risk genotype of 9p21 is associated with differential *ANRIL* expression. While the risk allele is characterized by upregulation of linear *ANRIL* and downregulation of *circANRIL*, the protective allele shows inverse effects.^{37,38} Functional cell studies showed that linear *ANRIL* is associated with atherogenic reactions such as increased proliferation and decreased apoptosis, whereas *circANRIL* acts in an atheroprotective way by restraining proliferation and apoptosis (at least deduced from overexpression data).^{37,38} Thus, an unbalanced linear/circular *ANRIL* expression in favour of linear *ANRIL* is thought to be indicative for an atherogenic state.³⁸ This might be achieved by epigenetic regulation of gene expression in *cis* and *trans* via *ANRIL*, as shown by independent biochemical studies.³⁷ The underlying molecular mechanisms which would show how linear and *circANRIL* act are still being investigated in recent studies.

On a more general level it is important to note that most genetic variants identified in GWAS fall into non-coding regions of the human genome. Only 4.9% of SNPs lie in protein-coding regions, suggesting that the majority of effects leading to disease is caused by regulation of gene expression or

by noncoding-RNA-dependent regulation of RNAs and proteins, rather than by directly altering codons of protein-coding genes (and thus directly impacting structure and function of proteins through amino acid changes or truncations etc.).³⁹ However, knowledge of the exact functions and underlying pathophysiology and architecture of non-coding regions is still not satisfactory.

1.3 Functional elements in the human genome and the way to functional studies

1.3.1 Protein-coding and noncoding elements of the human genome

The human genome consists of over 3.2 billion base pairs (haploid) organized in 23 chromosomes. The genome can be divided into protein-coding and non-coding DNA sequences. A protein-coding sequence represents a gene with an ORF, that when transcribed into mRNA, can in turn be translated into protein. Less than 2% of the genome represents coding sequences, giving rise to roughly 20 000 protein-coding genes.^{40,41} By alternative splicing, the number of actual proteins is much higher than the number of underlying genes. The non-coding sequences represent more than 98% of the total genome. Noncoding sequences are diverse and range from noncoding intronic and untranslated regions of coding-genes, over eroded pseudogenes, repeats and many diverse heterochromatic noncoding sequences.^{40,41} But there are also transcribed genes located in the noncoding portion of the genome that do not have an ORF and can be divided into three major categories: genes transcribed into short non-coding RNA (short ncRNA), genes for long non-coding RNA (lncRNA) and pseudogenes.⁴² The group of short ncRNAs is defined by a length < 200 bases and includes miRNA, rRNA, siRNA, snRNA, snoRNA, and tRNA. The group of lncRNAs is defined by a length > 200 bases and can be further subgrouped into intergenic lncRNAs (same strand, convergent, divergent with respect to neighbouring genes) and genic lncRNAs (within exonic, intronic sequences and/or overlapping them).^{43,44}

1.3.2 Functions of long non-coding RNAs

Many lncRNAs have become known to carry out functions and are also a main topic of this thesis. Many play a regulatory role in controlling expression of other genes. Thereby they affect many processes of cell function, from X chromosome inactivation, through transitions in embryogenesis, to control of cell pluripotency, cell development, and differentiation.⁴⁵ The majority of lncRNAs are highly tissue specific and transcribed by RNA polymerase II. As other normal mRNAs, many lncRNAs are spliced and receive a methylated cap at the 5' end and a polyadenylated 3' end.⁴⁵ After processing, they are escorted to their place of action in the cytosol, in the nucleus, or in both. By numbers, the main mechanism of lncRNAs seems to be the modification of gene expression of target genes via guiding or assembling chromatin modifications or by other transcriptional and post transcriptional regulations.⁴⁵

In one example, chromatin regulation is in part directed by histone modifications, and chromatin remodelling, and the polycomb complexes PRC1 and PRC2 as major repressive regulators are regulated by RNA and by noncoding RNAs in particular: Various lncRNAs were found to recruit PRC to silence gene expression.^{46,47} Further, DNA methylation is achieved by association of lncRNAs with certain DNA methyltransferases, for which some lncRNAs act as scaffolds, guiding them to desired genomic regions. Chromatin remodelling is achieved by association of lncRNAs with SWI/SNF complexes, which affects how they adjust nucleosome positions, thereby altering access to the DNA.⁴⁷ Chromatin organization is altered by promoting chromosome looping between enhancer and promoter regions, thus influencing gene expression, and noncoding RNAs play a central role for enhancer functionality.⁴⁷

Even post transcriptional regulation (e.g. affecting splicing, mRNA stability, translation, and miRNA levels) can be affected by lncRNAs⁴⁷: Some lncRNAs intervene in alternative splicing by competing for binding sites of splicing regulatory proteins and influencing the distribution of splicing factors.⁴⁸ Overall mRNA stability can be influenced in both directions: First, perfect base pairing of lncRNA and mRNA can protect transcripts from decay. Second, partial base pairing of lncRNA and the 3' UTR of the mRNA can activate degradation of the transcript by Staufen mediated decay (SMD).⁴⁹ Translation can be influenced in both ways as well. First, inhibition of translation can be accomplished by imperfect base pairing of lncRNA and mRNA and subsequent recruitment of inhibitory translation factors.⁵⁰⁻⁵² Second, translation can be promoted by perfect base pairing of lncRNA and mRNA, resulting in recruitment of polysomes.⁵³ Finally, lncRNAs with sequences complementary to miRNAs can act as miRNA sponges, trapping them and thereby preventing them from binding to their targets such as mRNAs or transcription factors.⁵⁴ To summarize, there is strong evidence that SNPs, when located within noncoding RNAs, can engage in a number of effector pathways and ultimately end up also controlling the expression strength and dynamics of protein-coding genes. Therefore, noncoding RNAs should not be neglected as potential causal disease modifiers. Consequently, this thesis focused on a locus with large predicted effect size where risk SNPs do lie in noncoding regions.

1.3.3 Functions of long non-coding RNAs in atherosclerosis and related mechanisms

Atherosclerosis can be influenced by lncRNAs both globally by modulating risk factors such as dyslipidaemia, obesity, diabetes, or hypertension, as well as locally by intervening with various steps of the pathophysiological process.^{55,56}

Influence of lncRNAs on lipid levels, as a major risk factor for atherosclerosis, is achieved in a number of different ways: some lncRNAs were increasing cholesterol levels by inhibition of reverse transport of HDL or by increasing expression of HMG-CoA reductase.^{57,58} While some lncRNAs were found to increase triglyceride levels by promoting Acyl-CoA synthetase activity, other studies described lowered triglyceride levels in absence of a lncRNA by increased lipoprotein esterase activity.^{59,60} Taken together, disruption of lncRNA regulatory circuits can alter the balance of lipid levels.

Further, lncRNAs have been linked to metabolic alterations important for atherosclerosis: Linkages of obesity and lncRNAs mainly rely on associations, as pathophysiological mechanisms are poorly understood. Still, it has been revealed that loss of the lncRNA at Chr15q11-q13 leads to Prader-Willi syndrome and subsequently to an obese phenotype.⁶¹ Also, a lncRNA has been found to function as a coactivator of PPAR γ , one of the main regulators of adipogenesis.⁶² Indications of lncRNAs influencing insulin resistance and diabetes mellitus are for example based on the lncRNA H19, which acts on various miRNAs responsible for inactivating certain proteins.⁶³ This ultimately results in decreased insulin sensitivity, reduced glucose uptake, and impaired insulin signalling.⁶⁴ Also, GWAS led to discovery of a SNP in a lncRNA, associated with elevated fasting glucose levels.⁶⁵ Moreover, blood pressure is altered by lncRNAs influencing VSMC proliferation as well as angiotensin II response.^{66,67} Finally, GWAS discovered an association of the lncRNA ANRIL with systemic hypertension.⁶⁸ The exact underlying pathomechanisms are still poorly understood.

Last but not least, there were several studies showing differential expression of lncRNAs inside an atherosclerotic lesion, thus locally influencing mediator cells of the disease. There, lncRNAs influence these cells mainly by regulating functions like apoptosis, adhesion, proliferation, migration, or inflammatory state in cell types associated with atherosclerosis such as ECs, VSMCs or macrophages.⁵⁶ Figure 3 depicts some examples of lncRNAs and their effects in atherosclerosis.

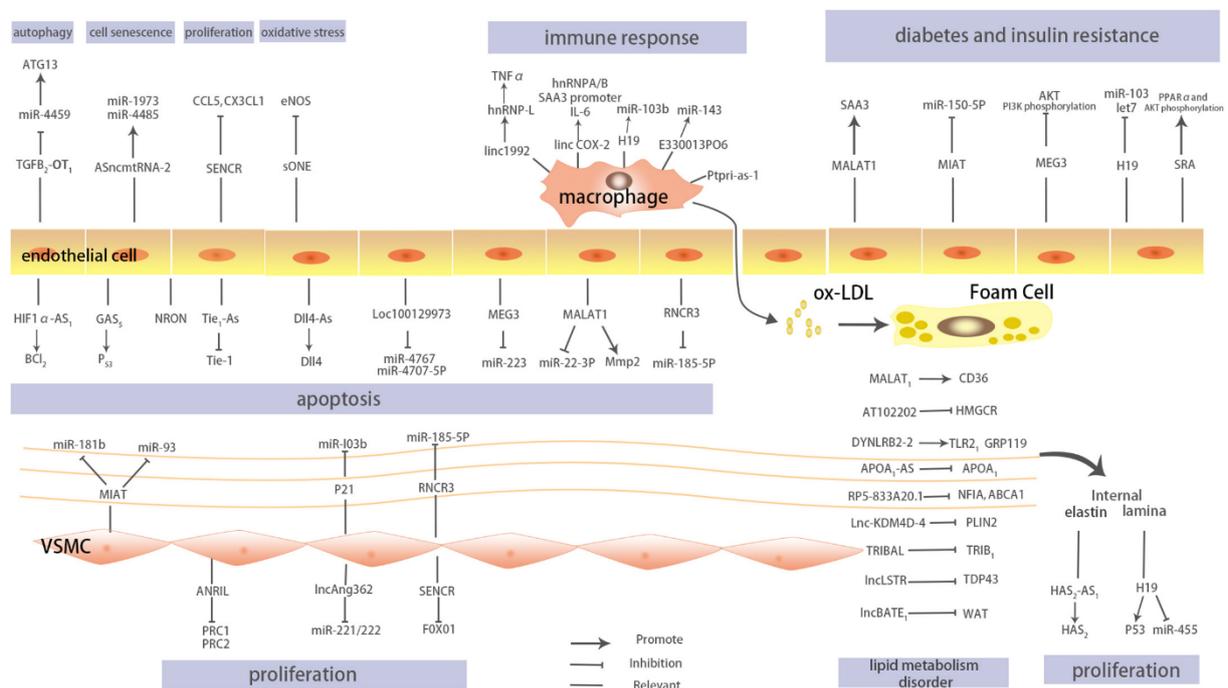


Figure 3 | Functions of lncRNAs in atherosclerosis. Effector cells of atherosclerosis (SMCs, ECs, macrophages/foam cells) expressing different lncRNAs with inhibitory and stimulating effects on apoptosis, proliferation, autophagy, cell senescence, immune response, diabetes and insulin resistance, and lipid metabolism disorder. Reprinted from Wang⁵⁶, Copyright 2018, with permission from Elsevier.

The observed effects can either be caused by gain or loss of function due to SNPs or as a result of altered expression levels of lncRNAs resulting from other causes.

1.3.4 From locus to biology: approaches and challenges of functional studies

Although GWAS were able to detect various loci associated with atherosclerosis, the identification of the causal variant, causal gene, the directionality of their effect and the underlying pathophysiological mechanisms still remains a challenging task. As GWAS detect only correlation, the causal SNP could be any variant in LD with the tagged SNP. With the majority of SNPs falling into non-coding regions, exerting regulatory functions *in cis* and *in trans*, the causal gene might as well be proximal or distal to the identified locus. Furthermore, there may be not only one, but multiple genes affected by the SNP, the interactions of which are ultimately responsible for the effect. Finally, the directionality (atherogenic/atheroprotective) and the cell type in which the effect is observed are important to be determined but can be context-specific.⁶⁹

One approach of associating genotype with gene expression levels are eQTL studies that are of relevance also for this thesis. Instrumental for eQTL analysis, changes in gene expression levels in a number of individuals (e.g. obtained by RNA-Seq) are statistically associated with a specific genotype or SNP they carry.⁷⁰⁻⁷² This method is able to detect the directionality of effects (SNP → mRNA level), and by screening of local and distant genes, can reveal both potential *cis* as well as *trans* effects of the examined SNPs.⁷³⁻⁷⁵ However, interindividual noise originating from different genetic backgrounds requires large sample sizes in order to detect significant effects.⁶⁹

Chromatin state, gene expression, and transcriptional regulation, especially for lncRNAs, are highly cell specific.^{69,70} Thus, for any of the stated analyses to test whether a given SNP is indeed causal, the appropriate cell type should be used for investigations. In samples originating from patients, there is often a heterogeneous mixture of cell types in different cell cycles and developmental stages, increasing background noise and therefore possibly masking subtle effects.^{72,73} In cell culture experiments, although well-established cell lines deliver homozygous and easy to handle material, immortalized cells represent a rather immature phenotype. The use of induced pluripotent stem cells (iPSCs) for cell culture experiments offers an ideal approach for reliable disease modelling, as they can be differentiated into any desired cell type associated with atherosclerosis.⁶⁹ For these reasons, we are using iPSCs and iPSC-derived ECs and VMSCs for this thesis.

Cell culture experiments are needed in order to validate GWAS findings and to assign a causal relationship to associated variants and downstream findings as gene expression. One powerful tool for investigation is genome editing, especially in the form of CRISPR/Cas9 technology, which was also applied in this thesis. The clustered regularly interspaced short palindromic repeats (CRISPR) and its guided nuclease, CRISPR-associated protein 9 (Cas9), are derived from a bacterial system and were adapted for experimental use. The CRISPR/Cas9 technology is capable of introducing double-strand breaks, repaired by either homology-directed repair (HDR) or non-homologous end joining (NHEJ) which creates less precise ends.⁷⁶ HDR can be used to introduce specific mutations by offering donor DNA as template, thereby mimicking the observed GWAS SNP. NHEJ can be used to introduce insertions or deletions of various length at the site of the double-strand breaks. In fact, introducing two double-strand breaks and allow for NHEJ can be used to create knockout models for desired large

genomic sites, as in this thesis, e.g. haplotype blocks of atherosclerosis. Nevertheless, genome editing using CRISPR/Cas9 technology remains a challenging task, as efficiency and tolerance, especially in iPSCs and primary cell lines, is quite low. Also, NHEJ is generally favoured over HDR and the underlying architecture of a genomic site can be hard to access by CRISPR/Cas9 because of certain sequence requirements that have to be fulfilled in order for a site to be potentially suitable for targeting. Though, if CRISPR/Cas9 modified iPSC lines are generated, they can be differentiated into desired somatic cells associated with atherosclerosis, such as SMCs or ECs. These can be used for downstream analysis, such as altered gene expression and functional experiments like adhesion, apoptosis, proliferation, and migration, which ought to be altered in atherosclerotic cells.⁷⁷ In this thesis, a new atherosclerosis risk locus was investigated in an iPSC model via CRISPR/Cas9-mediated deletion of relevant large sequences in the locus.

2

Aims of the thesis

It was the aim of this thesis to (i) create a comprehensive overview of atherosclerosis susceptibility loci from published atherosclerosis GWAS and, subsequently, to (ii) focus on an as yet unexplored atherosclerosis risk locus containing exclusively noncoding RNAs and to start exploring how risk was affected by genetic, molecular, and cellular experiments in a cell culture model. The central locus is termed the chromosome 21 atherosclerosis risk locus, and the thesis was therefore divided into two main subprojects.

2.1 Compilation of atherosclerosis susceptibility loci

The aim of the first part of this thesis was to create a comprehensive overview of published atherosclerosis susceptibility loci and to extract a promising locus containing lncRNAs for further experimental analysis in the second part of the thesis. Therefore, GWAS were analysed in order to assemble SNPs associated with atherosclerosis and its risk factors. The task was not trivial as information is spread out over many studies and laid down in different formats and is based on different populations and controls. To classify the vast number of atherosclerosis risk SNPs into a smaller number of functional units, a subsequent definition of haplotype blocks as equivalents for atherosclerosis susceptibility loci was performed and SNPs were allocated to their haplotype blocks. Finally, colocalization with risk SNPs and the genetic content of the loci was performed, to narrow down the number of candidate loci for further experimental studies and to pinpoint as yet uncharacterized atherosclerosis risk loci with large effect size that did not contain any protein-coding genes as possible primary effectors.

2.2 Experimental study of a selected atherosclerosis locus

The aim of the second part was to explore the possible effector mechanism by which a locus selected in the first part of the thesis affects atherosclerosis. This locus came out as one of the top regions of atherosclerosis risk and was located on chromosome 21, containing 2 lncRNAs as potential primary effectors. Therefore, knockouts of the sequences corresponding to the relevant atherosclerosis haplotype

block on chromosome 21 were generated in hiPSCs serving to establish a cellular framework where cells of otherwise identical genetic background could be used to test the specific effects of the knockout on the differentiation of iPSC-derived atherosclerosis-relevant VSMCs and ECs. The role of the lncRNA and genes surrounding the locus were investigated by qPCR analysis and eQTL analysis and cellular functions known to be involved in atherogenesis, like apoptosis, proliferation, migration, and adhesion were explored when studying the knockout cell lines.

3

Materials

3.1 Laboratory Equipment

4D-Nucleofector System.....	Lonza
CERTOMAT R	Sartorius BBI Systems
FlashGel Dock.....	Lonza
FlashGel Systems	Lonza
Fusion-FX7 imaging system	Vilber Lourmat
IncuCyte Cell Migration Software.....	Essen BioScience
IncuCyte ZOOM.....	Essen BioScience
LightCycler 480 II	Roche Molecular Diagnostics
MACS Seperator.....	Miltenyi Biotec
MoFlo Astrios EQ.....	Beckman Coulter
NanoDrop 2000c UV-Vis spectrometer.....	ThermoFisher Scientific
SpectraMax Paradigm.....	Molecular Devices
Thermomixer comfort.....	Eppendorf
Veriti 96-Well Thermal Cycler.....	Applied Biosystems
ViiA 7 Real-Time PCR System.....	ThermoFisher Scientific
WoundMaker.....	Essen BioScience

3.2 Chemicals and Consumables

360 GC Enhancer.....	ThermoFisher Scientific
AccuPrime PCR Buffer II (10X)	ThermoFisher Scientific
AccuPrime Taq DNA Polymerase, High Fidelity (5 U/ μ l)	ThermoFisher Scientific
Accutase	STEM CELL Technologies
Activin A.....	R&D Systems
AmpliTaq Gold 360 Master Mix	ThermoFisher Scientific
AmpliTaq Gold DNA Polymerase (5 U/ μ l)	ThermoFisher Scientific
autoMACS Rinsing Solution.....	Miltenyi Biotec
B-27 Supplement without Vitamin A	ThermoFisher Scientific

BbsI (10000 U/ml).....	New England BioLabs
BigDye Terminator 3.1 5X Sequencing Buffer.....	Applied Biosystems
BigDye Terminator 3.1 Ready Reaction Mix.....	Applied Biosystems
BSA (10 mg/ml).....	New England BioLabs
Buffer AE.....	Qiagen
Buffer AL.....	Qiagen
Buffer AW1.....	Qiagen
Buffer AW2.....	Qiagen
Buffer B (10X).....	Solis BioDyne
Buffer P1.....	Qiagen
Buffer P2.....	Qiagen
Buffer P3.....	Qiagen
Buffer QBT.....	Qiagen
Buffer QC.....	Qiagen
Buffer QF.....	Qiagen
Casein Yeast Magnesium Broth.....	Sigma-Aldrich
Caspase-Glo Reagent.....	Promega
CHIR99021.....	R&D Systems
Chloroform 100%.....	Sigma-Aldrich
Collagen.....	Sigma-Aldrich
Complete Lysis Solution.....	5-prime
DEPC water.....	Sigma-Aldrich
Diluted Wash Buffer.....	5-prime
DMEM GlutaMAX-I.....	ThermoFisher Scientific
DMEM/F-12.....	ThermoFisher Scientific
DMSO 100%.....	Sigma-Aldrich
DNA Binding Buffer.....	Zymo Research
DNA Clean & Concentrator-5 Kit.....	Zymo Research
DNA Loading Dye (6X).....	ThermoFisher Scientific
DNA Wash Buffer.....	Zymo Research
DNeasy Blood & Tissue Kit.....	Qiagen
DNeasy Mini spin column.....	Qiagen
dNTPs (10 mM).....	Promega
DTT (0.1 M).....	ThermoFisher Scientific
EDTA (0.5 M, pH = 8.0).....	Sigma-Aldrich
Elution Buffer.....	5-prime
Endothelial Cell Growth Medium.....	PromoCell
Ethanol.....	Merck Chemicals

Ethidium bromide (10 mg/ml).....	Serva
ExoSAP-IT	ThermoFisher Scientific
Falcon Cell Strainer	ThermoFisher Scientific
FBS	Lonza, Biochrom
Fibronectin from human plasma.....	Sigma-Aldrich
FIREPol DNA Polymerase.....	Solis BioDyne
First-Strand Buffer (5X).....	ThermoFisher Scientific
FlashGel DNA Cassette	Lonza
FlashGel Loading Dye (1X).....	Lonza
Forskolin	R&D Systems
Gelatine	Sigma-Aldrich
GeneRuler DNA Ladder Mix	ThermoFisher Scientific
Gentamycin sulfate/amphotericin-B (GA-1000)	Lonza
Gentle Cell Dissociation Reagent.....	STEM CELL Technologies
GlutaMAX Supplement	ThermoFisher Scientific
H ₂ O	ThermoFisher Scientific
Isopropyl alcohol.....	Merck Chemicals
KAPA SYBR FAST 50X ROX High Reference Dye	Kapa Biosystems
KAPA SYBR FAST qPCR Master Mix (2X) Kit	Kapa Biosystems
KAPA SYBR FAST qPCR Master Mix (2X) Universal	Kapa Biosystems
Low adhesion tubes	Sigma-Aldrich
MACS BSA Stock Solution	Miltenyi Biotec
MACS columns	Miltenyi Biotec
Matrigel hESC-qualified Matrix.....	Corning
MgCl ₂ (0.15 M).....	Integrated DNA Technologies
MgCl ₂ (1 M).....	Merck Chemicals
MgCl ₂ (25 mM).....	ThermoFisher Scientific
MiniPrep Kit.....	5-prime
mTeSR 1 5X supplement	STEM CELL Technologies
mTeSR 1 medium.....	STEM CELL Technologies
N-2 Supplement.....	ThermoFisher Scientific
NaCl (5 M).....	Merck Chemicals
NEBuffer 2.1 (10X).....	New England BioLabs
Neurobasal Medium.....	ThermoFisher Scientific
Nucleofector Solution (Amaxa P3 Primary Cell Kit)	Lonza
One Shot Chemically Competent E. coli	ThermoFisher Scientific
Opti-MEM I Reduced Serum Medium	ThermoFisher Scientific
PBS Dulbecco.....	Biochrom

pCas9_GFP	Kiran Musunuru
PCR Buffer (10X).....	ThermoFisher Scientific
pCR II-TOPO vector.....	ThermoFisher Scientific
PEG 40% (w/w).....	Sigma-Aldrich
PenStep (10000 U/ml Penicillin, 10000 µg/ml Streptomycin).....	ThermoFisher Scientific
pGuide.....	Kiran Musunuru
Plasmid Maxi Kit.....	Qiagen
Pre-Separation Filters.....	Miltenyi Biotec
Proteinase K (600 mAU/ml).....	Qiagen
Qiagen-tip column	Qiagen
Random hexamer primers p(dN)6 (500 µg/ml).....	Roche
Recombinant RNAsin Ribonuclease Inhibitor (40 U/µl).....	Promega
rh bFGF	R&D Systems
rh EGF.....	Lonza
rh FGF-B	Lonza
rh insulin	Lonza
rh PDGF-BB.....	ReproTech
rh VEGF	R&D Systems
ROCK inhibitor (Y-27632 dihydrochloride, 10 mM).....	Santa Cruz Biotechnology
RPMI medium	Biochrom
Single Nucleocuvette	Lonza
SmBM.....	Lonza
SmGM-2.....	Lonza
Sodium acetate (3 M).....	Sigma-Aldrich
Staurosporine (0.25 µM).....	Merck Chemicals
StemPro-34 medium	ThermoFisher Scientific
StemPro-34 Nutrient Supplement.....	ThermoFisher Scientific
Stop Solution.....	Integrated DNA Technologies
SuperScript II Reverse Transcriptase (200 U/µl)	ThermoFisher Scientific
Supplement (Amaxa P3 Primary Cell Kit).....	Lonza
Surveyor Enhancer S.....	Integrated DNA Technologies
Surveyor Nuclease S.....	Integrated DNA Technologies
T4 DNA Ligase (400000 U/ml).....	New England BioLabs
T4 DNA Ligase Buffer (10X)	New England BioLabs
T4 Polynucleotide Kinase 10000 U/ml.....	New England BioLabs
TAE.....	ThermoFisher Scientific
TE buffer (1X).....	G-Biosciences
TOPO TA Cloning Kit.....	ThermoFisher Scientific

TransIT-2020 Reagent	Mirus Bio
Tris (1 M, pH = 8.0).....	Carl Roth
TRIzol Reagent.....	ThermoFisher Scientific
Trypsin/EDTA Solution (0.05%/0.02%).....	Biochrom
UltraPure Agarose	ThermoFisher Scientific
Vacutainer CPT Mononuclear Cell Preparation Tubes.....	BD Biosciences
XL10-Gold Ultracompetent cells (Escherichia coli)	Stratagene
Zymo-Spin Column	Zymo Research
β -mercaptoethanol	ThermoFisher Scientific, Stratagene

4

Methods

4.1 Molecular Biological Methods of DNA

4.1.1 Isolation and purification of DNA

Genomic DNA was isolated from HEK-293 and hiPSC cells using the DNeasy Blood & Tissue Kit from Qiagen. Cells were removed from the plates using PBS Dulbecco followed by centrifugation at 300x g for 5 minutes. The supernatant was aspirated, and the cell pellet was resuspended in 200 μ l PBS Dulbecco. Lysis was achieved in lysis buffer under addition of 20 μ l Proteinase K (600 mAU/ml). After adding 200 μ l Buffer AL, the lysate was mixed and incubated at 56 °C for 10 minutes. 200 μ l 100% ethanol was added and the mixture was pipetted into the DNeasy Mini spin column placed in a 2 ml collection tube. After centrifugation at 6000x g for 1 minute, the flow-through was discarded. The DNeasy Mini spin column membrane was washed with 500 μ l Buffer AW1, followed by centrifugation at 6000x g for 1 minute. A second wash-step with 500 μ l Buffer AW2 and centrifugation at 20000x g for 3 minutes was performed. The flow-through was discarded and the DNeasy Mini spin column was placed in a clean collection tube. After pipetting 200 μ l Buffer AE onto the DNeasy Mini spin column membrane, the DNA was eluted by centrifugation at 6000x g for 1 minute and stored at 4 °C for further analyses.

The PCR products were purified using the DNA Clean & Concentrator-5 Kit from Zymo Research. In an Eppendorf tube, 5 volumes of DNA Binding Buffer were added to 1 volume of PCR product. The mixture was transferred into a Zymo-Spin Column in a collection tube. After centrifugation at 10000x g for 30 seconds, the flow-through was discarded. The columns were washed twice by adding 200 μ l DNA Wash Buffer and centrifugation at 10000x g for 30 seconds. The columns were placed onto new Eppendorf tubes. The DNA was eluted by adding 10.5 μ l H₂O, incubation at room temperature for 1 minute and centrifugation at 10000x g for 30 seconds. The purified DNA was stored at 4 °C for further analyses.

4.1.2 Analysis of concentration and purity of nucleic acids

The nucleic acid samples were analysed for concentration and quality using the NanoDrop 2000c UV-Vis spectrometer from ThermoFisher Scientific. Blank measurements were performed in solutions similar to solved nucleic acids. Concentrations were determined by measuring absorbance at $\lambda = 260$ nm with baseline normalisation at $\lambda = 340$ nm. The purity was determined by calculation of absorbance

ratios. A 260/280 ratio of ~ 1.8 for DNA and ~ 2.0 for RNA was accepted as pure and deviations indicate a contamination in form of proteins, phenols or other particles that absorb strongly at or near $\lambda = 280$ nm. As a second measurement, a 260/230 ratio of 1.8-2.2 was accepted as pure and deviations indicate a contamination in form of guanidine, magnetic beads, carbohydrates, or other particles that absorb strongly at or near $\lambda = 230$ nm.

4.1.3 Polymerase chain reaction and gel electrophoresis

The amplification of desired DNA sequences and was achieved by PCR, using AmpliTaq Gold 360 from ThermoFisher Scientific according to Tables 1 and 2.

Table 1 | PCR reaction mix

Reagent	Final concentration
AmpliTaq Gold 360 Master Mix	1X
360 GC Enhancer, optional	4%
Forward Primer	0.2 μ M
Reverse Primer	0.2 μ M
Template DNA	10 ng/ μ l

Table 2 | PCR thermal cycling conditions

Step	Temperature [°C]	Time [min:sec]	Cycles
Initial Denaturation	95	10:00	1
Denaturation	95	00:30	40
Annealing	primer-specific see Supplementary Table 1-4	00:30	
Extension	72	01:00	
Final Extension	72	07:00	1
Hold	4	∞	1

For DNA templates harbouring GC-rich regions or producing nonspecific products, 360 GC Enhancer was used in a final concentration of 4%.

The design of the primers for amplification was done with Primer3 software. The synthesis of the desired primers was performed by Eurofins Genomics in high purity salt free quality. Sequences and annealing temperatures of all primers are given in Supplementary Tables 1-4.

Gel electrophoresis was used to separate and identify PCR products. This was achieved by using agarose gel electrophoresis and FlashGel Systems by Lonza.

Agarose gels were cast with UltraPure Agarose in densities ranging between 0.7% and 3.0%, depending on the size of the fragment aimed for. TAE was used as a buffer in a final concentration of 40 mM Tris, 20 mM acetic acid and 1 mM EDTA. The DNA samples were loaded using 6X DNA

Loading Dye. The size of the fragments was determined by means of GeneRuler DNA Ladder Mix. Electrophoresis was run at 9 V/cm for 45-60 minutes and afterwards was stained ethidium bromide for 5 minutes. If a higher contrast was required, the gels were additionally destained by resting in TAE buffer for 10 minutes.

FlashGel DNA Cassettes with 2.2% Agarose were used for fragments with short size (< 100 bp) or low expected concentrations. Precasted and prestained cassettes were flooded with 10 µl H₂O per well. The DNA fragments were loaded with 1X FlashGel Loading Dye and pipetted into the wells. Electrophoresis was run at 275 V for 7 minutes in FlashGel Dock.

The finished agarose gels, as well as the FlashGel Systems, were photographed using the Fusion-FX7 imaging system with UV-transilluminator.

4.1.4 Genotyping

4.1.4.1 Genotyping using polymerase chain reaction

Identification of heterozygous and homozygous knockout cells was achieved by using the KAPA SYBR FAST qPCR Master Mix (2X) Kit from Kapa Biosystems according to Tables 3 and 4.

Table 3 | PCR reaction mix

Reagent	Final concentration
KAPA SYBR FAST qPCR Master Mix (2X) Universal	1X
KAPA SYBR FAST 50X ROX High Reference Dye	1X
Purified BSA (10 mg/ml)	1 µg/µl
Forward Primer	100 nM
Reverse Primer	100 nM
Template DNA	< 1 ng/µl

Table 4 | PCR thermal cycling conditions

Step	Temperature [°C]	Time [min:sec]	Cycles
Initial Denaturation	95	03:00	1
Denaturation	95	00:10	40
Annealing	see Supplementary Table 1	00:30	
Extension	72	00:05	
Melting curve	95	00:15	1
	60	01:00	1
	95	00:15	1
Hold	4	∞	1

The design of primers for amplification was done with Primer3 software. The synthesis of desired primers was performed by Eurofins Genomics in high purity salt free quality. Sequences and annealing temperatures of all primers are given in Supplementary Tables 1-4.

The designed primers were positioned 5' upstream and 3' downstream from the knockout region, forming a product only if the large region had been successfully deleted. A successful knockout resulted in generation of a PCR fragment with 503 bp for knockout 1 and 624 bp for knockout 2, respectively. There was no PCR product observed in case of an unsuccessful knockout, as the section of amplification would be 130 kb or 58 kb for knockout 1 or knockout 2, respectively. The initial identification of positive knockouts was achieved by qualitative detection of fluorescence signals of SYBR Green Dye I contained in KAPA SYBR FAST qPCR Master Mix (2X) on a ViiA 7 Real-Time PCR System. Verification was achieved by electrophoresis of PCR fragments using a 2% agarose gel, followed by imaging with Fusion-FX7 as described.

4.1.4.2 Genotyping using melt curve-based analysis

Genotypes of cultured cells regarding the atherogenic SNP rs9982601 were determined by melting curve analyses. The principle of the method was to run a PCR reaction, followed by a melting curve with a single dual-labeled probe. As a reporter, FAM was attached to the 5' terminus and as a quencher TAMRA was attached to the 3' terminus of the probe. As opposed to classical TaqMan probes, probes in this method were designed with a melting point (T_m) 10 °C below T_m of used primers to prevent participation during PCR and thus hydrolysis. After PCR, hybridization of the probe was permitted, followed by gradual increase of temperature during which the fluorescence signal was detected continuously. The sequence of the probe was designed for perfect matching with the wild type allele, thus dissociating at lower temperature when bound to mutant allele because of mismatch binding. After detachment of the probe, the mean distance between reporter and quencher was shortened, resulting in a diminished fluorescence signal due to fluorescence resonance energy transfer (FRET). Therefore, melting curve analysis showed a first peak of difference in fluorescence emission ($-dF/dT$) as the probe dissociated from the mutant allele, followed by a second peak about 10 °C later, as the probe dissociated from the wild type allele. PCR, and melting curve analysis with fluorescence detection was performed on a LightCycler 480 II from Roche Molecular Diagnostics.

Amplification was achieved by using FIREPol DNA Polymerase as described in Tables 5 and 6. The concentrations of the forward and the reverse primer differed in order to shift amplification in favour of the probe-binding strand.

Table 5 | PCR reaction mix

Reagent	Final concentration
FIREPol DNA Polymerase	0.05 U/ μ l
10X Buffer B	1X
25 mM MgCl ₂	4 mM
10 mM dNTPs	250 μ M
Purified BSA	500 ng/ μ l
Forward Primer	100 nM
Reverse Primer	900 nM
Probe	100 nM
Template DNA	< 10 ng/ μ l

Table 6 | PCR and melt curve thermal cycling conditions

Step	Temperature [°C]	Time [min:sec]	Cycles
Initial Denaturation	95	03:00	1
Denaturation	95	00:20	40
Annealing	see Supplementary Table 1	00:30	
Extension	72	01:00	
Final Extension	72	07:00	1
Melting curve	95	03:00	1
	30	02:30	
	90	00:15	
Hold	4	∞	1

The design of primers for amplification was done with Primer3 software and the design of probes was done with MeltCalc software. The synthesis of desired oligonucleotides was performed by Eurofins Genomics in high purity salt free quality. Sequences and annealing temperatures of all primers and probes are given in Supplementary Table 1.

4.1.5 Sanger sequencing

Sequencing was performed to confirm the exact genotypes of amplified PCR fragments and the correct insertion of ligated oligonucleotides in plasmids. Prior to sequencing, DNA products were purified by removing excess primers and nucleotides with ExoSAP-IT. To this end, 7 μ l DNA was mixed with 1.3 μ l ExoSAP-IT, followed by incubation at 37 °C for 15 minutes for degradation. After treatment, the enzyme was inactivated by incubation at 80 °C for 15 minutes. The purified DNA was next used for sequencing reaction with BigDye Terminator 3.1 Ready Reaction Mix according to Tables 7 and 8.

Table 7 | PCR reaction mix

Reagent	Final concentration
BigDye Terminator 3.1 Ready Reaction Mix	0.15X
BigDye Terminator 3.1 5X Sequencing Buffer	0.75X
Sequencing primer (forward or reverse)	500 nM
Template DNA	10-50 ng/ μ l

Table 8 | PCR thermal cycling conditions

Step	Temperature [°C]	Time [min:sec]	Cycles
Initial Denaturation	96	01:00	1
Denaturation	96	00:10	25
Annealing	60	00:05	
Extension	60	04:00	
Hold	4	∞	1

The design of primers for sequencing reaction was done with Primer3 software. The synthesis of desired primers was performed by Eurofins Genomics in high purity salt free quality. Sequences and annealing temperatures of all primers are given in Supplementary Tables 1-4.

After the sequencing reaction, amplified products were purified by ethanol precipitation. Precipitation was achieved by addition of 10 μ l H₂O, 2 μ l 3 M sodium acetate (pH = 5.2), and 55 μ l 100% ethanol. The resulting mixture was centrifuged at 14000x g for 20 minutes. The supernatant was discarded, and remaining salts and ethanol were removed by washing with 180 μ l 70% ethanol, followed by centrifugation at 14000x g for 10 minutes. After removing the supernatant, the pellet was dried at 68 °C for 10 minutes. Samples were submitted to Eurofins Genomics for subsequent sanger sequencing.

4.1.6 Quantitative Real Time PCR

Quantitative Real Time PCR (qRT-PCR) was performed in order to analyse RNA expression of genes of interest (as cDNA) and amounts of sample DNA for genotyping, respectively. Absolute quantification of gene expression was achieved by using standard curve method. Relative quantification of sample DNA was achieved by using internal standards as control.

4.1.6.1 Absolute quantification of gene expression using standard curve method

The method is based on the FRET phenomenon: a fluorochrome can be stimulated by light of certain wavelength S1, resulting in emission of light with another wavelength E1. If there is a second fluorochrome which can be stimulated by light of wavelength E1 and is close to the first fluorochrome, E1 will be absorbed by the second fluorochrome, resulting in emission of light with wavelength E2. Therefore, besides regular primers for PCR, dual-labelled TaqMan probes spanning across exon-exon

boundaries of target genes were designed. The TaqMan probes were positioned between forward and reverse primers. Labelling was accomplished by attachment of FAM as reporter dye at the 5' terminus and TAMRA as quencher dye at the 3' terminus of the probe. The AmpliTaq Gold DNA Polymerase used for amplification during PCR has 5'-3'-exonuclease activity. As long as the probe is intact, light of wavelength E1 is not detectable. As soon as exonuclease activity releases the reporter dye (FAM), the distance to the quencher (TAMRA) increases, resulting in the increased emission of light with wavelength E1. The emission increases proportional to the released probes and therefore to the amount of cDNA and genes, respectively. All absorbed and emitted wavelengths are highly characteristic and were measured using the ViiA 7 Real-Time PCR System. Quantification was performed by determination of a cycle threshold (CT), which is the number of cycles when the measured fluorescence signal is getting detectable above the background. Parallel measurement of cDNA with known quantities and their CT allowed establishment of a standard curve which can be used for determining quantity of cDNA copies by comparison of CT values.

The design of primers for amplification was done with Primer3 software and the design of probes was done with MeltCalc software. The synthesis of desired oligonucleotides was performed by Eurofins Genomics in high purity salt free quality. Sequences and annealing temperatures of all primers and probes are given in Supplementary Tables 1-4.

Quantification standards were established by amplification of targeted regions, followed by cloning into pCR II-TOPO vector (ThermoFisher Scientific). Cloning was performed using TOPO TA Cloning Kit from ThermoFisher Scientific. Therefore, a mixture of 4 µl PCR product of the target gene, 1 µl Salt Solution, 1 µl TOPO vector and 1 µl H₂O was incubated at room temperature for 5 minutes and subsequently chilled on ice. Next, 4 µl of the resulting mixture was added to a tube of One Shot Chemically Competent E. coli (ThermoFisher Scientific), followed by incubation at room temperature for 5 minutes. 50 µl of transformed cells were then spread onto LB-ampicillin agar plates (100 µg/ml ampicillin). Minipreparation, sequencing, and Maxipreparation were subsequently performed as described. After confirmation of correct plasmid standards, standard curves were diluted, ranging from 10² to 10⁷ copies. QRT-PCR was performed in 25 µl reactions according to Tables 9 and 10.

Table 9 | qRT-PCR reaction mix

Reagent	Final concentration
AmpliTaq Gold DNA Polymerase (5 U/µl)	25 mU/µl
10X PCR Buffer	1X
25 mM MgCl ₂	5 mM
10 mM dNTPs	800 nM
Forward Primer	300 nM
Reverse Primer	300 nM
Probe	200 nM
cDNA	< 100 ng/µl

Table 10 | qRT-PCR thermal cycling conditions

Step	Temperature [°C]	Time [min:sec]	Cycles
Initial Denaturation	95	10:00	1
Denaturation	95	00:15	45
Annealing and Extension	60	01:00	
Final extension	60	07:00	1

4.1.6.2 Relative quantification of DNA using internal standards

Relative quantification of DNA was performed for determination of heterozygous and homozygous status of potential knockout cells, using KAPA SYBR FAST qPCR Master Mix (2X) Kit from Kapa Biosystems. The designed primers were positioned within the knockout sequence. Thus, if the knockout was realized on both alleles (homozygous), no template for PCR was available, resulting in no PCR product. For heterozygous clones, one allele served as template and for wild type cells, both alleles contributed to PCR as templates. For each sample, the exact same amount of DNA was used for amplification. So, differences in resulting amounts of PCR fragments were due to genotype status and could be used for determination of heterozygous and homozygous clones.

KAPA SYBR FAST qPCR Master Mix (2X) Universal contained SYBR Green I fluorescent dye. SYBR Green I was intercalating between double-stranded DNA, emitting a fluorescence signal on binding, measured by ViiA 7 Real-Time PCR System. Amplification was performed in 10 μ l reaction volume in quadruplicates according to Tables 11 and 12. The design of primers for amplification was done with Primer3 software. The synthesis of desired oligonucleotides was performed by Eurofins Genomics in high purity salt free quality. Sequences and annealing temperatures of primers are given in Supplementary Tables 1-4.

Table 11 | qRT-PCR reaction mix

Reagent	Final concentration
KAPA SYBR FAST qPCR Master Mix (2X) Universal	1X
KAPA SYBR FAST 50X ROX High Reference Dye	1X
Forward Primer	200 nM
Reverse Primer	200 nM
Template DNA	1 ng/ μ l

Table 12 | qRT-PCR reaction mix

Step	Temperature [°C]	Time [min:sec]	Cycles
Initial Denaturation	95	03:00	1
Denaturation	95	00:10	40
Annealing	60-62	00:30	
Extension	72	00:05	
Melting curve	95	00:15	1
	60	01:00	1
	95	00:15	1
Hold	4	∞	1

For each sample, C_T values were generated. The median of C_T was calculated. S100 PCR was used as reference for experimental PCRs described above. Analysis of genotype was performed using normalized expression fold change $\Delta\Delta CT$ method, calculated as follows:

$$\Delta CT_{sample} = \overline{CT}_{experimental\ sample} - \overline{CT}_{reference\ sample}$$

$$\Delta CT_{control} = \overline{CT}_{experimental\ control} - \overline{CT}_{reference\ control}$$

$$\Delta\Delta CT = \Delta CT_{sample} - \Delta CT_{control}$$

$$normalized\ expression\ fold\ change = 2^{-\Delta\Delta CT}$$

Normalized expression fold change values ~ 1 indicated wild type samples while values ~ 0.5 indicated heterozygous samples and values < 0.5 suggested homozygous samples.

4.1.7 Production of plasmids for CRISPR/Cas9 systems

The CRISPR technology is capable of introducing precise double-strand breaks. CRISPR/Cas9 systems consist of two parts: a guide RNA (gRNA) and the CRISPR-associated protein 9 (Cas9).

The gRNA can be divided into two sequences, starting with the CRISPR RNA (crRNA) at the 5' end and the transactivating CRISPR RNA (tracrRNA) at the 3' end. The crRNA starts with a sequence of 20 nt, also known as protospacer, which can be altered as needed in order to target the desired genomic site. Both crRNA and tracrRNA complex with Cas9, the nuclease responsible for setting double-strand breaks. The 20 nt protospacer sequence at the 5' end of the gRNA directs Cas9 to a specific DNA target site by complementary RNA-DNA base pairing rules. A cut only occurs if certain short sequences, known as protospacer adjacent motifs (PAMs), are located immediately 3' of the sequence to be cleaved. The required PAM sequence was 5'-NGG. Therefore, it is possible to target any genomic site in form of N20-NGG by altering the first 20 nt of the gRNA to match the targeted DNA sequence.

Cas9 and gRNA both were delivered into cells in form of plasmids. Cas9 plasmid (pCas9_GFP, in cooperation with Kiran Musunuru) was containing a CAG promoter and coexpressing GFP for

subsequent FACS sorting. The plasmid for gRNA (pGuide, in cooperation with Kiran Musunuru) contained a U6 promoter followed by an incomplete gRNA sequence. The missing 20 bp protospacer had to be inserted individually. Usage of two different gRNAs for targeting two specific genomic regions at once resulted in two double-strand breaks yielding a fragment. Through NHEJ, a knockout was created.

4.1.7.1 Design of guide RNA plasmids

Individual gRNAs were produced for every genomic site targeted. Sequences surrounding the desired cutting site were extracted from UCSC Genome Browser (hg19). Corresponding sequences were entered into CRISPR Design Tool (<http://crispr.mit.edu/>) to search for optimal protospacers. Care was taken to obtain the protospacers physical structure of 5'-G(N)20-3', as the U6 promoter of the gRNA plasmid needed a guanine to start transcription. The targeted genomic strand was scanned for 5'-NGG-3' immediately next to the protospacer, as PAM is a requirement for nuclease activity. Additionally, the aim was to stay near the desired cutting site, while keeping off-target effects low. In order to correspond to sticky ends of the open pGuide plasmid, 5'-CACC-3' was added at the 5' end of the oligonucleotide representing the protospacer, while 5'-AAAC-3' was added at the 5' end of the complement oligonucleotide. The synthesis of desired oligonucleotides was performed by Eurofins Genomics in high purity salt free quality. Sequences of all oligonucleotides are given in Supplementary Table 7.

4.1.7.2 Generation and transformation of guide RNA plasmids in *Escherichia coli*

Generation of gRNA plasmids was achieved by annealing corresponding oligonucleotides, followed by ligation into empty pGuide vector. First, the empty pGuide vector was opened by digestion with BbsI according to Table 13.

Table 13 | Reaction mix for digestion with BbsI

Reagent	Final concentration
BbsI (10 000 U/ml)	1.5 U/ μ l
10X NEBuffer 2.1	1X
Empty pGuide	500 ng/ μ l

The digestion was performed at 37 °C overnight, followed by heat inactivation at 65 °C for 30 minutes. Second, 10X Annealing Buffer was mixed according to Table 14 for subsequent annealing.

Table 14 | Composition of Annealing Buffer

Reagent	Final concentration
1 M Tris, pH = 8.0	0.4 M
1 M MgCl ₂	0.2 M
5 M NaCl	0.5 M
0.5 M EDTA, pH = 8.0	0.01 M

Besides Annealing Buffer in a final concentration of 1X, the master mix for annealing contained corresponding oligonucleotides for protospacers in a final concentration of 100 nM each. Annealing was performed with Veriti 96-Well Thermal Cycler by heating up to 95 °C for 3 minutes and then cooling down to 20 °C with 0.5% of maximal cooling ramp rate. As ligation requires phosphate residues at the 5' end of annealed oligonucleotides, a kinase reaction according to Table 15 had to be performed.

Table 15 | Reaction mix for kinase reaction

Reagent	Final concentration
T4 Polynucleotide Kinase 10 000 U/ml	0.4 U/μl
10X T4 DNA Ligase Buffer	1X
Annealed oligonucleotides	80 nM

The mixture was incubated at 37 °C for 30 minutes, followed by heat inactivation at 65 °C for 20 minutes. The finished insert was subsequently ligated into the empty and open pGuide vector by incubation at room temperature for 30 minutes according to Table 16.

Table 16 | Reaction mix for ligation

Reagent	Final concentration
T4 DNA Ligase (400 000 U/ml)	40 U/μl
10X T4 DNA Ligase Buffer	1X
40% (w/w) PEG	4%
Empty and open pGuide vector	2 ng/μl
Kinase-annealed oligonucleotides	52 nM

Transformation of generated pGuide plasmid into *E. coli* was performed subsequently with XL10-Gold Ultracompetent cells (*E. coli*) from Stratagene. NZY+ medium was created by mixing 210 mg Casein Yeast Magnesium Broth with 10 ml H₂O and incubation at 37 °C until needed. In an Eppendorf tube, 50 μl of thawed *E. coli* were mixed with 4 μl β-mercaptoethanol and afterwards incubated on ice for 10 minutes. Then, 10 ng finished pGuide were added, followed by incubation on ice for 30 minutes. Ligation was performed next by a heat pulse at 42 °C for 30 seconds and subsequent cooling on ice for 2 minutes. After adding 900 μl of preheated NZY+ medium, the mixture was incubated in a

Thermomixer, running 300 rounds per minute at 37 °C for 60 minutes. After centrifugation at room temperature with 110x g for 10 minutes, the supernatant was removed except of 100 µl. The cell pellet was resuspended and subsequently streaked onto LB-ampicillin agar plates (100 µg/ml ampicillin). Agar plates were incubated at 37 °C overnight to grow single clones to pick for Minipreparation.

4.1.7.3 Mini- and Maxipreparation of plasmids

Minipreparation was performed to isolate plasmids from transformed cells using the MiniPrep Kit from 5-prime. Plasmids were subsequently checked for correct insertion via Sanger sequencing. In case of a proper insertion, Maxipreparation was performed using the Plasmid Maxi Kit from Qiagen to increase plasmid DNA yield for following transfections in cell culture experiments.

For Minipreparation, grown bacterial colonies were picked from their agar plates the day after transformation. After applying them to 4 ml LB-ampicillin medium (100 µg/ml), they were incubated in CERTOMAT R, running 130 rounds per minute at 37 °C overnight. Next, 2 ml bacterial culture was pipetted into an Eppendorf tube, followed by centrifugation at 20 000x g for 1 minute. The supernatant was decanted, and the remaining cell pellet was resuspended in 400 µl ice-cold Complete Lysis Solution. After constant vortexing for 30 seconds, the lysate was incubated at room temperature for 3 minutes. For precipitation, 125 µl 100% isopropyl alcohol was added. The mixture was transferred to a Spin Column Assembly, followed by centrifugation at 20 000x g for 1 minute. The Spin Column Assembly was washed with 400 µl Diluted Wash Buffer and centrifugation at 20 000x g for 1 minute. The flow-through was discarded. The Spin Column Assembly was dry-spun at 20 000x g for 1 minute and placed onto a new collection tube. DNA was eluted by adding 50 µl Elution Buffer directly to the centre of the Spin Column membrane followed by centrifugation at 20 000x g for 1 minute. Eluted plasmid DNA was quantified using NanoDrop 2000c UV-Vis spectrophotometer and subsequently used for downstream applications or stored at -20 °C for further analyses.

For Maxipreparation, a single bacterial colony was picked and added to 5 ml LB-ampicillin medium (100 µg/ml). After incubation in CERTOMAT R, running at 190 rounds per minute at 37 °C for 8 hours, the 1 ml of cell suspension was mixed with 1000 ml of LB-ampicillin medium (100 µg/ml). A bacterial culture was grown overnight in CERTOMAT R, running at 190 rounds per minute at 37 °C. The next day, the cells were centrifuged at 4 000x g for 20 minutes and the supernatant was removed. The pellet was resuspended with 5 ml ice-cold Buffer P1, followed by addition of 5 ml Buffer P2 and incubation at room temperature for 5 minutes. Then, 5 ml ice-cold Buffer P3 was added and the mixture was incubated on ice for 20 minutes. Next, while the solution was centrifuged at 4 °C and 4 000x g for 45 minutes, a Qiagen-tip column was equilibrated by applying 10 ml Buffer QBT. After centrifugation, the supernatant containing plasmid DNA was transferred to the equilibrated column. After the supernatant run through the column via gravity flow, the column was washed twice using 30 ml Buffer QC. The columns were placed onto new falcon tubes. DNA was eluted using 15 ml Buffer QF. The columns were removed, and DNA precipitation was achieved by adding 10.5 ml isopropyl alcohol. After centrifugation at 4 °C and 4 000x g for 30 minutes, the supernatant was discarded and the DNA pellet

was washed with 70% ethanol, followed by another step of centrifugation at 4 °C and 4 000x g for 15 minutes. The supernatant was removed, and the pellet was air-dried until transparent. The pellet was solubilised using 1X TE buffer and quantified using NanoDrop 2000c UV-Vis spectrophotometer. Isolated plasmid DNA was stored at -20 °C for further analyses and cell culture experiments.

4.1.8 Surveyor Nuclease Assay

The Surveyor Nuclease Assay was used to assess correct function of produced guide RNAs. CRISPR/Cas9 systems induced double-strand breaks were mainly repaired via NHEJ pathway, leading to introduction of insertion/deletion mutations (indels). As rates of successful transfection and cutting of CRISPR/Cas9 system is not 100%, there is always a heterogenous mixture of DNA, containing mutated alleles and wild type alleles. Amplification of affected genomic sites, followed by controlled melting and annealing, resulted in formation of homoduplex and heteroduplex DNA, containing mismatch pairings. Treatment with mismatch-specific DNA endonuclease resulted in cleavage of these fragments. Analysis of cleaved fragments by gel electrophoresis was performed to confirm the expected sizes of cleaved fragments and thus functionality of gRNA.

Cultivation of HEK-293 cells followed by transfection with CRISPR/Cas9 system was performed as described. DNA of transfected cells was isolated accordingly. Amplification of DNA was achieved by PCR according to Tables 17 and 18.

Table 17 | PCR reaction mix

Reagent	Final concentration
AccuPrime Taq DNA Polymerase, High Fidelity (5 U/μl)	0.02 U/μl
10X AccuPrime PCR Buffer II	1X
100% DMSO	2%
Forward Primer	0.2 μM
Reverse Primer	0.2 μM
Template DNA	4 ng/μl

Table 18 | PCR thermal cycling conditions

Step	Temperature [°C]	Time [min:sec]	Cycles
Initial Denaturation	94	02:00	1
Denaturation	94	00:30	40
Annealing	62	00:30	
Extension	68	01:30	
Final Extension	72	07:00	1
Hold	4	∞	1

The design of primers for amplification was done with Primer3 software. To ensure cleavage products > 70 bp, primers were placed > 50 bp away from cutting sites of CRISPR/Cas9. The synthesis of desired primers was performed by Eurofins Genomics in high purity salt free quality. Sequences and annealing temperatures of all primers are given in Supplementary Table 5.

A part of the PCR product was subsequently analysed using FlashGel Systems as described, in order to ensure amplification of a DNA fragment appearing as a single sharp band of expected size without background. Next, the remaining DNA was purified as described to remove interfering contaminants from PCR. Formation of heteroduplex/homoduplex DNA was performed by hybridization of purified DNA, using Veriti 96-Well Thermal Cycler according to Table 19.

Table 19 | Hybridization protocol for thermal cycler

Temperature [°C]	Time [min:sec]	Temperature ramp [°C/s]
95	10:00	
95-85		-2.0
85	01:00	
85-75		-0.3
75	01:00	
75-65		-0.3
65	01:00	
65-55		-0.3
55	01:00	
55-45		-0.3
45	01:00	
45-35		-0.3
35	01:00	
35-25		-0.3
25	01:00	
4	∞	

The resulting mixture, containing heteroduplex/homoduplex DNA, was subsequently treated with Surveyor Nuclease S according to Table 20.

Table 20 | Reaction mix for digestion

Reagent	Final concentration
Surveyor Nuclease S	1 µl/reaction
Surveyor Enhancer S	1 µl/reaction
0.15 M MgCl ₂	15 mM
Hybridized DNA	< 20 ng/µl

The digestion was performed by incubation at 42 °C for 60 minutes. The reaction was stopped by adding 1/10th volume of Stop Solution. The digested fragments were subsequently analysed using FlashGel Systems. The expected sizes of undigested PCR products and digested fragments are given in Supplementary Table 6.

4.2 Molecular Biological Methods of RNA

4.2.1 Isolation of RNA

Total RNA of cultivated cells was isolated via phenol-chloroform extraction, using TRIzol Reagent. The growth media was removed, and cells were lysed by adding 750 µl TRIzol Reagent per 10×10^6 cells directly onto the cell culture dish. After homogenization, the lysate was transferred into an Eppendorf tube. The lysed cells were incubated at room temperature for 5 minutes to allow complete dissociation of nucleoprotein complexes. The lysate was mixed with 150 µl 100% chloroform, followed by incubation at room temperature for 3 minutes. After centrifugation at 12 000x g and 4 °C for 15 minutes, the mixture separated into 3 layers: an upper aqueous phase containing RNA, a middle interphase and lower red phenol-chloroform organic phase containing DNA, proteins and other contaminants. The aqueous phase was transferred into a new tube and 375 µl isopropyl alcohol was added for precipitation of RNA. After incubation at room temperature for 10 minutes, the mixture was centrifuged at 12 000x g and 4 °C for 10 minutes. The supernatant was discarded. The remaining RNA pellet was washed twice by adding 1 ml 75% ethanol and centrifugation at 12 000x g and 4 °C for 10 minutes. The clean RNA pellets were air dried and solubilized at 55 °C for 10 minutes using DEPC water. The isolated RNA was quantified using NanoDrop 2000c UV-Vis spectrophotometer and stored at -80 °C for downstream applications.

4.2.2 Reverse transcription of RNA into complementary DNA (cDNA)

Generation of cDNA was achieved by RT (reverse transcription) of total RNA, using a RNase H-Moloney murine leukemia virus reverse transcriptase (M-MLV RT). Therefore, 2 µg total RNA – dissolved in 10 µl DEPC treated H₂O – was mixed with 1 µl random hexamer primers p(dN)₆ (500 µg/ml). After incubation at 68 °C for 10 minutes, in order to unfold secondary structures, the mixture was cooled down to room temperature to allow for the annealing of primers. Subsequently, 9 µl of the mixture given in Table 21 was added.

Table 21 | RT reaction mix

Reagent	Final concentration
SuperScript II Reverse Transcriptase (200 U/ μ l)	10 U/ μ l
5X First-Strand Buffer	1X
0.1 M DTT	10 mM
dNTPs (10 mM each)	500 nM
Recombinant RNAsin Ribonuclease Inhibitor (40 U/ μ l)	2 U/ μ l

Reverse transcription was performed by incubation at 42 °C for 60 minutes. The reaction was inactivated by heating at 70 °C for 15 minutes. The cDNA was diluted at a ratio of 1:12 using 1X TE and stored at -20 °C until qRT-PCR.

4.3 Cell culture

4.3.1 Cultivation of cells

4.3.1.1 Cultivation of human embryonic kidney cells

Human embryonic kidney cells 293 (HEK-293 cells, ACC 305) from Leibniz-Institut DSMZ were cultivated as a stable cell line to perform pretesting of transfection with used gRNAs. HEK-293 cells were cultivated using DMEM GlutaMAX-I, supplemented with 10% FBS and 1% PenStep (10 000 U/ml Penicillin, 10 000 μ g/ml Streptomycin). Cells were allowed to grow up to 80-90% confluence before passaging. To this end, the medium was removed and after washing with PBS Dulbecco, the cells were detached using Trypsin/EDTA Solution (0.05%/0.02%). After incubation at 37 °C for 3-5 minutes, the detached cells were resuspended with fresh DMEM GlutaMAX-I medium and seeded in ratios of 1:10 to 1:100. Medium was renewed every 48-72 hours.

4.3.1.2 Cultivation of human induced pluripotent stem cells

The hiPSC line (HMGU1c) from Helmholtz Zentrum München was cultivated as experimental cell line for further differentiation into desired cell types associated with atherosclerosis, such as endothelial cells and smooth muscle cells. HiPSCs were cultured using mTeSR 1 medium completed with mTeSR 1 5X supplement and 1% PenStep (10 000 U/ml Penicillin, 10 000 μ g/ml Streptomycin). The complete mTeSR 1 medium contained rh bFGF and rh TGF β for maintenance of pluripotency. For optimal growth, medium was replaced every 24 hours.

The maintenance of pluripotency was, inter alia, achieved by using appropriate matrix conditions. Therefore, Matrigel hESC-qualified Matrix was used as a soluble basement membrane extract of Engelbreth-Holm-Swarm (EHS) tumor. With its components (e.g. laminin, collagen IV, growth factors) the function of a genuine basement membrane was imitated. Thus, Matrigel hESC-qualified Matrix was

diluted with DMEM GlutaMAX-I according to the lot-specific dilution factor. After coating with diluted Matrigel, the plates were incubated at room temperature for 60 minutes. Coated plates were subsequently used or stored at 4 °C for future application.

HiPSC cultures were allowed to grow up to 80-90% confluence until passaging. Then, mTeSR 1 medium was removed and, after washing with PBS Dulbecco, the cells were detached using Accutase. After incubation at 37 °C for 5-10 minutes, the detached cells were resuspended in fresh medium containing 10 µM ROCK inhibitor (Y-27632 dihydrochloride, 10 mM). As inhibitor of Rho-associated, coiled-coil containing protein kinase (ROCK), Y-27632 was crucial for maintenance of pluripotency and enhanced survival of hiPSCs. After removal of Matrigel coating medium, resuspended cells were subsequently seeded to a coated plate containing fresh mTeSR 1 medium with 10 µM ROCK inhibitor. Cells were passaged in split ratios of 1:6 to 1:20, depending on past colony density and individual needs.

4.3.1.3 Cultivation of vascular endothelial cells

Vascular endothelial cells (ECs) derived from hiPSCs were generated following a published protocol and as described below in detail in chapter 4.3.5.⁷⁸ Shortly after establishment, the cells were cultivated using Endothelial Cell Growth Medium supplemented with 6% FBS and 1% PenStep (10 000 U/ml Penicillin, 10 000 µg/ml Streptomycin). The complete Medium contained all growth factors and supplements necessary for optimal growth (e.g. rh bFGF, rh EGF). Coating of used plates was performed at room temperature for 60 minutes using 2 µg/cm² fibronectin from human plasma solubilized with PBS Dulbecco. After removal of fibronectin solution, plates were washed with H₂O and the Endothelial Cell Growth Medium was added. For passaging, the medium was removed, and the cells were washed with PBS Dulbecco. Detachment of cells was performed using Trypsin/EDTA Solution (0.05%/0.02%) at 37 °C for 3-5 minutes. The detached cells were resuspended with fresh Endothelial Cell Growth Medium and were seeded onto fibronectin-coated plates at densities of 26 000 cells/cm². Cells were allowed to grow up to 80-90% confluence before the next passage.

4.3.1.4 Cultivation of vascular smooth muscle cells

Vascular smooth muscle cells (VSMCs) were derived from hiPSC following a published protocol and were cultivated using SmGM-2, consisting of Smooth Muscle Cell Basal Medium (SmBM), supplemented with rh FGF-B, rh EGF, rh insulin (Lonza), FBS and gentamycin sulfate/amphotericin-B (GA-1000).⁷⁸ The coating of used plates was performed at room temperature for 15 minutes using 0.1% gelatine solubilized in H₂O. After removal of the gelatine solution, SmGM-2 was added, and the cells were seeded. For passaging, the medium was removed, and the cells were washed with PBS Dulbecco. Detachment of cells was performed using Trypsin/EDTA Solution (0.05%/0.02%) at 37 °C for 3-5 minutes. The detached cells were resuspended with fresh SmGM-2 and were seeded onto gelatine-coated plates at densities of 30 000 cells/cm². Cells were allowed to grow up to 80-90% confluence before the next passage. The medium was replaced every 48-72 hours.

4.3.2 CRISPR/Cas9 methodology and transfection of cells

4.3.2.1 Test of CRISPR/Cas9 efficiency by transfection of human embryonic kidney cells

Lipofection of HEK-293 cells was performed to examine the functionality of the designed CRISPR/Cas9 system. Therefore, HEK-293 were plated onto collagen-coated dishes (0.5 mg/ml collagen) at densities of 50 000 cells/cm² followed by incubation at 37 °C for 6-8 hours. To check for functionality of gRNAs itself, pCas9_GFP vector (coexpression of human codon-optimized *Streptococcus pyogenes* type II CRISPR Cas9 gene together with GFP from one CAG expression plasmid vector) was transfected with one pGuide vector, containing the sequence for the gRNA driven by a U6 polymerase III promoter.^{79,80} To check the functionality of the knockout, pCas9_GFP was transfected with both pGuide vectors. The transfected amount of DNA did not exceed 1 µg per transfection. After incubation, for every transfection, a solution was prepared according to Table 22.

Table 22 | Reaction mix for lipofection

Reagent	Quantity for	
	single	double
	gRNA transfection	
Opti-MEM I Reduced Serum Medium	100 µl	100 µl
TransIT-2020 Reagent	3 µl	3 µl
pGuide vector 1	500 ng	333 ng
pGuide vector 2	-	333 ng
pCas9_GFP vector	500 ng	333 ng

Prepared solutions were incubated at room temperature for 20 minutes and subsequently added drop-wise to different areas of one well. After gentle rocking of the culture vessel back-and-forth to distribute the solution, the cells were incubated at 37 °C for 48-72 hours. Successful lipofection was confirmed by visualizing coexpressed GFP using fluorescence microscopy.

The transfected cells were then harvested, and DNA was isolated as described to check for functionality of used CRISPR/Cas9 systems.

4.3.2.2 Electroporation of CRISPR/Cas9 and gRNA expression plasmids into human induced pluripotent stem cells

Electroporation of hiPSCs was performed using the 4D-Nucleofector System to knockout desired genomic regions. To this end, hiPSCs were detached and filtered through 30 µm Pre-Separation Filters to ensure single cell suspension. For each electroporation, 800 000 cells were added in low adhesion tubes, followed by centrifugation at 200x g and room temperature for 10 minutes. The supernatant was removed, and the remaining cell pellet was resuspended in a previously prepared solution for electroporation, according to Table 23.

Table 23 | Reaction mix for electroporation

Reagent	Quantity
Nucleofector Solution (Amaxa P3 Primary Cell Kit)	82 μ l
Supplement (Amaxa P3 Primary Cell Kit)	18 μ l
pGuide vector 1	2 μ g
pGuide vector 2	2 μ g
pCas9_GFP vector	4 μ g

Resuspended cells were transferred into 100 μ l Single Nucleocuvettes, followed by subsequent electroporation in the 4D-Nucleofector System, running with CD-118 pulse program. Immediately after electroporation, 500 μ l of fresh pre-warmed mTeSR 1 medium, containing 10 μ M ROCK inhibitor was added to the cells. Nucleocuvettes were then incubated at 37 °C for 5 minutes to promote recovery. The electroporated cell suspension was subsequently transferred onto matrigel-coated plates, covered with fresh mTeSR 1 containing 10 μ M ROCK inhibitor. The cells were distributed evenly by gentle rocking of the cell culture vessel. The plates were put into the IncuCyte ZOOM for continued videomicroscopy while the cells were kept in a humidified 37 °C/5% CO₂ incubator to monitor GFP fluorescence of successful transfected cells. Thereby, IncuCyte ZOOM was configured for green fluorescence and phase contrast. The medium was renewed every 24 hours.

4.3.3 Fluorescence-activated cell sorting of human induced pluripotent stem cells

FACS of hiPSCs was performed using MoFlo Astrios EQ to obtain successfully electroporated GFP-positive Cas9-carrying cells to increase the yield of knockout clones. Therefore, hiPSCs were detached and centrifuged at 100x g for 5 minutes. The supernatant was removed, and the remaining cell pellet was resuspended in FACS Buffer, consisting of PBS Dulbecco and 10 μ M ROCK inhibitor. The resuspended cells were filtered through 30 μ m Pre-Separation Filters into a FACS tube to assure single cell suspension. The concentrations of cell suspensions were adjusted to 5×10^6 cells/ml for FACS. Unelectroporated cells were used to set gating of FSC and SSC. Debris and dead cells were indicated by low FSC versus high and low SSC. Doublets and clumps were indicated by high FSC versus high FSC area. To exclude these adverse events and to maintain single cells, gates were set accordingly. Next, the green fluorescence gate was set using the unelectroporated control. The final sorting of electroporated cells was subsequently performed. Single cells expressing GFP were collected in a new Eppendorf tube containing a post-FACS recovery medium, consisting of 1 volume conditioned mTeSR 1, 1 volume fresh mTeSR 1, 10 μ M ROCK inhibitor and 12 ng/ml rh bFGF. The post-FACS cells were seeded at a density of 15000 cells per matrigel-coated 10-cm dish containing post-FACS recovery medium. This was a limited density, sufficient to ensure growth of colonies from single GFP-positive cells and yet avoiding colonies growing into each other. Cells were incubated in a humidified

37 °C/5% CO₂ incubator for 7-10 days until colonies had grown large enough for picking. The medium was renewed every 24 hours.

4.3.4 Manual picking and passaging of human induced pluripotent stem cells

Post-FACS, several of the single cell colonies from the 10-cm dishes were picked manually for clonal expansion and screening for knockout clones. Once the colonies were big enough, the medium was removed and, after washing with PBS Dulbecco, the cells were partly detached by adding Gentle Cell Dissociation Reagent, followed by incubation at room temperature for 6-8 minutes or until the edges of the clones began to rise. Then, Gentle Cell Dissociation Reagent was carefully removed and mTeSR 1 medium with 10 µM ROCK inhibitor was slowly pipetted onto the plate to avoid tearing off the colonies. As a picking setup, the plate was placed under a microscope with a 10X objective. The colonies were picked using a 200 µl pipette, carefully scraping off small clumps of cells from a single clone and collecting them with the pipette. The cells were subsequently pipetted into a previously prepared matrigel-coated 96-well plate, containing fresh mTeSR 1 with 10 µM ROCK inhibitor. For each colony a new pipette was used. After picking, the plates were put in a humidified 37 °C/5% CO₂ incubator for ~ 4-7 days, requiring medium changes every 24 hours.

Once colonies were grown to ~80% confluence, they were passaged and split into two 96-well plates, as described. One plate of cells was used for DNA isolation and subsequent analysis of knockout clones as described. The other plate was used for further cultivation and clone expansion. When PCR-screening of clones revealed a knockout genotype, respective cells were passaged onto 6-well plates and 10-cm dishes. Knockout clones were frozen or subsequently used for differentiation into desired cell lines.

4.3.5 Generation of vascular endothelial and smooth muscle cells from human induced pluripotent stem cells

HiPSCs were differentiated into ECs and VSMCs to analyse gene expression over time and to perform functional *in vivo* experiments following an established protocol.⁷⁸

As vascular progenitor cells originate from lateral and posterior mesoderm, the first step of differentiation was to induce mesoderm by canonical Wnt signalling activation through inhibition of GSK3β and treatment with BMP4. For differentiation into ECs, mesodermal progenitors were exposed to VEGF-A, whereas differentiation into VSMC was achieved by treatment with PDGF-BB and Activin A.

The differentiation procedure was started by seeding hiPSCs at a density of 32 000 cells/cm² on day 0. Lateral mesoderm induction was performed on day 1 by replacing mTeSR 1 medium with priming medium, consisting of N2B27 medium supplemented with 6 µM CHIR99021 as inhibitor of GSK3 and 25 ng/ml rh BMP4 as additional inducer of mesoderm. N2B27 medium itself was set up by mixing equal parts of DMEM/F-12 with Neurobasal Medium and supplementing it with 50 µM β-mercaptoethanol, B-27 Supplement without Vitamin A and N-2 Supplement (each in final concentration of 1X). Priming

medium was remaining on the plates for 3 days without medium change. After generation of hiPSC-derived mesoderm by day 4, the next steps of differentiation varied depending on desired cell type.

Vascular endothelial cell induction was performed on day 4 by replacing priming medium with StemPro-34 SFM medium supplemented with 200 ng/ml rh VEGF and 2 μ M Forskolin, as vascular maturation is promoted by enhanced protein kinase A activity via increased cyclic AMP. StemPro-34 SFM medium itself was set up by mixing StemPro-34 medium with StemPro-34 Nutrient Supplement and adding 1% PenStep (10 000 U/ml Penicillin, 10 000 μ g/ml Streptomycin) and GlutaMAX Supplement. The medium was renewed on day 5. Endothelial cell differentiation was completed by day 6. As vascular endothelial cadherin (VE-cadherin, CD144) served as marker for mature endothelial cells, CD144 Microbead purification was performed on day 6 to remove cells other than ECs. To this end, the cells were dissociated and centrifuged at 300x g for 10 minutes. The supernatants were removed, and the cell pellets were resuspended in a mixture of 80 μ l buffer and 20 μ l CD144 MicroBeads per 10^7 cells for magnetic labeling. Buffer was prepared by mixing MACS BSA Stock Solution 1:20 with autoMACS Rinsing Solution. The resuspended cells were incubated at 4 °C for 20 minutes, followed by washing with 2 ml buffer per 10^7 cells and centrifugation at 300x g for 10 minutes. After removing the supernatant, the pellet was resuspended in 500 μ l buffer per 10^8 cells. MACS columns were placed in the magnetic field of a MACS Separator, and rinsing was performed with 1 ml buffer before actual sorting. The cell suspension was subsequently applied onto the column, retaining the magnetically labelled CD144⁺ cells. After washing three times with 500 μ l buffer, the flow-through was discarded and the columns were placed onto new collection tubes. After removing the columns from the magnetic field, the magnetically retained CD144⁺ cells were eluted using 1.5 ml buffer. The cell suspension was subsequently centrifuged at 300x g for 10 minutes. The supernatant was discarded. The cells were resuspended in StemPro-34 SFM medium supplemented with 50 ng/ml rh VEGF, followed by seeding on fibronectin-coated dishes (containing identical medium) at a density of 26 000 cells/cm². Fibronectin-coating was achieved by adding 2 μ g/cm² fibronectin onto the plates followed by incubating for 60 minutes. Before the medium was applied, the fibronectin solution was removed, and plates were washed with H₂O. On day 8, the medium was replaced by Endothelial Cell Growth Medium supplemented with 6% FBS and 1% PenStep (10 000 U/ml Penicillin, 10 000 μ g/ml Streptomycin) and the cells were subsequently treated as endothelial cells.

VSMC induction was performed on day 4 by replacing priming medium with N2B27 medium supplemented with 10 ng/ml rh PDGF-BB and 2ng/ml Activin A. The medium was renewed on day 5. VSMC differentiation was completed by day 6. At that time, cells were dissociated and seeded at a density of 30 000 cells/cm² onto gelatine-coated plates, containing N2B27 medium supplemented with 10 ng/ml rh PDGF-BB to promote synthetic state of VSMC. Gelatine-coating was achieved by adding 0.1% gelatine solubilized in H₂O onto plates followed by incubating for 15 minutes. Before the medium was applied, the gelatine solution was removed, and the plates were washed with H₂O. On day 8, the medium was replaced by Smooth Muscle Cell Growth Medium-2 (SmGM-2) and the cells were subsequently treated as smooth muscle cells.

4.3.6 Functional experiments in cell culture

Functional experiments of ECs and VSMCs derived from hiPSCs were performed to compare the behaviour of knockout cells against controls. For each experiment, cells were seeded onto 0.1% gelatine-coated 96-well plates (SMCs) or 2 $\mu\text{g}/\text{cm}^2$ fibronectin coated 96-well plates (ECs), respectively. ECs and VSMCs were analysed separately. During experiments, cells were cultivated as described above.

4.3.6.1 Apoptosis

Apoptosis was measured using a commercial luminescence assay detecting activated caspase 3 and 7. Therefore, cells were seeded at a density of 5 000 cells/well each in quadruplicates for samples and controls. Apoptosis was measured on a plate reader after 24 hours, or separately 24 hours after apoptosis was experimentally induced using 0.25 μM staurosporine, while controls remained untreated. For the measurement, 1:1 volume of Caspase-Glo Reagent was added to all cell cultures. Caspase-Glo Reagent provided a caspase-3/-7 substrate, Z-DEVD-aminoluciferin. After cleavage and release of Z-DEVD, resulting aminoluciferin served as a substrate for luciferase reaction, resulting in a luminescence signal, detected by the SpectraMax Paradigm 60 minutes after adding Caspase-Glo Reagent. The luminescence signal was proportional to the amount of caspase activity.

4.3.6.2 Proliferation

Proliferation was measured by imaging-based videomicroscopy using the IncuCyte ZOOM label-free proliferation assay. Therefore, cells were seeded in quadruplicates for samples and controls each at two different densities (of 2 500 or 5 000 cells/well). The plates were put into the IncuCyte ZOOM in a humidified 37 °C/5% CO₂ incubator to track cell numbers via calculation of how much of well area was covered by cell bodies (“confluence”). This was done by photographing every well once every 4 hours for the next 5 days. The medium was renewed every 48 hours without interruption of the assay. Collected phase contrast live cell images (10X objective) were analysed using the integrated object algorithm which isolates cells from the background. Thus, kinetic graphs of the cell-occupied area (% confluence) were created using the median of quadruplicates.

4.3.6.3 Adhesion

Adhesion was measured by imaging-based videomicroscopy using the IncuCyte ZOOM label-free adhesion assay. Therefore, ECs were seeded in quadruplicates at a density of 20 000 or 30 000 cells/well, respectively. The plates were put into the IncuCyte ZOOM in a humidified 37 °C/5% CO₂ incubator overnight.

After 24 hours, the PBMCs required for the assay were isolated from human whole blood from a healthy donor. After whole blood was collected in Vacutainer CPT Mononuclear Cell Preparation Tubes, centrifugation was performed at 1 650x g for 20 minutes. The supernatant, containing the

PBMCs, was transferred into a new tube and PBS Dulbecco was added up to 50 ml. After centrifugation at 400x g for 15 minutes, the supernatant was discarded, and the remaining cell pellet was resuspended in PBS Dulbecco. After filtration through 40 µm Falcon Cell Strainer, cells were centrifuged at 300x g for 15 minutes and the remaining cell pellet containing PBMCs was resuspended in RPMI medium supplemented with 10% FBS, and 1% PenStep (10 000 U/ml Penicillin, 10 000 µg/ml Streptomycin).

PBMCs for adhesion assays were obtained from a healthy blood donor in the context of the Munich Study of Preanalytical and Analytical Laboratory Test Establishment (MyTEST) after written informed consent. The MyTEST study has been approved by the Ethics Committee of the Medical Faculty of the University Munich Germany (Reference number 17-012).

Endothelial cell plates were washed using Endothelial Cell Growth Medium supplemented with 6% FBS and 1% PenStep (10 000 U/ml Penicillin, 10 000 µg/ml Streptomycin). Isolated unlabelled PBMCs were subsequently pipetted onto EC cultures at a density of 30 000 cells/well. After an incubation period of 45 minutes to allow for adherence of PBMCs to endothelial cells, the remaining non-adherent cells were washed off the plates. The washed plates were subsequently put into the IncuCyte ZOOM in a humidified 37 °C/5% CO₂ to capture phase contrast images.

The collected phase contrast live cell images (10X objective and 20X objective) were analysed using the integrated object algorithm which achieves differentiation between endothelial cells and PBMCs due to z-position and cell shape/granularity. The confluence mask (%) was used to calculate the area of adherent PBMCs relative to the total area of the well. Graphs of occupied area were generated using the mean of quadruplicates.

4.3.6.4 Migration

Migration was measured by imaging-based videomicroscopy using the IncuCyte ZOOM label-free Scratch Wound assay. Therefore, cells were seeded in quadruplicates for samples and controls each at densities of 15 000 or 30 000 cells/well, respectively. After incubation for 48 hours, the plates were processed with the 96-pin WoundMaker which created homogenous 700-800 µm wide cell-free wounds in the cell monolayer according to manufacturer's instructions. After washing twice with PBS Dulbecco, the medium was added, and plates were put into the IncuCyte ZOOM in a humidified 37 °C/5% CO₂ incubator to track migration of wells by photographing every well once every 2 hours for the next 3 days. The medium was renewed every 48 hours. The collected phase contrast live cell images (10X objective) were analysed using the integrated IncuCyte Cell Migration Software. Therefore, a group of images was selected to define, train and test parameters used for analysis of all images within the experiment. Parameters were refined to achieve clear differentiation between the background and the cells within the images. Adjusted parameters were subsequently used to analyse all recorded images of the experiment. Analysis was performed by applying 3 masks followed by calculation of 3 metrics. The initial scratch wound mask was generated for the first image of each well to define the initial borders of the wound region. Further scratch wound masks identifying the leading edge of migrating cells were computed for each subsequent image in the series. The confluence mask was used in combination with

the initial scratch wound mask to determine the cell confluence within the wound region for calculation of confluence metrics. Three integrated metrics were subsequently calculated based on processed images. The wound width metric (μm) represented the average distance between migration edges of the cell lawn at the wound. The wound confluence metric (%) represented the confluence of cells within the wound region given as percentage of wound area occupied by cells. The relative wound density metric (%) represented the cell density in the wound area relative to cell density outside the wound. The relative wound density was accounting for background density of the initial wounded area. As each reported value was background subtracted, the metric is self-normalizing and designed to be 0% at the beginning of the experiment and 100% when cell density inside the wound is equal to cell density outside the wound. Kinetic graphs of the cell migration were created using the mean of quadruplicates.

4.4 Bioinformatics analysis

4.4.1 Analysis of haplotype blocks

4.4.1.1 Definition of haplotype block margins

Haplotype blocks harbouring SNPs associated with atherosclerosis were defined manually. Therefore, genotype data of a Caucasian population (Utah Residents with Northern and Western European Ancestry, CEU, $n_{\text{total}} = 183$, $n_{\text{male}} = 89$, $n_{\text{female}} = 94$) were extracted from 1000 Genomes Project Phase 3 (<http://www.internationalgenome.org/>). Genotype data were filtered for SNPs with a MAF ≥ 0.05 and ≤ 2 alleles. Filtered genotype data were used to generate plots of linkage disequilibrium (LD) for D' and r^2 using R software. Genomic regions inspected spanned over 1 Mb with a physical distance of 500 kb to each side of the SNP of interest. If necessary, genomic regions were extended. SNPs associated with atherosclerosis were tagged within generated LD plots. The physical extent of each haplotype block was determined by visual inspection of D' and r^2 pattern at the blocks borders. Resulting data were used for further analysis.

4.4.1.2 Determination of annotated genes within haplotype blocks

Determined haplotype blocks were analysed for genes located inside these blocks. Thereby, a gene was acknowledged as within a haplotype block, as soon as parts of the gene reached into the block. Genes were extracted from GENCODE version 19 (<https://www.gencodegenes.org/>). Extracted genes were grouped in following subsets: protein coding genes, long non-coding RNAs (lncRNAs), micro RNAs (miRNAs), ribosomal RNAs (rRNA), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), pseudogenes and other RNAs.

4.4.1.3 Gene expression analysis

Expression profiling of identified genes was performed to gain insight into genetics of haplotype blocks in order to screen for possible candidate loci to be modified. Transcriptome data of tissues (liver and plaque), primary cell lines (PBMC, EC, FB, SMC) and immortalised cell lines (MonoMac, THP-1, HEK-293, U937, Hep G2, CaCo-2) were processed using R software. Cut-offs for gene expression of protein coding genes was set to ≥ 5 reads per kilobase of transcript per million mapped reads (RPKM) and to $\text{RPKM} \geq 2$ for lncRNAs. The absolute and relative expression of genes was visualised using heatmaps generated for every haplotype block via R software.⁸¹

4.5 Statistical analysis

4.5.1 Systematic review of published genome-wide association studies

Published GWAS of atherosclerosis and its risk factors, such as elevated lipid levels, high blood pressure, diabetes mellitus, and smoking, were gathered to create a unified library of SNPs associated with the corresponding phenotype. Therefore, the NHGRI-EBI Catalog of published GWAS (GWAS Catalog, <https://www.ebi.ac.uk/gwas/>) was accessed (05/2016) and phenotypes of interest were defined by arrangement of default traits according to Supplementary Table 8. In addition, relevant GWAS not featured by GWAS Catalog were added manually. Compiled GWAS are given in Supplementary Table 9. Reported SNPs of published GWAS were filtered for $\text{MAF} \geq 0.05$ and $p\text{-value} \leq 5 \times 10^{-8}$ to be statistically significant. Combined lists of filtered SNPs for atherosclerosis and its risk factors served as source for future analyses.

4.5.2 Statistical analyses of functional assays in cell culture

For statistical analysis of data, the GraphPad Prism Software (Version 6.0) was used. Data are given as mean \pm SEM, or median and interquartile range, as indicated. Normality of distribution was assessed using the D'Agostino and Pearson test. Comparison of multiple groups was done using ANOVA and Tukey was performed as a post-test. Comparison of 2 groups of normally distributed samples was done using the t-test.

Data for assays in time-series experiments were tested for normality distribution and analysed by repeated-measures 1-way ANOVA with Dunnett's multiple comparison test. Data are plotted in the figures as mean \pm SEM (Figures 16, 17, 18, 19), as indicated. Significances are highlighted by asterisks (* equals $p \leq 0.05$). Statistical analyses were plotted in R or GraphPad.

5

Results

5.1 Survey of genetic loci linked to human atherosclerosis by genome-wide association studies in populations

5.1.1 Assembly of genetic variations associated with atherosclerosis risk from existing GWAS

Exploring the heritable genetic component of diseases by GWAS is a hypothesis-free approach that potentially allows the discovery of many different aspects, as well as as yet unknown routes to disease onset. We set out to bioinformatically compile a list of all known atherosclerosis and atherosclerosis risk factor-associated SNPs, which would then allow us to focus on a functionally unexplored locus with large effect size and to start to decipher it mechanistically in this thesis.

The first step was to access the publicly open GWAS Catalog (<https://www.ebi.ac.uk/gwas/>) (query state 05/2016) in order to create a list of all SNPs so far known to be associated with atherosclerosis, and, more generally, also with its risk factors, such as increased blood pressure, diabetes mellitus, smoking and known factors of dyslipidaemia, e.g. high levels of LDL or low levels of HDL (see Supplementary Table 8 for complete list of traits used in the query). Specifically, the GWAS Catalog offers default traits to set a query, and we selected dozens of relevant default traits to best possibly catch all GWAS reporting on phenotypes related to aspects of atherosclerosis and its risk factors as given in Supplementary Table 8 and Figure 4. The query then resulted in a total of $n = 124$ published GWAS, of which 30 GWAS dealt with atherosclerosis related traits (ARTs), 32 with lipids, 16 with blood pressure, 46 with diabetes mellitus and 2 with smoking. As indicated in Figure 4, two GWAS were dealing with overlapping traits (#49 and #89). Figure 4 gives a visual overview of the 124 GWAS and their corresponding traits as well as phenotypes. The detailed list of all GWAS found is given in Supplementary Table 9.

The second step was to select and gather the content of the discovered GWAS in order to create a list of corresponding SNPs. A total of $n = 968$ SNPs with a $p\text{-value} \leq 5 \times 10^{-8}$ were extracted, given that this is a stringent and commonly used threshold for prioritizing hits with genome-wide significance in common-variant GWAS.²⁴ Thereof, 216 SNPs were associated with atherosclerosis, 337 with lipids, 202 with blood pressure, 211 with diabetes mellitus and 2 with smoking. The assembled lists are given in Supplementary Tables 10-14.

5.1.2 Definition of haplotype block borders

In the next step, we explored where the identified SNPs associated with atherosclerosis (Figure 4) were located on the human genome and in relation to each other. For example, it could be that several SNPs, even from different traits, clustered together closely in genetic loci tend to be inherited together. Therefore, for each SNP we manually determined in which haplotype block from the 1000 genomes Caucasian CEU population it fell. This was a means to determine the number and location of all genetic loci of relevance to atherosclerosis (see methods for details): For every SNP associated with atherosclerosis from Figure 4, a plot of linkage disequilibrium was created in D' and r^2 . In order to determine the borders of LD as the ends of the haplotype blocks, we inspected the patterns D' and r^2 in the plots visually over 500 kb on each side of a SNP of interest (see method for details): In most cases, multiple SNPs associated with atherosclerosis mapped into one haplotype block. The methodical approach of determining the exact borders of a haplotype block is demonstrated in Figure 5, using the locus on chromosome 9p21 as an example. The processing of all 216 SNPs associated with atherosclerosis resulted in a total of $n = 120$ haplotype blocks, containing these SNPs. The complete list of all haplotype blocks with their defining borders, absolute length in bp, physical position on chromosomes, and containing SNPs is given in Supplementary Table 15. Given the nature of recombination and genetic inheritance within these haplotype blocks, every SNP or genomic site (also beyond the lead SNPs from GWAS) could be the causal variant responsible for the observed phenotype.

5.2 Prioritization of haplotype blocks for functional studies

In order to narrow down the list of 120 haplotype blocks associated with atherosclerosis to few of the most interesting candidate loci, we assessed if any colocalization existed between haplotype blocks that we had annotated as atherosclerosis-haplotype blocks (Supplementary Table 15) and those SNPs that we had associated with atherosclerosis risk factors. Further, the genetic content of the haplotype blocks was determined as well as the strength of association with atherosclerosis and effect size of lead SNPs.

5.2.1 Colocalization of haplotype blocks and SNPs associated with risk factors

The colocalization of haplotype blocks and risk SNPs can be assessed in two different ways. First, it was evaluated which risk SNPs mapped inside the atherosclerosis haplotype blocks. Then, the physical position of a risk SNP was compared to the physical positions of the haplotype blocks. The results are displayed in Figure 6.

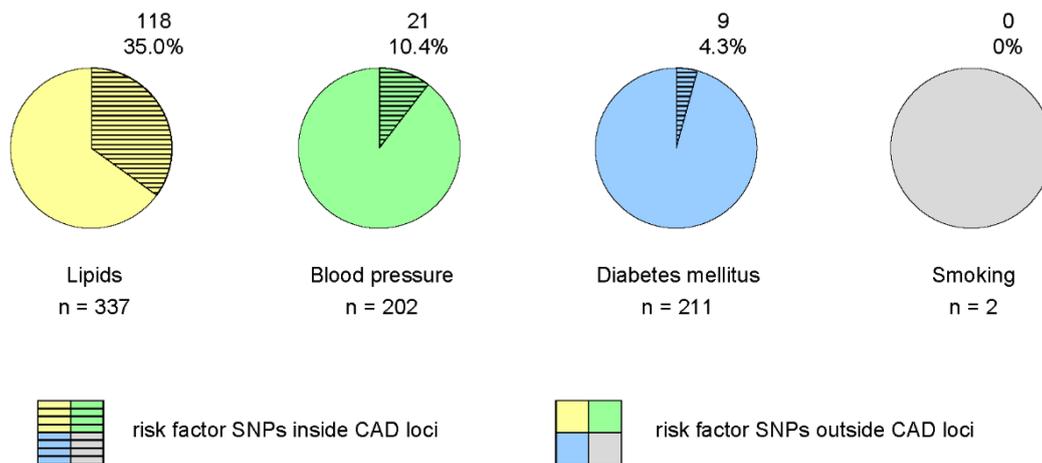


Figure 6 | Degree of colocalization of atherosclerosis haplotype blocks and risk SNPs. Displayed are total numbers (n) of risk SNPs (related to distorted lipid metabolism, blood pressure, diabetes or smoking) and the proportion (%) of these risk-associating SNPs to map loci that have additionally been defined to be associated with atherosclerosis through GWAS (hatched coloured area), as opposed to mapping elsewhere in the genome (non-hatched coloured area). Note that the majority of atherosclerosis risk loci does not contain known genetic variation associated with the classical systemic risk factors.

Here, 35.0% of SNPs associated with lipids map inside the haplotype blocks, while only 10.4% of SNPs associated with blood pressure and 4.3% of SNPs associated with diabetes mellitus map into haplotype blocks, and SNPs associated with smoking are located solely outside atherosclerosis haplotype blocks. With blood pressure and diabetes mellitus being one of the classic risk factors, the proportion of their SNPs mapping into atherosclerosis haplotype blocks remains noticeably low for as yet unknown reasons.

Of the 120 atherosclerosis haplotype blocks, we next investigated a potential overlap with one or more haplotype blocks with or without SNPs associated with atherosclerosis risk factors. The results and the workflow up to that point are displayed in Figure 7.

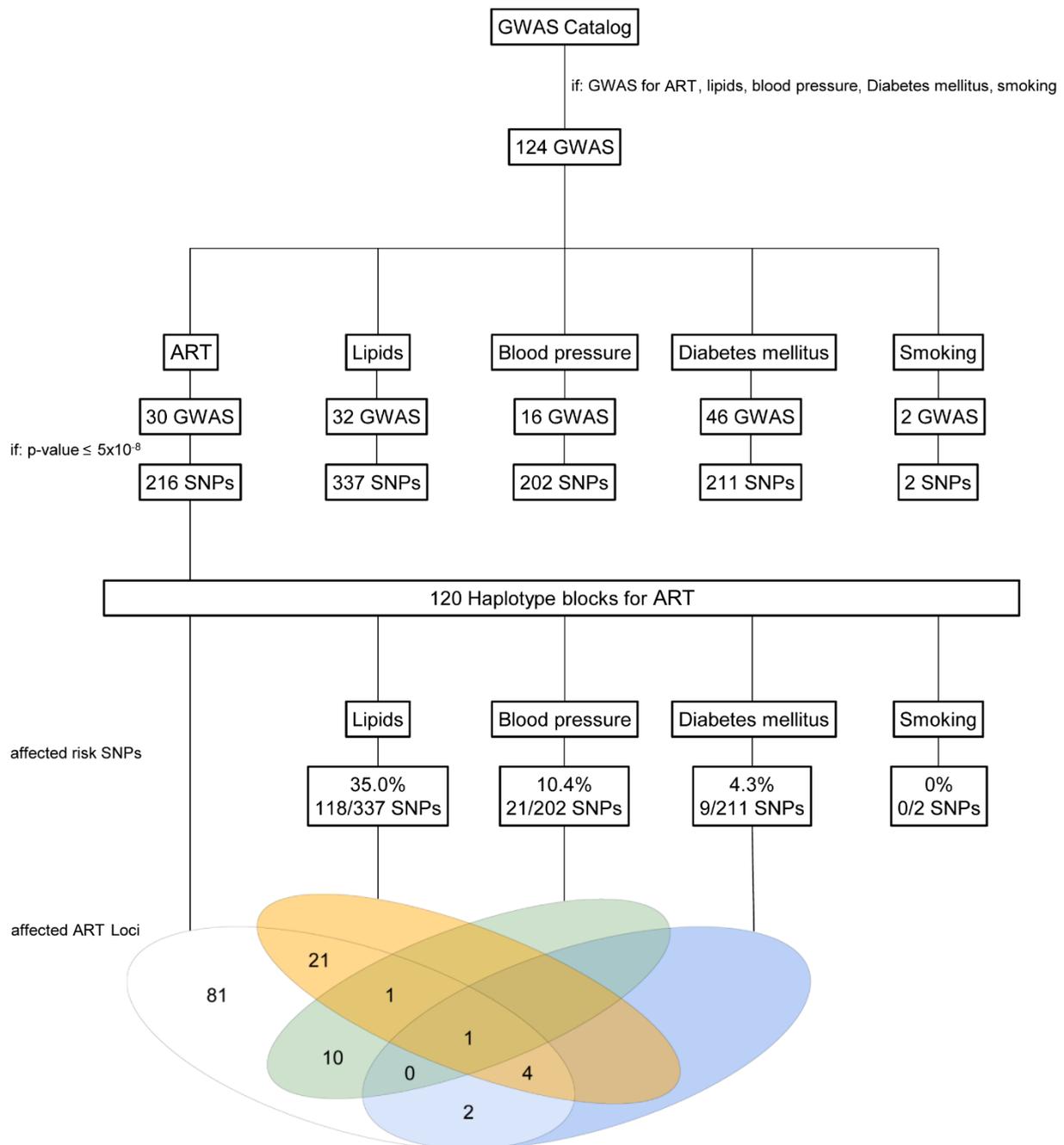


Figure 7 | Progression and methodology of work. GWAS Catalog (<https://www.ebi.ac.uk/gwas/>) was assessed for GWAS related to ART, lipids, blood pressure, diabetes mellitus, and smoking. Resulting GWAS ($n = 124$) were screened for SNPs with p-values $\leq 5 \times 10^{-8}$ ($n = 968$). ART SNPs were used to define haplotype blocks ($n = 120$) based on D' and r^2 data of CEU population ($n = 183$) of 1000 Genomes Project Phase 3 (<https://www.internationalgenome.org/>). Subsequently, mapping of risk SNPs into ART haplotype blocks was determined. Note the presence of ART haplotype blocks devoid of any association with known systemically acting risk factors.

By this second measure, 39 of 120 atherosclerosis haplotype blocks (32.5%) contain risk SNPs. Hereof, ART haplotype blocks contain solely SNPs associated with lipids (21 haplotype blocks) or SNPs associated with blood pressure (10 haplotype blocks). A combination of different risk SNPs (from different traits) mapping into one haplotype block remains an exception. For example, only 4 haplotype blocks contain both SNPs associated with lipids as well as with diabetes mellitus (Figure 7). The majority of published haplotype blocks (81 out of 120, 67.5%) were clear of any risk SNPs. In these

haplotype blocks, the mechanism of atherosclerosis is unlikely mediated by one of the classical risk factors. In fact, it is suggestive of an alternative mechanism, related to the growing number of cellular processes in the cells of the arterial walls that are relevant for atherosclerosis onset (see introduction) or to related other less well understood atherosclerosis risk-promoting settings such as systemic, metabolic or inflammatory processes.

5.2.2 Genetic content of haplotype blocks

All 120 haplotype blocks were then screened for the identity of genes located inside of the block or at least reaching into it. Therefore, the physical positions of each haplotype block (hg19) were compared with the physical positions of annotated genes as listed in the Gencode Database V19 (hg19). The database contains subclasses of protein-coding RNA, lncRNA, miRNA, rRNA, snRNA, snoRNA, pseudogenes and other noncoding RNA. The absolute and relative annotated gene content of every individual haplotype block as well as the overall gene content of all 120 haplotype blocks are displayed in Figure 8. The length distribution of identified haplotype blocks is also shown.

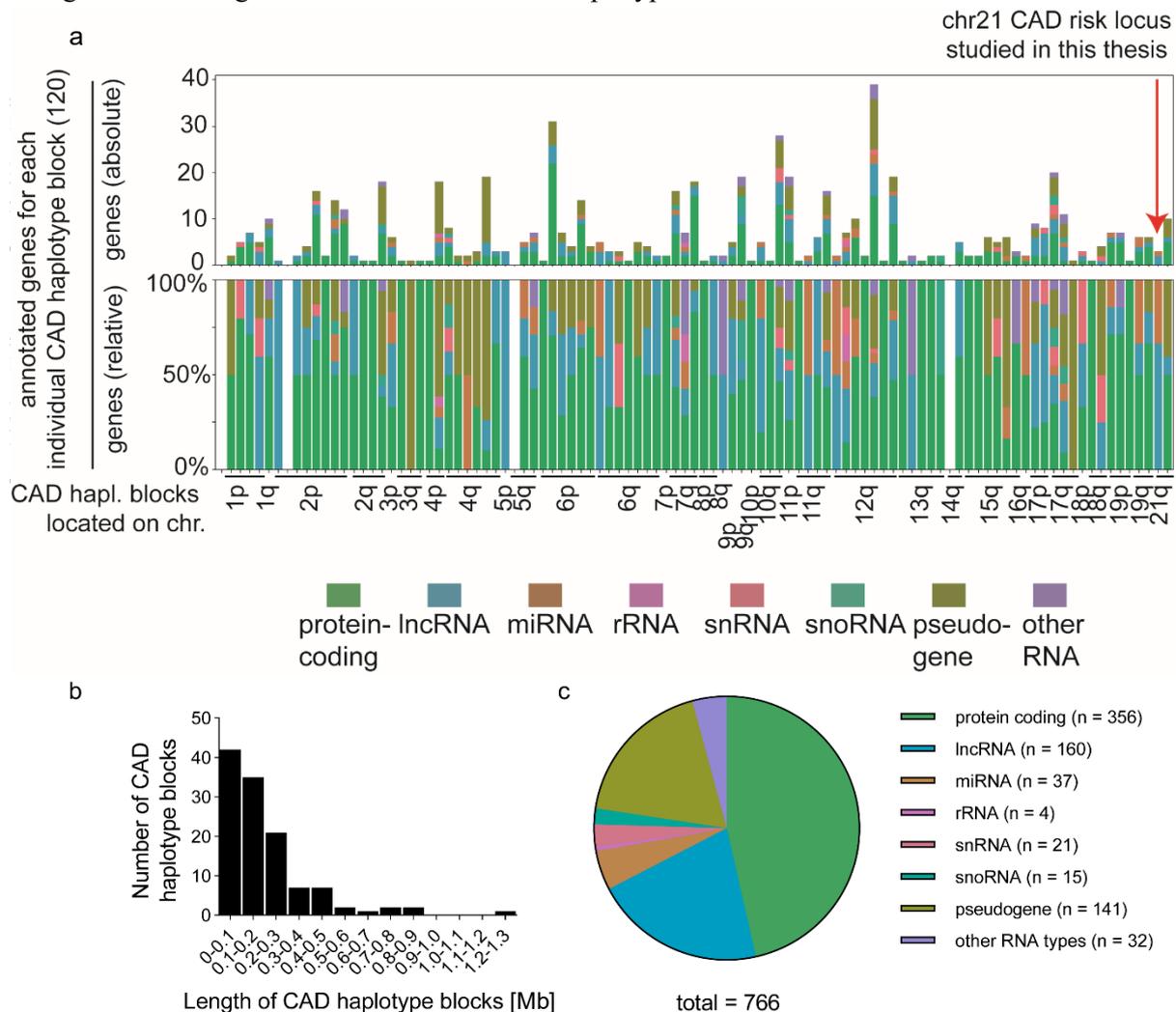


Figure 8 | Annotated genes in haplotype blocks. Gencode Database V19 (hg19) was used to identify annotated genes in the defined 120 haplotype blocks. Absolute (top panel) and relative (bottom panel) distribution of annotated genes within each individual haplotype block (a). See Supplementary Table 15 for defined chromosomal

position of all 120 atherosclerosis risk haplotype blocks shown in the figure. Length distribution of all 120 haplotype blocks in megabases (b) and proportion of total annotated gene classes (n = 766 genes) over all 120 haplotype blocks (c). Note the significant presence of non-protein-coding elements, and especially of lncRNAs (blue) to atherosclerosis risk loci.

In total, 766 genes distributed over all 120 haplotype blocks were found. The majority of genes were protein coding (46.5%), followed by lncRNAs (20.9%) and pseudogenes (18.4%), with the remaining RNA classes collectively contributing only 14.2% of all annotated genes. The mean length of a haplotype block is ~ 200 kb, while the median length is ~ 150 kb. Most of the haplotype blocks (35%) are < 100 kb, while lengths > 500 kb remain an exception (5%). The individual haplotype block contains a mean of 6.4 genes and a median of 3.5 genes. The majority of haplotype blocks (61.7%) contain ≤ 5 genes, while some blocks contain up to 38 genes. The number of genes correlates with the length of the associated haplotype block.

5.2.3 Candidate haplotype blocks for future experimental studies

Most of the haplotype blocks (99 of 120, 82.5%) contain at least one protein-coding gene. Given that noncoding regions have recently found to often exert regulatory roles in expression control of neighbouring genes, of special interest were the haplotype blocks containing long non-coding RNAs or no annotated genes (termed “gene deserts”). Since our lab focuses mostly on regulatory mechanisms controlling atherosclerosis onset by alterations in cells of the vascular wall, we also focused on such noncoding haplotype blocks that contained no SNPs associated with any risk factors from our list (which are all of systemic nature). According to these criteria, the top candidate haplotype blocks and their estimated effect size are listed in Table 24 arranged with p-values in descending order.

Table 24 | List of non-protein-coding candidate haplotype blocks, used for selecting a locus for experimental studies in the thesis

Locus	Content	GWAS cohort size cases/controls	Lead SNPs	p-value	Effect size (OR)
Chr13q22.1	Non-coding RNA	88/517	rs12429889	5.28×10^{-20}	1.64
Chr21q22.11	Non-coding RNA	60 801/123 504	rs28451064 rs9982601	1.33×10^{-15} 1.33×10^{-13}	1.14 1.12
Chr12q21.2	Non-coding RNA	88/517	rs7307780	5.03×10^{-15}	1.45
Chr5p15.33	Non-coding RNA	194/1539	rs11748327	5.30×10^{-13}	0.80
Chr2p24.1.00	Gene desert	1515/5019	rs2123536	6.83×10^{-11}	1.12
Chr11q22.3	Non-coding RNA	60 801/123 504 21 408/19 185 60 801/123 504	rs2128739 rs974819 rs2019090	7.05×10^{-11} 2.41×10^{-9} 3.60×10^{-9}	1.07 1.07 1.07
Chr8q24.13	Non-coding RNA	63 746/130 681	rs2954029	4.75×10^{-9}	1.06
Chr12q24.31.1	Gene desert	88 192/162 544	rs11057830	5.65×10^{-9}	1.07
Chr2q22.3.1	Non-coding RNA	88 192/162 544	rs2252641	6.77×10^{-9}	0.96
Chr1p32.3	Gene desert	2967/3075	rs11206510	9.60×10^{-9}	1.15
Chr2p24.1.0	Non-coding RNA	60 801/123 504	rs16986953	1.45×10^{-8}	1.09
Chr1q25.3	Non-coding RNA	669/1641	rs10911021	2.04×10^{-8}	1.36
Chr14q31.3	Gene desert	88/517	rs11624056	3.00×10^{-8}	1.43
Chr18q21.32	Non-coding RNA	60 801/123 504	rs663129	3.20×10^{-8}	1.06
Chr5q22.2	Gene desert	88/517	rs4621553	4.12×10^{-8}	1.40
Chr17q21.32.1	Gene desert	88 192/162 544	rs17608766	4.14×10^{-8}	1.07
Chr6q13	Non-coding RNA	34 622/88 766	rs9351814	4.93×10^{-8}	NA

All listed haplotype blocks contain no protein-coding RNAs but at least one lncRNA (11 haplotype blocks) or no genes at all (6 haplotype blocks). As soon as at least 1 bp of a gene reached into the haplotype block or if a gene was lying in the haplotype block completely, it was considered a gene of the block. The according GWAS are both studies with smaller sample sizes as well as large consortia studies comprising over 100 000 cases and controls, thus being more powerful, and thus rule out the possibility that the haplotype blocks we focused on came from mediocre GWAS. The top five candidate haplotype blocks, as defined by the p-value for the lead SNP, did not contain any SNPs associated with risk factors.

The haplotype block on chromosome 21 (Chr21q22.11) is not yet functionally understood. It contains two SNPs associated with atherosclerosis and no SNPs associated with any risk factors. It shows a respectable effect size (1.14 and 1.12, respectively) while being identified in multiple large GWAS cohorts, one consisting of 60 801 cases and 123 504 controls. The other top candidate locus on chromosome 13 only was backed up by one SNP found in one GWAS with noticeable lower cohort size. This made the chromosome 21 locus more valid and an ideal candidate for downstream experimental studies in this thesis.

5.3 Characterization of Chr21q22.11 locus

5.3.1 Overview of locus on chromosome 21

The atherosclerosis haplotype block on Chr21q22.11 spans from rs28451064 (chr21:35 593 827) at the 5' end to rs4817645 (chr21:35 724 134) at the 3' end (light-red shaded area in Figure 9). The association with atherosclerosis in GWAS is largest in its 5' part. There is at least one major subblock visible at the 5' end of the haplotype block, expressed both in the D' and r^2 structure in the LD plots. The subblock spans from rs28451064 (chr21:35 593 827) to rs10470171 (chr21:35 652 644). The 3' end of the subblock is marked with a vertical red line within the light-red shaded area in Figure 9. Corresponding to the 3' end of this subblock, a relatively sudden decrease of strength of association of GWAS SNPs with atherosclerosis can be noted (red arrow in Figure 9). The entire haplotype block contains two lead SNPs associated with atherosclerosis, rs9982601 (chr21:35 599 128) and rs28451064 (chr21:35 593 827). Both are located within the haplotype block as well as the small subblock, more specifically within intron 2 of *AP000318.2* and intron 3 of *AP000320.7*, respectively. The SNP rs9982601 is marked as black line within the light-red shaded area in Figure 9, whereas rs28451064 marks the actual 5' border of the haplotype blocks. Both SNPs were found in the GWAS from Nikpay, M. et. Al (2015) with a p-value of 1.33×10^{-15} (rs28451064) and 1.33×10^{-13} (rs9982601). There are no SNPs associated with risk factors within the haplotype block.

The haplotype block contains (parts of) two lncRNAs *AP000318.2* and *AP000320.7*. The promoters of these two lncRNAs lie outside (5') of the haplotype block, with the haplotype block starting within intron 2 of *AP000318.2*. There is one more gene annotated within the haplotype block. It encodes a predicted miRNA (*AP000318.1*), which was, however, not confirmed bioinformatically until August 2019. There are multiple other annotated lncRNAs and protein coding genes in the regions surrounding the haplotype block. Among others, upstream there are the lncRNA *LINC00310* and the protein coding genes *MRPS6*, a mitochondrial ribosomal protein, *SLC5A3*, a membrane transporter, and *ATP5O*, a component of the mitochondrial membrane ATP synthase. Downstream, there are the lncRNA *AP000320.6* and the protein-coding genes *KCNE2*, a regulatory subunit of a potassium voltage-gated channel, *SMIM11*, an integral membrane protein, *KCNE1*, and *RCANI*, encoding a protein that interacts with calcineurin A. Figure 9 shows the overview of chromosome 21, the strength of associations with atherosclerosis, the gene bodies of lncRNAs and protein-coding genes in and around the haplotype block, as well as the haplotype block structure.

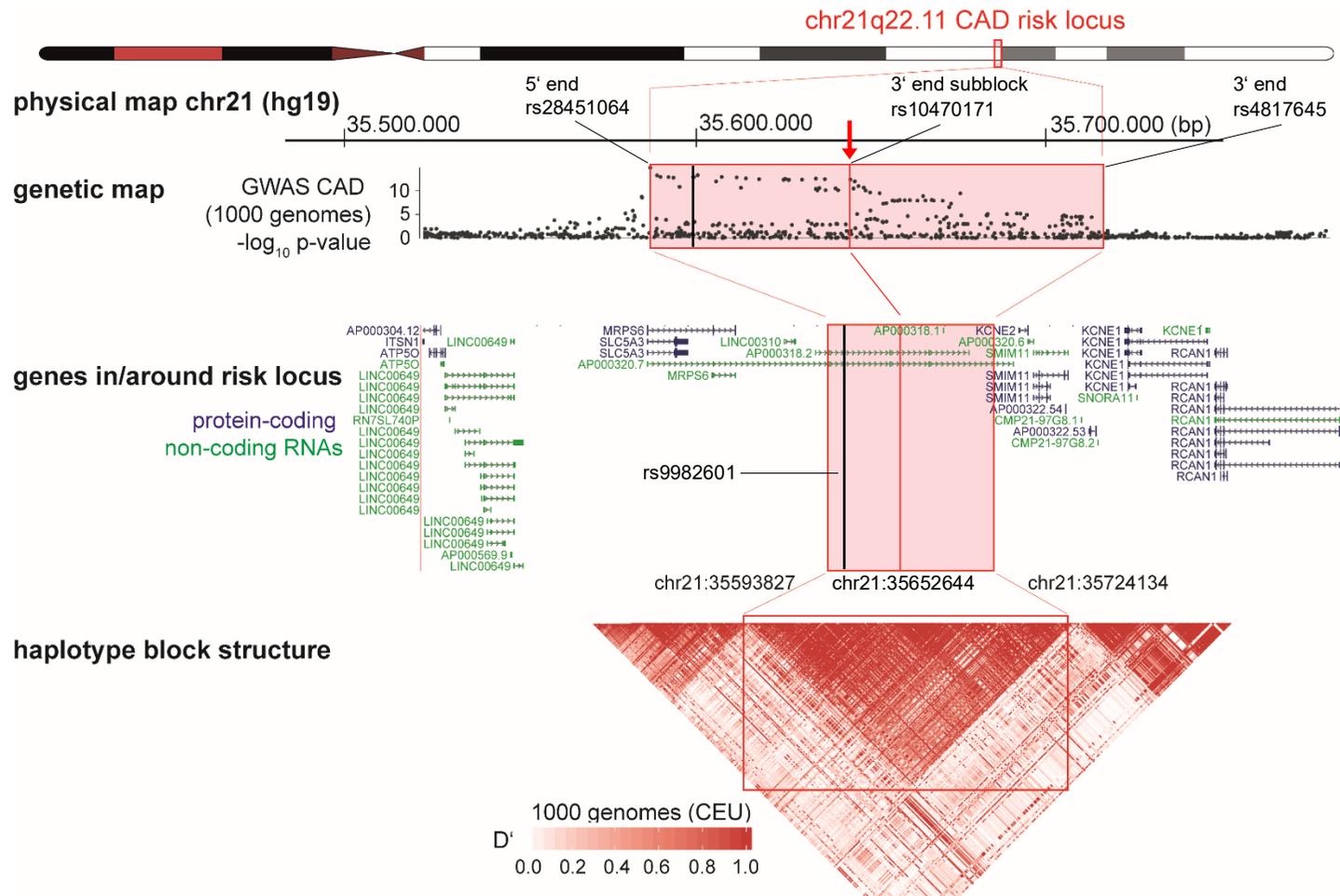


Figure 9 | Overview of the studied chromosome 21 atherosclerosis risk locus. Top, ideogram of chromosome 21 and positions of locus begin and end in base pairs (red boxed region). Below, atherosclerosis risk associations as measured by GWAS whereby strength of association is depicted as the logarithm of adjusted p-value in the GWAS as black dots. Note the drop in association strength highlighted by the red arrow. Below, genes encoded in the chromosome 21 risk locus and the immediate neighbouring genes. Note that atherosclerosis risk-associated genetic variation resides in an exclusively non-protein-coding core region. Bottom, map of linkage disequilibrium with marked borders of haplotype block in D' (indicator for linked inheritance). The haplotype block encompassing the chromosome 21 atherosclerosis risk locus spans from chr21:35 593 827 to chr21:35 724 134 bp (light-red shaded area). The strength of association with atherosclerosis in negative log-fold change (top) was extracted from Nikpay, M. et al. (2015).

To explore the RNA expression of the centrally located lncRNAs and transcripts from the neighbouring genes in vascular tissues and relevant cell types of the vascular wall, primers for qPCR were designed. Primers for individual exons of *LINC00310*, *AP000318.2*, and *AP000320.7* were designed based on EST annotations. EST-based transcript information was all information available at this locus, which had been considered a “gene desert” at the time this part of the thesis was started. A detailed overview of the used nomenclature for qPCR primers is shown in Figure 10. This nomenclature will be used for the remainder of the thesis.

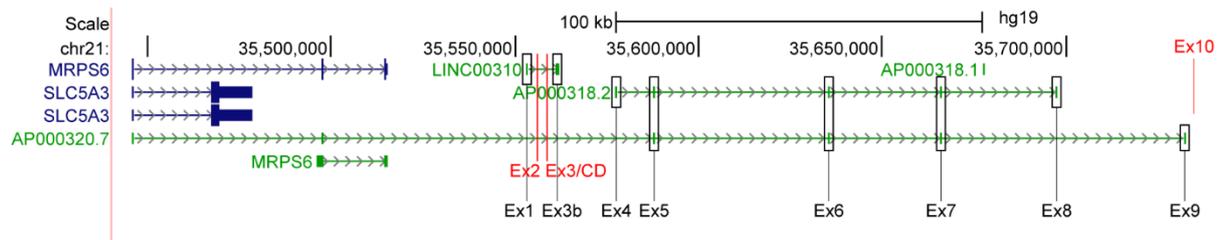


Figure 10 | Primers for qPCR to probe long noncoding RNA expression from introns and exons located within the core chromosome 21 atherosclerosis risk locus. Black font indicating PCR primers targeting exons of contained lncRNAs (*LINC00310* or *AP000318.2*), red font showing positions of primers lying in their intergenic or intronic sequences. Exon/intron structure is based on former EST annotation which was all information available at the start of this thesis. Up-to-date gene structure models for the lncRNAs at the locus are from the more recent hg19 genome annotation from the UCSC genome browser.

Primers Ex1 (forward) and Ex3b (reverse) both map on the first and second (= last) exon of *LINC00310*. Ex2 (reverse) and Ex3 (reverse) /CD (forward) map into intronic sequences of *LINC00310* based on former EST annotations. Ex4 (forward) maps on the first exon of *AP000318.2*, while Ex5 (forward and reverse), Ex6 (reverse), and Ex7 (reverse) map on exons both for *AP000318.2* and *AP000320.7*. The last exons of *AP000318.1* and *AP000320.7* are covered by Ex8 and Ex9, respectively (both reverse). Ex10 (reverse) maps into an intergenic area, based on former EST annotation. Using these primers, we performed qPCRs to probe whether noncoding RNA transcription occurred in the > 100 kb region containing *LINC00310*, *AP000318.2*, and *AP000320.7*. Indeed, we confirmed RNA transcribed from all tested regions in vascular cell types, and iPSC cells (data shown in detail in the context of functional analysis of the locus in chapters 5.5.1 and 6.3)

5.3.2 Expression of genes on chromosome 21 locus in tissues and cells

The expression levels of genes of the haplotype block on chromosome 21, including non-protein-coding transcripts from the previous chapter, was determined more comprehensively by exploring datasets from the GTex Portal (<https://gtexportal.org/>), a public resource collecting expression from hundreds of tissues in >1000 individuals based on high throughput RNA and genome sequencing efforts. We investigated our candidates manually in the dataset in order to evaluate their potential importance for pathophysiological processes specifically in vascular tissues and cells of the arterial wall (Figure 11). The lncRNAs *AP000318.2* and *LINC00310* showed expression especially in the arterial tissues of aorta,

coronaries, and tibial artery, while there was no gene expression found for the ultra-long lncRNA *AP000320.7* (Figure 10). Most protein-coding genes showed ubiquitous expression throughout all tissues covered by the GTex Portal. Solely for *SLC5A3*, there was no gene expression found. While *MRPS6* and *KCNE2* showed a light peak in arterial tissue (aorta, coronary, tibia), this observation could not be made for the remaining next-neighbouring genes, being *ATP5O*, *RCANI*, *KCNE1* and *SMIM11*. The expression levels of relevant genes are depicted in Figure 11.

To summarize, the observed expression of many genes in the chromosome 21 locus in arterial tissues like aorta, coronary and tibial artery grants the possibility of relevance in pathophysiological mechanisms of atherosclerosis, as will be examined in the following experiments.

5.4 Design of experimental studies

5.4.1 Knockout of haplotype block and subblock

For experimental examination of the atherosclerosis relevant chromosome 21 locus, knockouts of the haplotype block (chr21:35 593 827-35 724 134), as well as subblock (chr21:35 593 827-35 652 644), were performed (Figures 9 and 12). The nomenclature for the knockouts of the haplotype block and the subblock was KO 1 and KO 2, respectively (Figure 12).

The CRISPR/Cas9 system was used to introduce deletions in desired cell types. This was achieved by cutting sites at both the 5' end as well as 3' end of the haplotype block and subblock. This was first tested in HEK-293 cells and then applied in iPSCs. Therefore, multiple gRNAs per target site were designed and, subsequently, gRNA-containing DNA oligonucleotides hybridized and cloned into an expression plasmid. The resulting U6 promotor-driven gRNAs-encoding plasmids were produced via transformation in *E. coli*. Surveyor assay was used to confirm functionality of different gRNAs. For each cutting site, multiple gRNAs were designed and the best working combination with least reported off-target effects was used, as listed in Supplementary Table 7. Finally, knockouts were generated by concurrent transfection via electroporation of 5'-gRNA with 3'-gRNA-containing plasmids with the Cas9 expression plasmid, which contained GFP as a selectable marker. This resulted in the introduction of two double strand breaks (DSB), being repaired via NHEJ in order to complete the knockout. Validation of the knockouts was performed by PCR with primers sitting upstream of the 5' cutting site and downstream of the 3' cutting site. Without a positive knockout, the resulting PCR products would theoretically span 140 kb and 65 kb for KO 1 and KO 2, respectively, which was impossible to achieve. Therefore, a short PCR product of defined length (Supplementary Table 6) resulted only if the knockout was successful, as intervening genomic DNA had been deleted. Additionally, Sanger sequencing of received PCR products was performed to confirm joint cutting sites. In Figure 12, the exact regions of the knockouts, as well as affected exons and used primers with genomic sequence of PCR products, are shown.

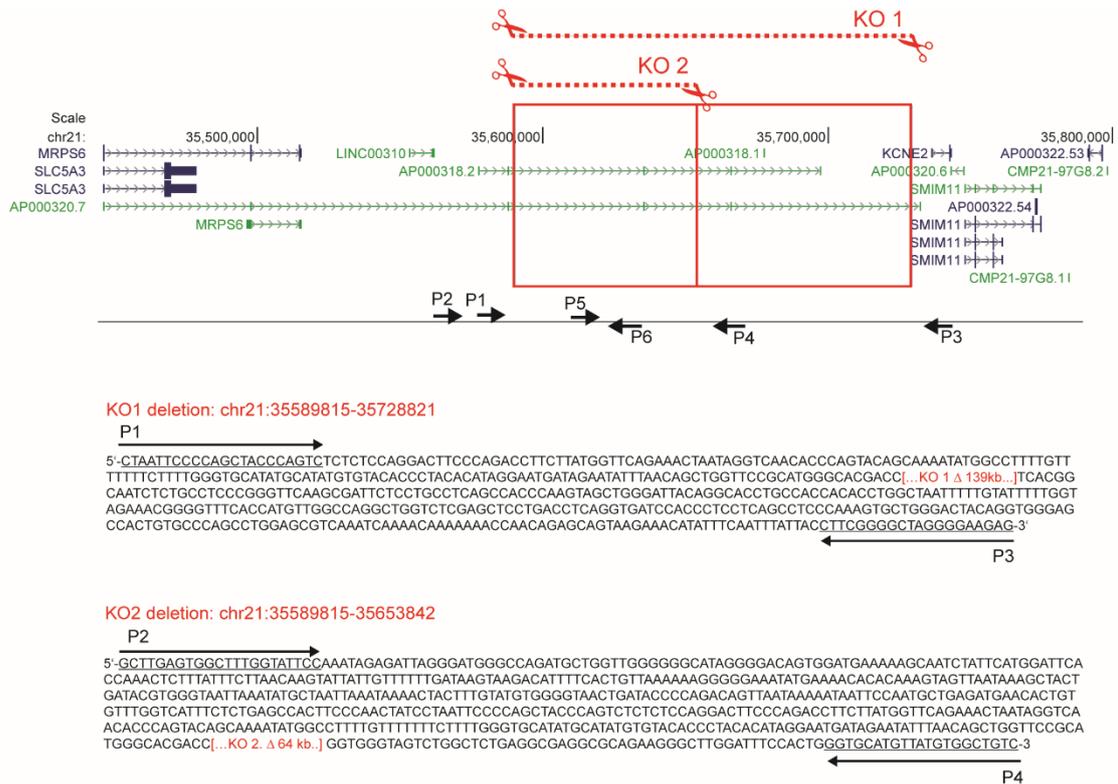


Figure 12 | Overview of CRISPR/Cas9-mediated knockout strategy on chromosome 21 locus. KO 1 spans across ~ 139 kb (chr21:35 589 815-35 728 821). KO 2 spans across ~ 64 kb nested within KO 1 (chr21:35 589 815-35 653 842). For KO 1, primer pairs P1 and P3 were used for documenting successful knockout by PCR. A PCR product only arises when the intervening sequence of an allele is deleted. For KO 2, primer pairs P2 and P4 were used for deletion validation by PCR. Successful knockout should theoretically result in PCR products of 503 bp for KO 1 and 624 bp for KO 2, respectively. Note that KO 1 results in knockout of Ex3-5 of *AP000318.2* and Ex4-5 of *AP000320.7*, while KO 2 results in knockout of Ex3 of *AP000318.2* and Ex4 of *AP000320.7*. PCR primer pair P5/P6 was used to screen for sequences deleted in both knockouts, KO 1 and KO 2, and hence for testing homozygosity or heterozygosity of knockout (Figure 13 for experimental results).

Before performing the knockouts in iPSCs, the genotype of rs9982601 in hiPSCs was assessed using melt-curve based analysis. This served to verify that the iPSC cell line we used did not by chance stem from a donor individual that carried the potentially atherosclerosis-risk conferring genetic variation but was wild type. The atherosclerosis risk SNP rs9982601, located on position chr21:35 599 128 is defined by genotype T (11.002%) versus the wild type allele C (88.998%). In hiPSCs the wild type allele C was present.

CRISPR/Cas9-treated (GFP-positive) hiPSCs were selected using FACS by green fluorescence in order to enhance outcome. Subsequently, GFP-positive hiPSCs were seeded in extreme low density to allow for single cell colonies to grow. After expansion, multiple single cell colonies per knockout were manually picked, expanded, and checked for positive knockout clones by analysis of their genomic DNA with the PCR assays P1/P3 for KO 1, and P2/P4 for KO 2, as indicated in Figure 12. In total, three biological replicates of KO 1 were created, which were named as follows: KO 1 – 1.A7, KO 1 – 1.C12, KO 1 – 4.H5. For KO 2, four biological replicates were created: KO 2 – 3.H1, KO 2 – 2.A1, KO 2 – 2.A7. The exact breakpoints of the knockouts were verified, with subsequent Sanger sequencing of genomic DNA: The knockouts were rather accurate, with two lines each showing some small erosion at

the breakpoints, and one line showing the exact breakpoint. The theoretically expected sequences after CRISPR/Cas9 treatment and the actual sequences, as well as the flash gels of KO 1, KO 2 and a wild type control, are depicted in Figure 13. Since the PCR strategy used to detect the knockouts did not distinguish whether one or both alleles had been targeted, we designed a second PCR primer pair located in an internal region commonly deleted by KO 1 and KO 2. This latter primer pair should have enabled the forming of a PCR product only if the locus had not been homozygously deleted (Figure 13f). Since we detected a PCR product from this region in all 6 KO iPSC cell lines (Figure 13f), we conclude that both KO 1 and KO 2 are heterozygous.

a KO 1 - 1.A7

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGACC [...139 kb...]TCACGGCAATCTCTGCCTCCCGGGTTCAAGCGATT.....-3'

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACG-----GCAATCTCTGCCTCCCGGGTTCAAGCGATT.....-3'

KO 1 - 1.C12

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGACC [...139 kb...]TCACGGCAATCTCTGCCTCCCGGGTTCAAGCGATT.....-3'

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGAC-----GCGATT.....-3'

KO 1 - 4.H5

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGACC [...139 kb...]TCACGGCAATCTCTGCCTCCCGGGTTCAAGCGATT.....-3'

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGACC-----TCACGGCAATCTCTGCCTCCCGGGTTCAAGCGATT.....-3'

b KO 2 - 3.H1

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGACC [...64 kb...]-----GGTGGGTAGTCTGGCTCTGAGGCGAGGCGCAGAAG.....-3'

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGACC-CTTCATGTTACTATGGTGGGTAGTCTGGCTCTGAGGCGAGGCGCAGAAG.....-3'

KO 2 - 2.A1

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGACC [...64 kb...]GGTGGGTAGTCTGGCTCTGAGGCGAGGCGCAGAAG.....-3'

5'-.....GAATATTTAACAGCTGGTTCCGCATGGG-----TGGGTAGTCTGGCTCTGAGGCGAGGCGCACAAG.....-3'

KO 2 - 2.A7

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGACC [...64 kb...]GGTGGGTAGTCTGGCTCTGAGGCGAGGCGCAGAAG.....-3'

5'-.....GAATATTTAACAGCTGGTTCCGCATGGGCACGACC-----GGTGGGTAGTCTGGCTCTGAGGCGAGGCGCAAAG.....-3'

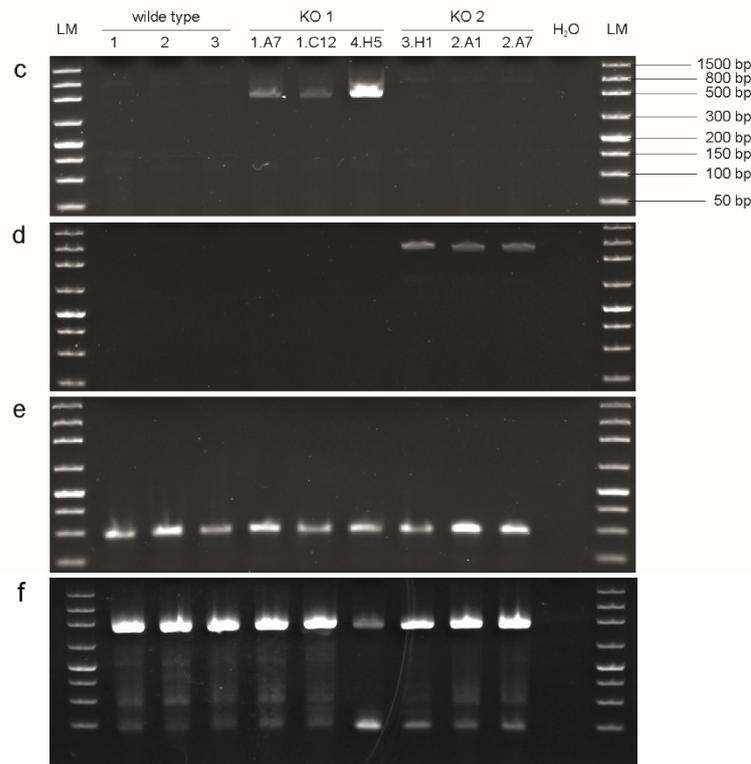


Figure 13 | Knockout sequences and PCR-genotyping of knockout (agarose flash gel). Theoretical (upper line) and actually sequenced (lower line) DNA sequences of the borders of the KO 1 (a) and KO 2 (b). Theoretical sequences were determined by assembly of expected CRISPR/Cas9 cutting sites. Actual sequences were obtained by Sanger sequencing. Cut-out fragments sized 139 kb for KO 1 and 64 kb for KO 2, respectively. For validating the knockouts by semiquantitative PCR, genomic DNA from targeted cells was obtained by single cell-picking and agarose flash gels of PCR products of expected size (Figure 12) using primer pair P1/P3 spanning the genomic

deletion in KO 1 (c) or using primer pair P2/P4 for verifying KO 2 (d) by spanning the deleted region, or using an unrelated primer pair to an untargeted locus elsewhere (S100) as positive control (e) for PCR. Note that we isolated three independent clones with a targeting event per knockout. See Figure 12 for the positioning of primers relative to the deleted region. For evaluation whether CRISPR/Cas9 knockout was heterozygous or homozygous in KO 1 and KO 2, PCR primer pair P5/P6 (Figure 12) was used to detect the absence or presence of a region commonly deleted by both knockouts, KO 1 and KO 2 (f). Note that a PCR product is formed, indicating the heterozygous nature of knockout for all of our 6 clones.

5.4.2 Differentiation of hiPSCs into cell types associated with atherosclerosis

For all future experiments, we chose to pool lines KO 1 – 1.A7, KO 1 – 1.C12, KO 1 – 4.H5 for KO 1, as well as KO 2 – 3.H1, KO 2 – 2.A1 and KO 2 – 2.A7 for KO 2. From now on, “KO 1” and “KO 2” refer to these pooled lines, and the term “knockout” refers to the CRISPR/Cas9 lines, created in this work (iPSC or differentiation derivatives), while “wild type” refers to the non-edited primary cells (iPSC or differentiation derivatives). To analyse whether deletion of the risk locus conferred any functional changes to cells we investigated effects on basic cellular functions: Specifically, differentiation of hiPSCs into SMCs and ECs was performed in order to gain insight into pathophysiological processes of cells relevant for atherogenesis. By inducing the knockouts in hiPSCs, later differentiation into different cell types with the exact same genetic background greatly improved the relevance for interpretations. We chose to differentiate iPSC to SMCs and ECs using a described protocol and proceeding via a common mesodermal progenitor as intermediate.⁷⁸ Both SMC and EC differentiation took 7 days to complete. Starting as hiPSCs at day 0, the cells were induced to reach a mesodermal state by day 4. Subsequently, the differentiation into SMCs and ECs was initiated with the final phenotype of SMCs and ECs being reached by day 7. During differentiation, different cell type-specific and expected molecular markers were analysed using total RNA isolation. Successful differentiation into the desired cell types was analysed using qRT-PCR. Therefore, molecular markers exclusively expressed by the desired cell types were determined. The pluripotent status of stem cells was analysed using the molecular markers *SOX2*.^{78,82} As markers for the mesodermal transitional period, *NKX2.5*, *ISL1*, and *TBX18* were used.⁸³⁻⁸⁵ Differentiation into ECs was analysed using endothelial cell specific markers *CD144*, *ICAM1* and *VEGF*.⁸⁶ Differentiation into VSMCs was analysed using smooth muscle cell specific markers *CNN1*, *MYH11* and α -*SMA*.⁸⁷ Therefore, total RNA was isolated on days 0, 3, 5, and 7 in biological replicates of 3 which were analysed in technical replicates of 4. In Figure 14, median of RNA expression values for differentiation markers during time course of differentiation are displayed in a heatmap.

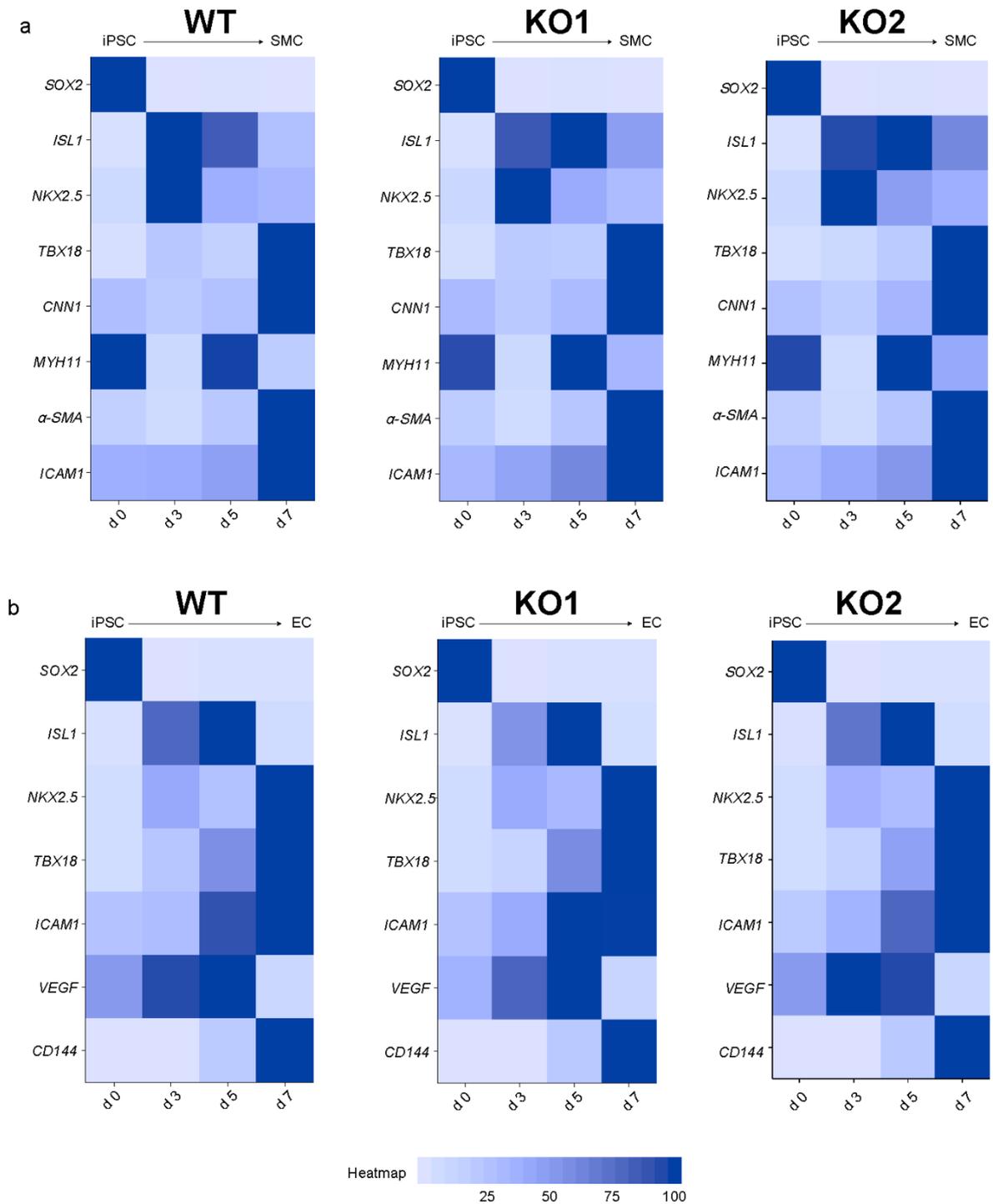


Figure 14 | Time course of differentiation markers. Differentiation of iPSCs into SMCs (a) and ECs (b). Wild type controls are displayed on the left side, followed by KO 1 (middle) and KO 2 (right). Established markers for stem cell fate (*SOX2*), for intermediate mesenchymal lineage and for EC and SMC fate were quantified by qPCR. The heatmap compares levels of iPSCs (day 0) up until their final EC/SMC forms. Hereby, the maximum expression of a single gene during that time course was set 100% and compared to the relative expression of the same gene during the differentiation at days 0, 3, 5 and 7.

The pluripotent state of hiPSCs was confirmed using *SOX2*, and the knockouts did not impair pluripotency as such, which was consistent with the fact that knockout iPSCs continued to proliferate

and grow in typical stem cell clones. After induction of mesoderm in the culture, *SOX2* decrease was observed as expected (~ 1000:1) in both control iPSCs and in the knockout iPSCs KO 1 and KO 2. During the transitional period in mesodermal state, *ISL1* as marker for progenitor cells of both SMCs and ECs and *NKX2.5* as marker for SMC progenitors both showed an increase as expected, also in the knockout lines. The differentiation of SMCs in the knockouts showed an increased expression of *CNN1* (~ 3:1) and *α -SMA* (~ 6:1) as general markers for SMCs as in the wild type control. Expression of *MYH11* varied during differentiation already in the control wild type iPSC, and was low in final SMCs, indicating the non-contractile phenotype form of the cultured SMCs. The differentiation of ECs showed an increased expression of *VEGF* during the transformation into ECs. Then, final EC phenotypes showed decreased *VEGF* expression, while *ICAM* and *CD144* expression as indicators for differentiated ECs both peaked (~ 1:4 and ~ 1:800, respectively). No differences in these expression patterns were observed in either knockout line compared to control or to each other. We have, however, only measured a small number of markers, and whether the knockouts affect SMC or EC differentiation capacity will have to be assessed in more detail by comparing entire transcriptional profiles (eg, by RNAseq), and in particular by performing dedicated functional cellular and biochemical and biomechanical test. In conclusion, both SMC and EC differentiation continued successfully in both wild type and knockout cell lines and levels and dynamics of differentiation markers were not altered. Therefore, knockouts, at least in their heterozygous state, did not influence proper differentiation into atherosclerosis-relevant cell types or change cell fate.

5.5 Comparison of wild type versus knockout cells

5.5.1 Altered gene expression of knockout cells

Next, the gene expression for the non-coding RNAs and the immediately adjacently located protein-coding genes was compared in SMCs and ECs between wild type and knockout cell lines in order to examine effects of the knockout on a molecular level. We performed this analysis because many transcribed non-coding loci are known to be functional as *cis*- and *trans*-acting elements (e.g. enhancers).⁸⁸ Therefore, RNA was isolated, and qRT-PCR was used to determine expression levels 5 days after differentiation of wild type and knockout iPSC lines into final SMC and EC cell types (day 12).

To analyse the two lncRNAs *AP000318.2* and *AP000320.7* encoded in the locus we used three different primer pairs (Figure 10 for their location). In addition, *SLC5A3*, *MRPS6*, *KCNE1*, *KCNE2*, *SMIM11*, *ATP5O* and *RCAN1* expression levels were detected. In SMCs, a general increase of gene expression in both knockouts KO 1 and KO 2 was observed when compared to wild type controls, and both noncoding and protein-coding genes were affected (Figure 15). For quantification, means of technical replicates (n = 4) for each biological replicate (n = 3) were compared. Expression levels were adjusted to equal amounts of input RNA from cells, quantified as copy numbers and then data were normalized relative

to levels in wild type control (100%). The significance was tested by multiple t-test after validating normality distribution of the data and correction for multiple testing using Holm-Sidak method.

In Figure 15, the RNA expression of selected genes in the targeted chromosome 21 locus is shown for both SMCs and ECs in each wild type, KO 1, and KO 2. Knockouts were compared relative to 100% wild type control for each of the selected genes in both SMCs and ECs.

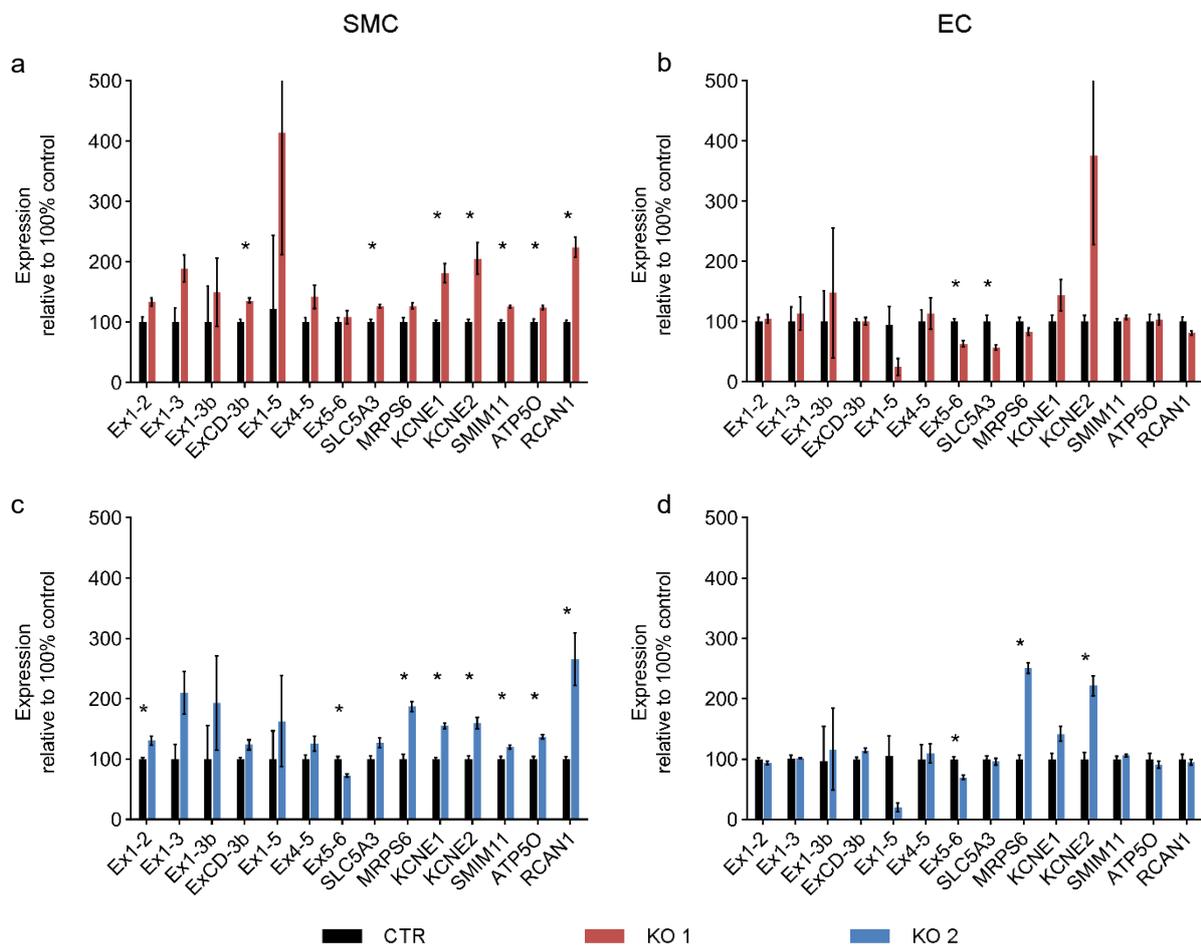


Figure 15 | RNA expression of genes within and near the chromosome 21 atherosclerosis risk locus in iPSC-derived SMCs and ECs. Abundance of steady-state RNA transcript levels (lncRNAs or protein-coding) was determined by qRT-PCR using Taqman probes. For quantification, expression levels were adjusted to equal amounts of input RNA from cells, quantified as copy numbers and then data were normalized relative to levels in wild type control (100%). Samples were total RNA isolated 5 days after differentiation of iPSCs into the respective SMC and EC lineages (day 12). KO 1 (a, b) is displayed in red. KO 2 (c, d) is displayed in blue. Wild type controls (parental iPSC cell line) are displayed in black. The expression levels of most tested genes in knockouts in SMCs (a, c) were increased, while knockouts in ECs (b, d) showed a mixed pattern (some increased and some decreased expression in both knockouts). Bars represent means from quadruplicate measurements of all three biological replicates each for control and both knockouts. * (adjusted $p \leq 0.05$). Significance was tested by multiple t-test after validating normality distribution of data and correction for multiple testing using Holm-Sidak method.

When analysing the effects in ECs, we found that KO 1 and KO 2 slightly differed in the effect on lncRNAs and in whether *SLC5A3* and *KCNE2* were affected or not. In conclusion, knockouts showed altered gene expression in SMCs and ECs for both KO 1 and KO 2. The different pattern in SMCs and

ECs suggest a cell type specific impact of the knockouts, with SMCs tending to show a uniformly increased expression pattern, while ECs present a more selective pattern and have effects on some local genes but not all, depending on which DNA region was deleted.

5.5.2 Altered cellular functions of knockout cells

5.5.2.1 Altered apoptosis phenotype in SMCs and ECs

Another atherosclerosis-relevant feature is the stage-dependent propensity to apoptosis. Functional studies of apoptosis sensitivity in iPSC-derived SMCs and ECs were performed, exploring the effect of chromosome 21 knockouts on this atherogenic behaviour. Therefore, caspase 3/7 activities, as indicators of engaged apoptosis, were measured in challenged conditions (proapoptotic 0.25 μ M staurosporine treatment) and in untreated baseline conditions (medium control) in both knockout cell lines and in controls. Comparison of treated versus untreated cell lines revealed that apoptosis could be detected in principle and validated the successful experimental approach. Then, untreated, and treated cell lines were compared as shown in Figure 16, using 1way ANOVA with Tukey correction for multiple comparisons after testing for normal distribution using D'Agostino-Pearson test.

First, in untreated conditions (no apoptosis induced) and using wild type cells as reference value (100%) we found that in SMCs, KO 2 showed a significantly increased apoptosis ($P = 0.03$). The mean increase of apoptosis was 149% in KO 1 (not significant) and 160% in KO 2. In ECs, the increased apoptosis of both KO 1 and KO 2 cell lines was not significant (Figure 16 a, b).

Second, treated (apoptosis induced) wild type cells were compared to treated knockout cells. Also, wild type cells served as reference value and were set to 100%. The cell lines of KO 1 and KO 2 were adjusted for their increased baseline apoptosis, revealed in the first calculations of untreated cell lines. For SMCs, this was a factor of 1.5 and 1.6 for KO 1 and KO 2, respectively. For ECs, this was a factor of 1.2 and 1.5 for KO 1 and KO 2, respectively. After adjustment in SMCs, both KO 1 ($P = 0.01$) and KO 2 cell lines ($P < 0.004$) showed a significantly decreased apoptosis. The mean decrease of apoptosis was 74% in KO 1 and 70% in KO 2. After adjustment in ECs, KO 2 cell line showed a significantly decreased apoptosis with a mean of 71% ($P = 0.002$), whereas the decreased apoptosis of KO 1 cell line (mean of 97%) remained not significant (Figure 16 c, d).

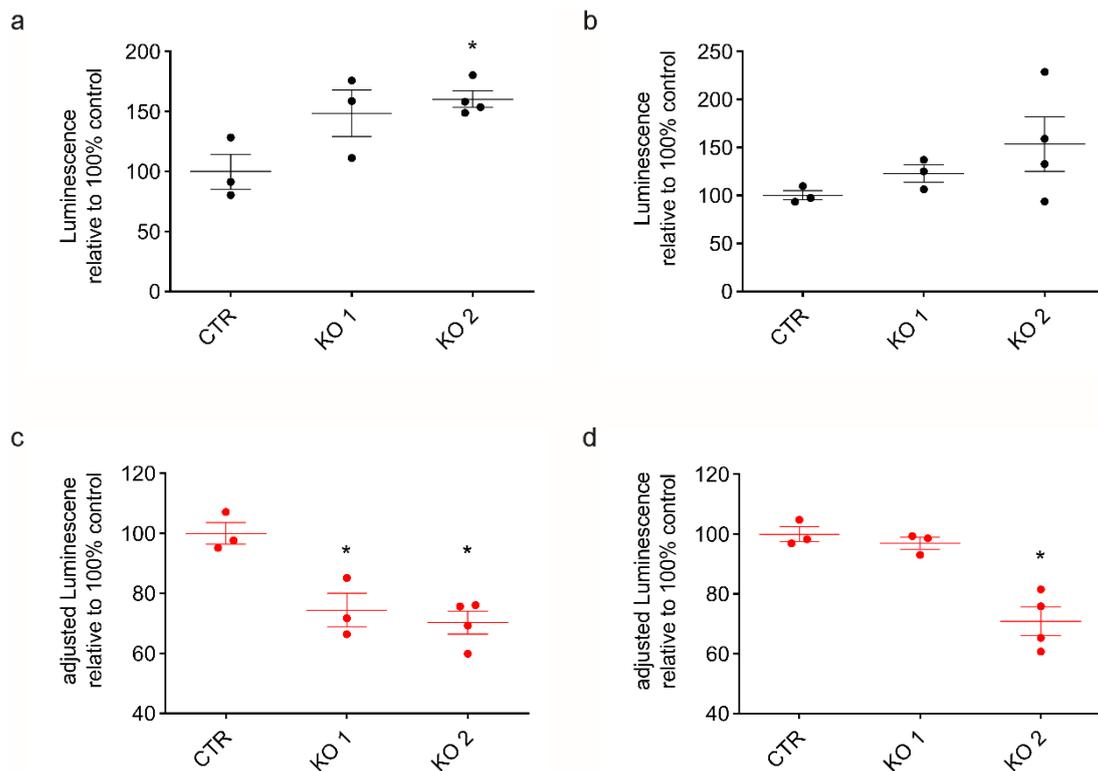


Figure 16 | Effect of chromosome 21 knockout on apoptosis levels in iPSC-derived SMCs and ECs.

In staurosporine induced apoptosis measuring phospho-caspase 3/7 activation (upper figures), an increased baseline apoptosis of KO 2 cell line was observed in SMCs compared to wild type SMCs (a), while for ECs (b) there was no significant increase in the knockout cell lines. In challenged conditions, after treating cells with 0.25 μM staurosporine as apoptosis trigger (lower figures), after adjustment for increased baseline apoptosis in the unchallenged knockouts, there was relatively decreased apoptosis observed in both knockout cell lines in SMCs (c), while ECs (d) showed decreased apoptosis solely in the KO 2 cell line. Assays were performed in quadruplicates for every one of the 3 biological replicates of controls and both knockouts. Scatter dot plots show means with SEM. * (adjusted $p \leq 0.05$). Significance of differences in means was tested by 1way ANOVA with Tukey correction for multiple comparisons.

To summarize, there was a trend of increased baseline apoptosis observed in untreated knockout cell lines when compared to wild type cells, most prominent in KO 2. Whereas, in challenging conditions of already ongoing high levels of apoptosis (at least in the context of cell death induced by the broad-range ATP-competitive kinase inhibitor staurosporine) knockout cell lines showed a decreased level of apoptosis, compared to wild type cells. These results suggest potentially dual roles in apoptosis execution, and for example, an inferior response to induction of apoptosis in knockout cells.

5.5.2.2 Decreased proliferation phenotype in SMCs and ECs

Another cellular function with implications for atherosclerosis is cell proliferation rate. Specifically, functional studies of proliferation activity in iPSC-derived SMCs and ECs were performed, exploring whether chromosome 21 knockouts might lead to increased proliferation, which is considered a context-dependent atherogenic behaviour. Therefore, knockout cell lines and wild type controls of both iPSC-derived SMCs and ECs were seeded in equal densities and cell numbers were tracked microscopically, periodically, over several days in culture and measured as cell confluence (relative area

of culture vessel covered by cell bodies). Since the surface aspect area of the individual cells were unchanged between controls and knockouts (as determined qualitatively by visual microscopic inspection), an increased confluence in the cell culture plate can be correlated with an increased number of cells, hence proliferation. But more specific experiments (cell counts, or single live cell-tracking, or genetic clonal marking) are required to formally document proliferative changes. Subsequently, knockout cell lines were compared to wild type controls for each cell line as shown in Figure 17, using 1way ANOVA with Dunnett correction for multiple comparisons after testing for normal distribution using D'Agostino-Pearson test.

In SMCs, KO 1 showed a reduced proliferation rate compared to wild type controls: The diminished growth rate became statistically significant 36 hours after seeding for KO 1 (adjusted $P = 0.04$). By that time, cell numbers showed a mean 2.3-fold increase ($SD \pm 3.4$) for KO 1 when normalized to cell numbers at the time point of seeding (0 h). This is lower, compared to wild type cells, whose numbers had increased 2.7-fold ($SD \pm 19.7$). The most pronounced reduction in proliferation in KO 1 was observed after 70.75 hours with a mean delta of 165% (CI 4.9 to 387.5) compared to control (adjusted $P < 0.04$). A significant difference in proliferation was observed only mid-time, while in the beginning the dividing cells had to gain a lead, respectively lack behind. KO 2 did not show a significant difference in proliferation. This may be because the distribution of the sampling error for KO 2 was large at many measured time points, which indicated that a proliferative deficit in KO 2 is not yet so clear. For confirmation, the phenotype of KO 2 should be investigated in follow-up experiments using larger number of replicates.

In ECs, also, KO 1 showed reduced proliferation rate compared to wild type control. The diminished growth rates became statistically significant 46.85 hours after seeding for KO 1 (adjusted $P = 0.03$). At that time, cell numbers relative to 100% at seeding showed a mean increase of 1.8-fold for KO 1 ($SD \pm 14.7$) at the time of seeding compared to 2.6-fold ($SD \pm 55.4$) for wild type controls, documenting again a proliferation deficit in cells. The largest decrease in proliferation in KO 1 was observable after 126.85 hours at a mean delta of 179.2% compared to control (CI 23.6 to 334.7) (adjusted $P < 0.03$). KO 2, as in SMCs, did not show a significant difference in proliferation.

In summary, KO 1 showed reduced proliferation rates in both SMCs and ECs, while observable trends of diminished proliferation in KO 2, remained not significant. Complex stage- and cell-specific increased proliferation of cells is typically seen in atherosclerotic lesions.^{4,89,90} Dedicated conditional knockout *in vivo* experiments would be required though to investigate, whether the reduced proliferative capacity of the knockouts *in vitro* would also occur in mutation carriers *in vivo*. Even more importantly, it remains to be functionally evaluated *in vivo*, whether, in which stage, and in which cell type, any reduced proliferation would lead to net atheroprotection. Formally it is possible that context-dependent proliferative changes even worsen aspects of disease progression.

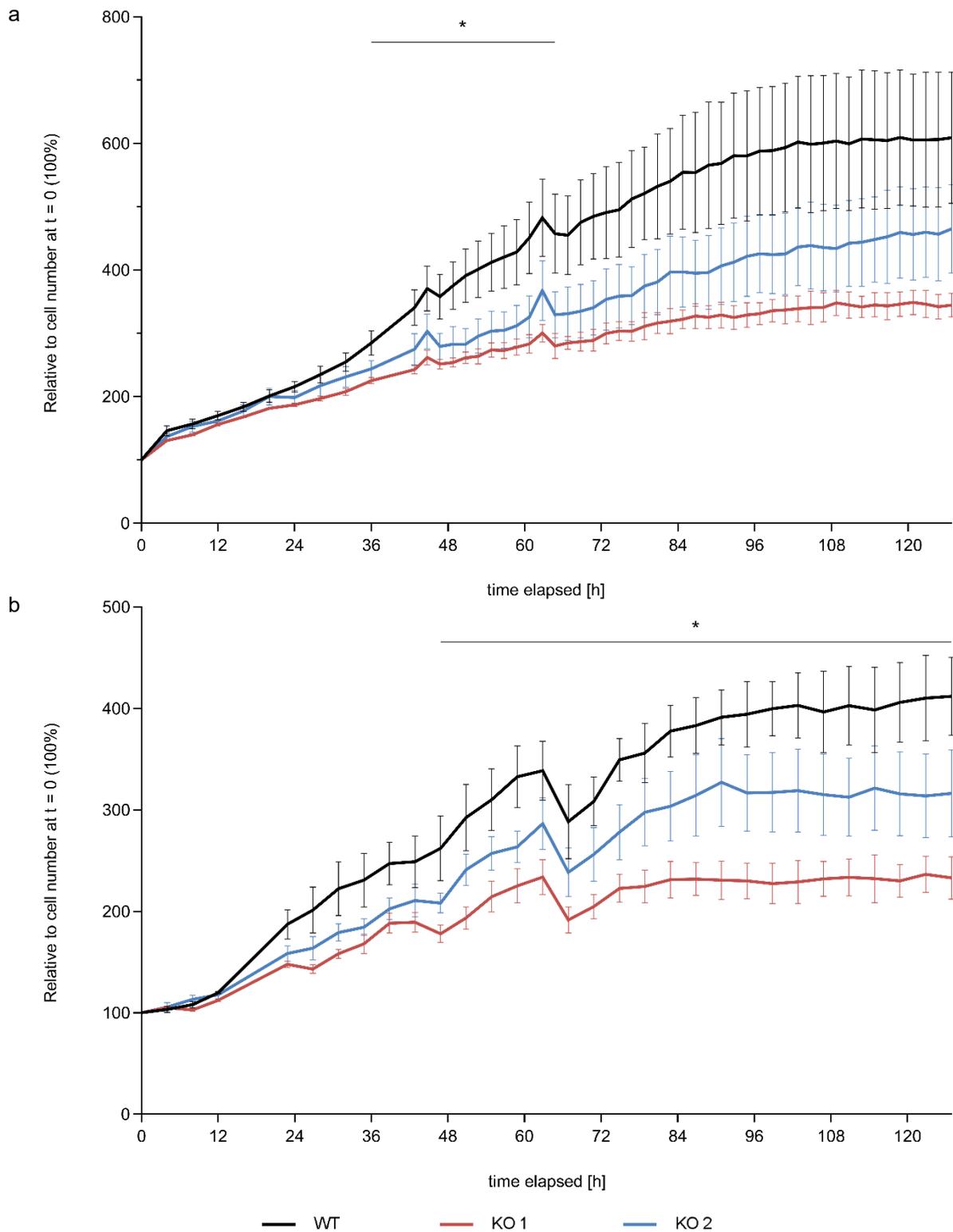


Figure 17 | Functional studies exploring proliferation rates in iPSC-derived SMCs and ECs. In SMCs (a), KO 1 (red) showed reduced proliferation, while the observed less pronounced reduction of proliferation for KO 2 (blue) was not significant. The significant deviance from wild type controls (black) was observable first at 36 hours after seeding for KO 1. In ECs (b), also KO 1 showed reduced proliferation, while KO 2 remained not significant. The significant deviance from wild type controls was observable 46.85 hours after seeding for KO 1. The time series shows means and \pm SEM from 3-4 experimental repetitions and quadruplicate measurements at each time point. For plotting, values for each replicate were normalized to the cell number at the time of seeding the cells for each replicate (0 h). * (adjusted $p \leq 0.05$). Dunnett's multiple comparison testing in repeated-measured 1way ANOVA.

5.5.2.3 Decreased migration phenotype in SMCs

Another cell phenotype we assayed was the migratory behaviour of SMCs, which is known to be increased in atherosclerotic lesions *in vivo* in humans as well as in many animal model systems.^{91,92} Functional studies of migration activity in iPSC-derived SMCs were performed, exploring the effect of chromosome 21 knockouts on this potentially atherogenic behaviour. Therefore, knockout cell lines and wild type controls of SMCs were seeded in similar cell counts and cultivated until fully confluent. An artificial “wound” (scratch assay) of defined diameter was created. Subsequently, migratory efforts of SMCs invading the wound regions and to close the wound were tracked periodically by videomicroscopy over time. Then, knockout cell lines were compared to wild type controls as shown in Figure 18, using 1way ANOVA with Dunnett correction for multiple comparisons after testing for normal distribution using D’Agostino-Pearson test.

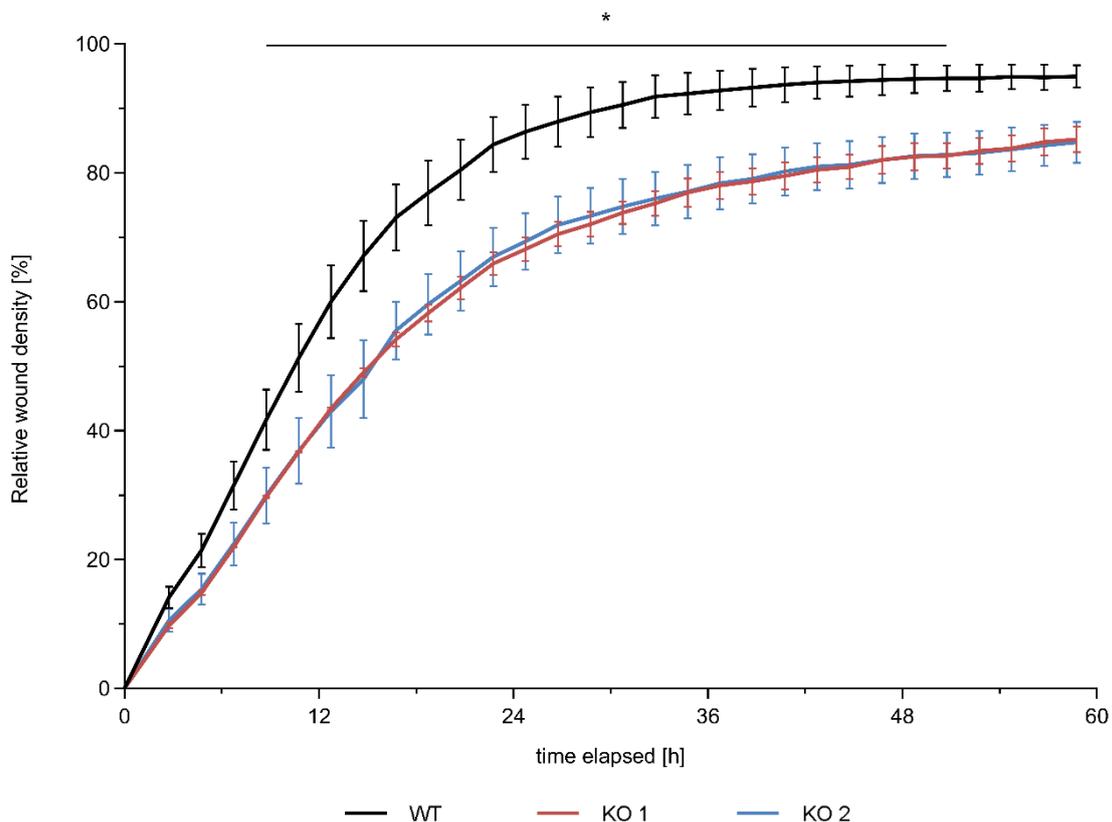


Figure 18 | Functional studies exploring migration phenotype of iPSC-derived SMCs. Measuring the number of SMCs migrating into cell-free area during a scratch-wound assay revealed reduced migration of KO 1 and KO 2 SMCs compared to wild type control SMCs. First significant differences can be observed 8.75 hours after wounding for both KO 1 and KO 2. Closure of the created wound (attained at 100% cellular density in the former wound) was also delayed in both KO 1 and KO 2. While the wound of wild type controls was already closed 30.75 hours after wound creation, both knockout cell lines did not close completely during the observed time frame. The time series shows means and \pm 95% confidence interval from 3-4 experimental repetitions and quadruplicate measurements at each time point. For plotting, values for each replicate were normalized to the cell density outside of the wound area at every time point for each replicate. * (adjusted $p \leq 0.05$). Dunnett’s multiple comparison testing in repeated-measured 1way ANOVA.

The relative wound density (RWD) showed a decreased cell density within the wound for both KO 1 and KO 2 (Figure 18). Here, both knockout cell lines acted approximately similarly. The statistically significant effect of decreased cell density was observable 8.75 hours after wound creation for KO 1 ($P = 0.04$) and KO 2 ($P = 0.03$) for KO 2. At this time, wild type controls showed a mean RWD of 41.7%, while the mean RWD was 29.7% for KO 1 and 29.9% for KO 2. The difference became even more pronounced, reaching its peak after 16.75 hours for KO 1 with a mean delta RWD of 18.9% ($P < 0.001$). For KO 2, the peak was reached after 14.75 hours with a mean RWD delta of 19.1% ($P < 0.001$).

In summary, knockout cell lines showed decreased migratory SMC efforts into created wound areas. As migratory efforts are part of atherogenesis, regarding the migration phenotype, knockouts seem to have atheroprotective properties. Of note, as proliferation has not been selectively blocked during the assay, a limitation of interpreting results from the scratch assay is the possible impact of altered proliferation of the cells during migration.⁹³ Thus, the assay readout is due to a compound effect of proliferation and/or migration, whereby the latter cannot be separated from the first. Therefore, it cannot be ruled out, that at least to some degree, the observed decreased wound density of knockout cell lines is partly because of the decreased proliferation in knockout cell lines, that we have described above.

5.5.2.4 Increased adhesion phenotype in ECs

Functional studies of adhesion characteristics in iPSC-derived ECs were performed, exploring whether the chromosome 21 knockouts showed defects in adhesive behaviour. An increase in EC adhesion is considered one of the first and main atherogenic behaviours.⁹⁴ To simulate the situation in nascent atherosclerotic lesions, where activated ECs recruit circulating immune cells of the blood, freshly isolated wild type PBMCs (Ethik-Kommissionsnummer 17-012, see methods) were pipetted onto knockout cell lines or wild type controls of ECs. After briefly allowing ECs to adhere, residual unattached PBMCs were washed off with medium and attached PBMCs were counted. Subsequently, variance between remaining adherent PBMCs in knockout cell lines and wild type controls were compared as shown in Figure 19, using multiple t-tests with Holm-Sidak correction for multiple comparisons after testing for normal distribution using D'Agostino-Pearson test.

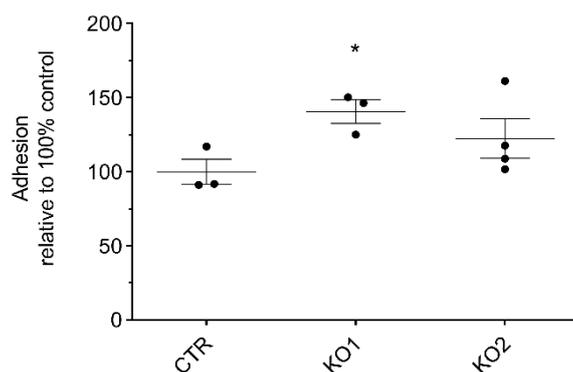


Figure 19 | Functional studies exploring adhesion phenotype of iPSC-derived ECs. Wild type PBMCs freshly isolated from blood of a healthy donor were pipetted onto confluent cell layers of wild type and knockout iPSC-derived endothelial cell lines in order to adhere. After removal of unattached PBMCs by rinsing, adhesion was measured by microscopically quantifying the number of PBMCs remaining attached on top of ECs. PBMCs showed increased adhesion to KO 1, compared to wild type controls. Data represents 3 replicates of the assay measured in quadruplicates each and normalized to adhesion in wild type controls (100%). Scatter dot plots show means with SEM. * (adjusted $p \leq 0.05$). Holm-Sidak multiple comparison testing in repeated-measured t-test after testing for normal distribution with D'Agostino and Pearson test.

In KO 1 cell line, adhesion of PBMCs to ECs was increased significantly when the chromosome 21 locus was knocked out in ECs. In detail, compared to wild type controls set to 100%, KO 1 showed augmented adhesion with a mean of 140.6% ($P = 0.02$). Meanwhile, adhesion of KO 2 cell line was not significantly increased at a mean of 122.4% ($P = 0.25$).

In conclusion, KO 1 EC cell lines showed an increased adhesion phenotype relative to PBMCs. Such a phenotype would conceptually enhance immigration possibilities for PBMCs and other immune cells in the blood, which are known to be crucial for creating and maintaining an atherosclerotic lesion. Thus, in regard to adhesion phenotype, knockout in ECs might be capable of contributing atherogenic properties.

6

Discussion

The thesis is built up in two main parts. In the first part, SNPs associated with atherosclerosis and its risk factors were assembled and used to define haplotype blocks associated with atherosclerosis. Out of all identified haplotype blocks, one top candidate locus was chosen using different selection criteria. These included checking for colocalization of haplotype blocks and risk SNPs, taking genetic content of the haplotype blocks into account, and considering effect sizes (large) and p-values (small) of according lead SNPs within the haplotype block. The haplotype block on chr21q22.11 became the top candidate locus for further downstream experimental analysis. With a total length of 130 307 bp, the relevant haplotype block contains two lead SNPs with robust disease associations at rs9982601 (p-value 1.33×10^{-13}) and rs28451064 (p-value 1.33×10^{-15}). Furthermore, a haplotype subblock structure with a length of 58 817 bp, containing both lead SNPs as well was found, thus begging the question of whether the entire region or the smaller subregion were relevant for vascular biology. The locus does not contain any protein-coding genes, but two lncRNAs. However, there are some protein-coding genes surrounding the haplotype block, mentioned in previous literature referring to chr21q22.11 and as characterized below. The locus was followed-up experimentally by knockouts of the actual haplotype block (KO 1) as well as the smaller subblock (KO 2), were generated in hiPSCs. Subsequently, knockout and wild type hiPSCs were differentiated into cell lines relevant for atherosclerosis (SMCs and ECs), allowing gene expression and effects on cell function to be compared in KO and wild type cell lines.

The main finding of the thesis was an enhanced expression of neighbouring genes, both 5' and 3' of the knockouts, when regions encompassing parts of the lncRNA-encoding core locus were deleted. The presence of an intact locus in the wild type, therefore, seems to repress (to a certain extent) the expression of neighbouring genes, an effect that is most prominently seen in SMCs. As of now, and in the absence of additional experiments, we do not know whether the effect is due to the presence of a DNA element in the non-coding region, or to the function of the lncRNA expressed from this DNA element, or both. Collectively, our data give rise to the hypothesis that the atherosclerosis-causing mutation, as inferred from our data on the deletion, deactivates a repressive element, leading to enhanced proatherogenic gene expression. With the effect being detectable in both KO 1 and KO2, it is indicative for the responsible effector to reside in the smaller subblock of KO 2 (Figure 12), which would also be consistent with the peak of genetic association in this region as compared to KO 1 (Figure 9).

6.1 Possible explanations for observed gene expression alteration

As both knockouts include relatively large genomic regions, the exact molecular mechanism behind the observed gene expression enhancement remains unclear, as is the question of whether the experimental deletions reflects effects that would be caused in human patients by one or few single nucleotide polymorphisms in an otherwise unchanged topology of the locus. However, there are certain reasonable possibilities as to how genetic variation in the locus, even if small, could affect gene expression in this way. First, it could be due to altered chromatin state. It is conceivable that through loss of a structural element essential for maintaining an intact 3D heterochromatin unit, gene expression in and around the locus is increased.⁹⁵⁻⁹⁷ Second, it could be due to a change in DNA sequence: Through loss of transcription factor binding sites, repressors could be enhanced, or activators/enhancers could directly lose their function and gene expression could be correspondingly increased on a broader scale.⁹⁸⁻¹⁰¹ Finally, the reason for enhanced gene expression could be at RNA level: Due to altered function or expression of lncRNAs, these might influence gene expression in *cis*, and, though it was not tested here yet, possibly also in *trans*.^{46-49,52} Our RNA data suggest at least a local effect, because of increased expression of genes surrounding the knockout. Moreover, multiple genes are affected unidirectionally (all up), implying that there are either multiple effectors, or (potentially more likely) one effector with multiple targets. The latter scenario is similar to the possibility that the expression in the whole larger region is linked and regulated collectively. In addition, the SNPs probably have cell-specific effects, as the detected gene expression levels differ already in the investigated SMCs and ECs. However, currently, and without further smaller deletions, specifically without directly engineering the risk SNPs into an otherwise wild type locus, we cannot be clearer what the exact molecular function of the deleted region and SNPs are and which DNA motifs, if any, are relevant to protect from disease in the wild type.

6.2 Interpretation of altered cellular functions

On the cellular level, we observed increased baseline apoptosis, as well as a reduced execution of apoptosis when high apoptotic stress was induced. Effects were seen in both SMCs (KO 1 and KO 2) and ECs (KO 2), which is seemingly contradictory at first sight. There are, however, ambiguous views on the net effect of apoptosis in atherosclerosis, depending on the exact stage of disease and cell type. First, for ECs in the initiating phase of atherosclerosis development, studies discovered increased turnover and apoptosis, reinforcing an inflammatory response in atherogenesis.¹⁰² Furthermore, apoptosis of ECs in late stages has adverse effects on plaque stability.^{102,103} Transferred to our findings, the increased baseline apoptosis observed in ECs (KO 2) are a sign for atherogenic behaviour. Nevertheless, the reduced vulnerability to induced and ongoing apoptosis would be an anti-atherogenic

feature and, thus, does not allow for an unequivocal statement.¹⁰²⁻¹⁰⁵ Second, apoptosis of SMCs within plaques comes with multiple consequences: The decreased output of structural proteins like collagen, crucial for stabilizing the fibrous cap, makes the plaque more prone to rupture^{87,102,106-108}, and loss of SMCs promotes calcification, enlargement of the necrotic core and maintains the inflammatory response^{109,110}. Also, studies showed SMCs from atherosclerotic lesions to be intrinsically more sensitive to apoptosis than SMCs from normal vessels.^{111,112} Therefore, our finding of increased baseline apoptosis could be interpreted as atherogenic. However, on the other hand, and as above for ECs, one could also argue that the reduced sensitivity to induced apoptosis could be an atheroprotective feature. Lastly, there are also ambiguous views on apoptosis of macrophages, which we did not examine.¹¹³ For early stages, apoptosis of macrophages has been shown to be atheroprotective, because their removal from the lesion attenuated the inflammatory response.¹¹³⁻¹¹⁵ On the other hand, a loss of macrophages comes in hand with reduced clearance of other apoptotic cell types or debris.^{113,116} In order to prevent the inflammatory consequences associated with the accumulation of apoptotic cells and debris, these materials are rapidly and efficiently cleared by macrophages in a phagocytic process known as programmed cell removal or 'efferocytosis'.¹¹⁶ Therefore, reduced efferocytosis based on increased macrophage apoptosis leads to accumulation of secondary necrotic cells, which result in formation of a highly inflammatory necrotic core, releasing toxic components, thus further promoting inflammatory response.¹¹⁶⁻¹¹⁸ These effects are mainly observed in advanced stage atherosclerotic lesions. In summary, despite many efforts in determining the potential mechanisms of apoptosis, the significance of apoptotic cell death in atherosclerosis in distinct cell types remains not fully clear. Therefore, it would be too much of a simplification to claim that increases in apoptosis are (by definition) atherogenic or atheroprotective.

As another cellular function, we observed decreased levels of proliferation in both SMCs and ECs in KO 1. There are different views on proliferation in atherosclerosis, depending on the cell type and stage of lesion. First, early stages of atherosclerosis show increased proliferation of SMCs, which build up and maintain the lesion and fibrous cap by production of extracellular matrix and inflammatory cytokines.^{89,119} Also, studies showed that the restenosis rate of stented or bypassed vessels is linked to increased proliferation of SMCs.⁹⁰ In advanced lesions, on the contrary, reduced proliferation rates are observed in SMCs, resulting in thinning of the fibrous cap and danger of rupture.¹²⁰ Here, proliferation would be beneficial for plaque stability. Altogether, the reduced proliferation rate of SMCs observed in both knockouts, can be atheroprotective in early stages but negative for disease outcome in later stages of atherogenesis. Second, increased proliferation of ECs, especially in early-stage lesions, helps limit atherosclerosis through replacement of defect ECs at predilection sites.^{121,122} In reference to the decreased proliferation observed in our experiments, it could be atherogenic especially in early stages of lesion formation, as EC replenishment is diminished. However, the consequences of proliferation of distinct cell types in atherosclerotic lesions needs to be further uncovered in future studies.

There are again different hypotheses of SMC migration in atherogenesis. First, the inflammatory environment of early atherosclerotic lesions promotes conversion of quiescent contractile SMCs into active synthetic SMCs capable of migrating and proliferating.^{123,124} Subsequent migration from the

media into the intima with local proliferation and production of extracellular matrix is a crucial step in atherogenesis.¹²⁵ Therefore, the reduced migration of SMCs observed in both knockouts may be interpreted as atheroprotective. Indeed, there are studies showing prevented lesion development after reduction of SMC migration.¹²⁶ Then again, plaque stability is dependent on SMCs in the fibrous cap, producing extracellular matrix proteins.^{127,128} If the equilibrium of SMC apoptosis and proliferation is shifted towards apoptosis, new SMCs need to migrate into the lesion to ensure plaque stability.¹²⁹ Especially fibrous caps with a reduced number of SMCs in the border are prone to rupture.^{91,123,124,129} Therefore, reduced migration of SMCs observed in both knockouts may be atherogenic in late stages of atherosclerosis.

In early stages of atherosclerosis, vascular stress and an inflammatory environment lead to increased production of cytokines and expression of adhesion molecules in ECs. The increase in adhesiveness and cytokines results in capturing of monocytes and T lymphocytes, maintaining, and promoting atherosclerotic development.^{94,130,131} Therefore, the ability of adhesion is a vital step in early lesion formation. Thus, the increased adhesion of PBMCs to ECs observed in KO 1 may be atherogenic in early stage lesions. It should be stated, however, that we measured here the adhesion of wild type PBMCs to mutated ECs, and the behaviour of mutant PBMCs as occurring in a human individual carrying the risk genotype in all (or at least the relevant) somatic cells in the body cannot yet be predicted on our current data alone.

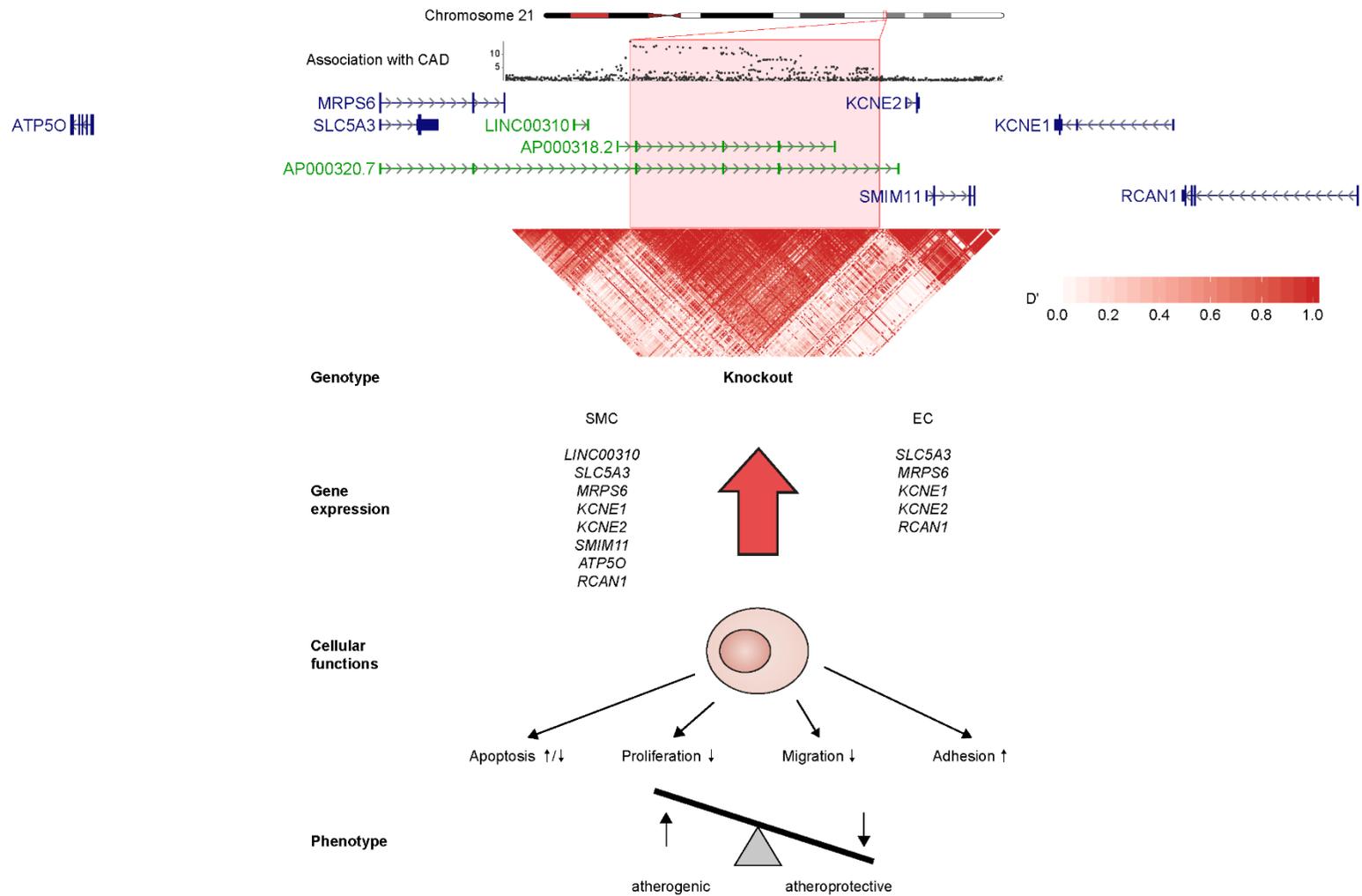


Figure 20 | Synopsis and impact of chromosome 21 locus in atherosclerosis . The upper part gives the details of the atherosclerosis susceptibility locus on chromosome 21. The association with atherosclerosis is mapped by the scatter blot and confirms the structure of the actual haplotype block below, shown in D'. Only non-coding RNAs (green) reside within the haplotype block, which is surrounded by other protein-coding genes (blue). Shown are only the genes examined in the thesis, extracted from Gencode V19. The effects of the knockout genotypes are shown below. In general, the knockout caused a general rise in expression of genes surrounding the locus. The cellular functional assays showed changes in an altered apoptosis, decreased proliferation and migration as well as an increased adhesion. All in all, representing an atherogenic phenotype, or at least a loss of atheroprotective features.

In a synopsis of all experiments on cellular function, we conclude that although the knockouts showed atherogenic behaviour, depending on the stage of atherosclerosis and cell type, atheroprotective features could be thinkable as well (Figure 20). Therefore, additional experiments targeted at mimicking the different stages and contexts of atherogenesis are necessary to further evaluate the net effect of the risk of the locus. These future experiments are needed to examine effects of SNPs/loci in separate cell types, as we observed cell type dependent outcomes. Also, *in vivo* experiments may be useful to further validate and classify the observed atheroprotective/atherogenic phenotype of the knockout. In addition, we cannot give a definitive statement on the effects of actual risk genotype (SNP), as we produced knockouts in genetic wild type cell lines, meaning the cells did not carry the relevant SNPs, and we also investigated heterozygous deletions but not homozygous mutant conditions.

6.3 Background and relevance of examined genes

There is little research of the genes surrounding the locus, so far. In fact, there are no studies indicating an association between *ATP5O*, *MRPS6*, *SLC5A3* and *SMIM11* expression in the vessel wall and atherosclerosis. The SNP rs28451064 on chromosome 21 locus is associated with increased expression of *MRPS6* and *SLC5A3* in blood and vascular tissue, but how atherosclerosis-relevant molecular and cellular transitions would be affected remained unaddressed in this study.¹³² *MRPS6*, as mitochondrial ribosomal protein, is involved in the generation of reactive oxygen species.¹³² These are molecules associated with proatherogenic properties.¹³³ The role of increased *SLC5A3* expression on atherogenesis remains unclear. Eventually, its increased expression is potentially due to both *MRPS6* and *SLC5A3* sharing an exon, which is the ORF for *MRPS6*. In concordance with our experiments resulting in both increased *MRPS6* and *SLC5A3* levels, the created knockout seems to provoke atherogenesis.

It was shown that expression of one neighbouring lncRNA of chromosome 21 risk locus, *LINC00310*, was increased in patients with breast cancer.¹³⁴ Here, *LINC00310* promoted cell proliferation in two ways: first by regulation of *c-Myc*, a well-known proto-oncogene and second by regulation *CDK4*, an important factor of cell cycle control.¹³⁴ Therefore, the increased expression of *LINC00310* in our experiments and the concomitant decreased proliferation in SMCs are not in line with these other findings in different cell types. However, effects might be specific for the vascular cell types, and as our knockout affected expression of multiple genes and we limited our research to *in vitro* experiments, these earlier studies are also not contradictory to our finding. Additional studies are required to further shed light on *LINC00310* function in atherosclerosis. The other lncRNA in the locus, *AP000318.2*, has only recently been annotated and has not at all been studied before.

More is known about the *KCNE* gene family: *KCNE1-5* are coding structural elements for potassium voltage-gated channels. The SNP rs9982601 on chromosome 21 locus is associated with increased expression of *KCNE2*.¹³² In a study of 2010, both *KCNE1* and *KCNE2* were shown to be elevated 1-2 weeks after myocardial infarction in mice.¹³⁵ With most arrhythmic events happening

during that time period, authors concluded a connection between *KCNE1/KCNE2* and arrhythmogenesis. Also, *KCNE* genes have previously been associated with Long QT syndrome.^{136,137} More interestingly, a recent study showed that deletion of *KCNE2* promotes atherosclerosis in mice.¹³⁸ Although the underlying mechanism is still unclear, *KCNE2* deletion led to increase in serum LDL and impaired glucose tolerance. In this context, our findings of increased levels of *KCNE2* in knockouts imply an atherogenic role of the knockouts. Though, another more recent study published the cardioprotective quality of *KCNE2* deletion following the reperfusion period after myocardial infarction in mice.¹³⁹ Therefore, *KCNE2* seems to have context dependent effects in cardiovascular biology but is a good candidate when trying to explore how chromosome 21 risk is executed. Additional studies are needed to define the exact molecular mechanisms.

For *RCANI*, an inhibitor of calcineurin, studies in mice found an increased expression within atherosclerotic plaques, with reduced plaque size after knockout of *RCANI*.¹⁴⁰ Expression of *RCANI* in VSMC, macrophages, and ECs was induced by treatment with oxidized LDL particles and in turn, uptake of oxidized LDL was mediated by *RCANI*.¹⁴⁰ Furthermore, these particles inhibited migration of macrophages, thus trapping them in the lesion.¹⁴⁰ Lastly, the presence of *RCANI* has been shown to promote an inflammatory phenotype in macrophages.¹⁴⁰ In turn, its deletion pushed for an anti-inflammatory phenotype.¹⁴⁰ All in all, their data strongly suggest that *RCANI* promotes disease progression in atherosclerosis.^{140,141} Taken together, the effects in this study and in our data are congruent in their direction, and we conclude that the increased *RCANI* expression in our experiments could have an atherogenic influence. Nevertheless, the previous study had a strong focus on macrophages, therefore a direct translation would be too short-sighted and further experiments of the role of *RCANI*, especially in SMC and ECs are necessary.

6.4 Outlook into *in vivo* relevance

The relevance of our findings *in vitro* need to be assessed in *in vivo* studies and experiments. In particular, the question arises of whether expression of the investigated protein-coding genes neighbouring the lncRNA-containing chromosome 21 core region is indeed also increased in human atherosclerosis patients. Only this would document that the effects we observe in our knockouts are of potential relevance for the disease *in vivo*. One possible approach would be to perform an eQTL analysis using RNA isolated from atherosclerosis relevant tissue of a large atherosclerosis cohort. Thus, the genotypes regarding the discovered risk SNPs of the chromosome 21 locus could be linked to a corresponding gene expression pattern of the genes examined in this thesis. This may be a first useful step to validate the *in vitro* findings and further explore how atherosclerosis risk is affected in carriers of chromosome 21 risk variations.

In the future, not only genes in the neighbourhood, but all genes elsewhere (and potentially regulated *in trans*) will have to be explored *in vivo* and in the created knockout iPSC model from this

thesis. To summarize, it is thus possible that *in vivo*, just as in our knockout models, gene expression in neighbouring genes increases eventually via disturbance of a repressive element of the locus. The exact molecular mechanism behind the de-repression is beyond the scope of the presented thesis. The main advance of this thesis (as compared to earlier work) is that we achieved the establishment of a cellular model that allows the functional analysis of a disease risk locus via knockouts based on the initial hypothesis-free selection of a GWAS hit. Furthermore, advancing earlier studies that had implicated one gene as effector via an eQTL in blood that remained functionally unclear (*MRPS6*¹⁴²), we functionally studied effectors candidates. We did so in cell lines like endothelial and smooth muscle cells, thus pertaining to cells that participate directly in the growth of atherosclerotic lesions and thereby going further than the initial analysis in whole blood.

6.5 Implications for future studies and clinical applications

The results of this thesis provide the bases for future experiments to further unravel the pathophysiology and exact molecular mechanisms behind the atherosclerosis region on chromosome 21. First, a detailed analysis of *cis*- as well as *trans*-eQTLs for both risk SNPs could yield possible candidate genes for further experimental analysis. Especially the relationship of long distant genes with *trans*-eQTLs were not covered in our work. Furthermore, the potential effects of the SNPs on transcription factor binding sites and other regulatory elements via bioinformatical prediction tools could provide explanations for the increased gene expression levels, observed in our experiments.

Future wet lab experiments may be suitable to narrow down the functional genetic segment by downsizing the respective knockout sites. As our experiments yielded comparable results for both KO 1 and KO 2, the functionally important part seems to concentrate within the genetic sequence of KO 2. One should remain careful though, as our data do not rule out that more than a single functional element is contained in the locus, and that KO 2 does not phenocopy KO 1 in all assays. The generation of two knockouts within KO 2 which separate both SNPs associated with atherosclerosis (rs9982601 and rs28451064) could be one possible approach. Another approach may be to directly and solely promote single base replacement of these SNPs, thus studying only the effect of the single SNP as they occur in patients *in vivo*, which might or might not lead to different consequences for cells. In our laboratory, we reached the limit of CRISPR/Cas9-dependent locus targeting available at this time, as the respective locus is a highly repetitive genetic site and difficult to edit. Future evolved or orthologous Cas enzymes will certainly allow additional deletions and modifications to be implemented. However, rescue experiments of specific knockouts with engineered transgenes could be another way of further analysing the locus. Also, knockdown experiments or functional experiments of lncRNAs themselves via RNAi should facilitate the understanding of molecular pathways of chromosome 21 locus.

The complex role of lncRNAs in various stages of gene expression as explained earlier, gives rise to the view that noncoding RNA is involved in many diseases. In fact, multiple SNPs causing alterations in structure and expression levels of lncRNAs, have been identified and have been linked to cancer, neurological disorders, cardiovascular diseases, and others.¹⁴³⁻¹⁴⁸ In cancer, lncRNAs were found to have oncogenic, tumour suppressive or bidirectional functions, mainly via dysregulation of genes involved in differentiation, proliferation, apoptosis, and cell cycle regulation. *HOTAIR*, *MALATI*, *ANRIL* and *PANDA* are just some examples of oncogenic lncRNAs involved in hepatocellular, colorectal, lung and breast cancer.¹⁴⁹⁻¹⁵⁹ Here, for example, overexpression of lncRNA *PANDA* immortalizes cancer cells in breast cancer via modulation of p53 pathway.¹⁵⁷⁻¹⁵⁹ Many lncRNAs were also found to be dysregulated in nervous system disorders like Alzheimer's disease (*BACE1-AS*)^{160,161}, Huntington's disease (*HTT-AS*)^{162,163}, Parkinson's disease (*UCHL1-AS*)¹⁶⁴, amyotrophic lateral sclerosis (*NEATI_2*)^{165,166} or multiple sclerosis (*Tmevpgl*)¹⁶³. Some of the most prominent examples of lncRNAs in cardiovascular diseases are probably *ANRIL*, *MIAT*, *MALATI* and *GAS5*.^{36-38,66,167-170} The role of *ANRIL* in atherosclerosis has already been presented above. SNPs located within the sequence of *MIAT* were found to be a risk factor for myocardial infarction.¹⁶⁸ The lncRNA *GAS5*, mainly expressed in ECs and VSMCs, regulates hypertension-induced vascular remodelling and is downregulated in hypertension.⁶⁶ Beyond causal roles in disease process, and the fact that lncRNAs could serve as therapeutic targets, their selective expression in disease states also makes lncRNAs interesting candidates as potential biomarkers yielding diagnostic and prognostic clinical information.^{171,172}

The better understanding of their role in pathophysiology and technological advances allowed RNAs (coding and non-coding) to be considered for therapeutic approaches. The two lncRNAs in the chromosome 21 locus might also be functional relevant, but more specific work (like RNAi) is necessary to address this possibility. The general goal is to alter gene expression of disease-relevant proteins via manipulating lncRNAs, thereby making a positive impact. This could be achieved by partially or completely turning off gene expression using siRNAs, ASOs or miRNA as RNA drugs.¹⁷³⁻¹⁷⁵

Despite the various applications and progress in the field, there are still major challenges and hurdles to overcome. The fundamental problems faced, are the instability of RNA as drugs, its vulnerability against nuclease degradation, the immunogenicity of RNAs and the problem of achieving selective cellular uptake into the desired cells and tissues. Yet, chemical modifications of RNA drugs and the usage of artificial carriers have facilitated and improved RNA therapy already.^{173,174}

The applications mentioned above are also used to alter expression levels of deregulated lncRNAs involved in pathological processes of diseases. The two main fields of lncRNA therapies, so far, are cancer and rare genetic disorders. In this context, as well as for atherosclerosis, the lncRNA *MALATI* was tested as druggable target, showing the feasibility of the approach.^{176,177} However, the inconsistency of lncRNA conservation throughout different species makes it difficult for researchers to develop and study pharmacological effects and side effects based on animal models. The iPSC system, as we use here, is thus an interesting model, but in the best case needs complementary testing in *in vivo* context.^{178,179}

Appendices



Supplementary Tables

Supplementary Table 1 | Oligonucleotides for genotyping of SNP rs9982601 by melt curve

Primer	Sequence in 5'-3'	Annealing temperature [°C]
Forward	CTGCAACAGACGCAGTATCTCTTGGT	60
Reverse	TTAAAACAAGCACAGGGCCCATCT	60
Probe	*CTGCTCCATGGCCTTG [#]	50

* FAM

[#] TAMRA

Supplementary Table 2 | Oligonucleotides for PCR of knockout positive clones

Detected Knockout	Primer	Sequence in 5'-3'	Annealing temperature[°C]
1	Forward	CTAATTCCCCAGCTACCCAGTC	62
	Reverse	CTCTTCCCCTAGCCCCGAAG	62
2	Forward	GCTTGAGTGGCTTTGGTATTCC	60
	Reverse	GACAGCCACATAACATGCACC	60

Supplementary Table 3 | Primer and Probes for qRT-PCR

Gene	Primer and Probe	Sequence in 5'-3'
<i>SOX2</i>	Forward	CCAGCGCATGGACAGTTACG
	Reverse	CTGGTCCTGCATCATGCTGTAG
	Probe	*TGAACGGCTGGAGCAACGGCA [#]
<i>ISL1</i>	Forward	TCATGATGAAGCAACTCCAGCA
	Reverse	TGGACTGGCAGCCACCAT
	Probe	*AGCAGCCCAATGACAAAATAATATCCAGGG [#]
<i>NKX2.5</i>	Forward	ACCCAGCCAAGGACCCTAGA
	Reverse	CAGCTCCACCGCCTTCTG
	Probe	*CGAAAAGAAAGAGCTGTGCGCGC [#]
<i>TBX18</i>	Forward	CATTTGCTAAAGGCTTCCGAGA
	Reverse	TATGATTCCACCAAGGCTTCCA
	Probe	*TCCGGGCGCAACAGAATGGGT [#]
<i>ICAM1</i>	Forward	CCTTCCTCACCGTGTACTGG
	Reverse	AGCGTAGGGTAAGGTTCTTGC
	Probe	*CAGAACGGGTGGAAGTGGCACCCTC [#]

<i>VEGF</i>	Forward	TCCTGCCCCGGCTCACC
	Reverse	GAGGCAGCTTGAGTTAAACAGAC
	Probe	*AAACACAGACTCGCGTTGCAA [#]
<i>CD144</i>	Forward	GCAGTGGTAGCCATCTTACTCTGC
	Reverse	GATCTCCGGCACGCTCTTG
	Probe	*CTCACCATCACAGTGATCACCTGCTCATC [#]
<i>CNN1</i>	Forward	TCCACCCTCCTGGCTTTG
	Reverse	TCACTCCCACGTTACCTTGT
	Probe	*CCAGCATGGCGAAGACGAAAGGA [#]
<i>MYH11</i>	Forward	GCGAGGTGAACGCACTCAAG
	Reverse	TTCAATAACTCTACGTCCTCCAGACC
	Probe	*CTCAGGCGAGGAAACGAGACCTCTTTTCG [#]
<i>α-SMA</i>	Forward	GACCGAATGCAGAAGGAGATCAC
	Reverse	GAGTATTTGCGCTCCGGAGG
	Probe	*CACCCAGCACCATGAAGATCAAGATCATTG [#]
<i>LINC00310 Ex1-2</i>	Forward	CTTTGGATTCTTTGTGAACAAATGAG
	Reverse	GAAGAAGGTGGTGACCTGCTCC
	Probe	*CAAGGCTGCAAATTCAGAGGATCCTCCCAG [#]
<i>LINC00310 Ex1-3</i>	Forward	CTTTGGATTCTTTGTGAACAAATGAG
	Reverse	GTCCTGCTGCAGCACGTGG
	Probe	*CAAGCTGCCAGAGGGATGCTAACACCG [#]
<i>LINC00310 Ex1-3b</i>	Forward	CTTTGGATTCTTTGTGAACAAATGAG
	Reverse	GTACTTTTCAATGGTTCCATTGCC
	Probe	*CAAGCCCCTTTTGTCTAAAATAGTGATGTACATTAAATG [#]
<i>LINC00310 ExCD-3b</i>	Forward	CTAGCAGTTCTGATGTCTCTCCAAGG
	Reverse	GCCCTTTGCATTTAATGTACATCAC
	Probe	*CCCACAGATCCTGGTCTACTAACTTAAGGCCCC [#]
<i>LINC00310 Ex1-5</i>	Forward	GATCGATGTTGCTGGCTTTGG
	Reverse	GATGAGAAGTCACAAGGAAGCTTGC
	Probe	*CTTTGTGAACAAATGAGACAAGTTGCAAGTGGATG [#]
<i>AP000318.2 AP000320.7 Ex4-5</i>	Forward	CAGCTCTGCAGGAGCCAGG
	Reverse	CAGGATGAGAAGTCACAAGGAAGC
	Probe	*CTGCTCTGGGTTGCAAGTGGATGGC [#]
<i>AP000318.2 AP000320.7 Ex5-6</i>	Forward	GGATGGCAAGCTTCCTTGTGAC
	Reverse	GGAAGACTTAACAGTCACTGTCGCC
	Probe	*CTCATCCTGGCAATTTTCTGGAGTGTTTGTG [#]
<i>AP000318.2 AP000320.7 Ex6-7</i>	Forward	GACGAGGGTTCCTGGGAGC
	Reverse	GATGTTGAGGATTCAATGATGAATTTG
	Probe	*CCAGGGCTGCAAGCATCAGGACAGC [#]
<i>AP000318.2 AP000320.7</i>	Forward	CTGGCATAAGGTGAAGAGAAGAAGAGTA
	Reverse	GAGTAGCTGAAGGTCATTGTATTTGATTG

<i>Ex7-8</i>	Probe	*TGAACTACCACGAAACTCATTTCCCTCCGTC [#]
<i>AP000318.2</i> <i>AP000320.7</i> <i>Ex7-9</i>	Forward	CAGGACAGCAGAGACCAAATTCATC
	Reverse	CAGCAAAAATATAGTCTGTACACCTTGATATATC
	Probe	*TGAATCCTCAACATCTAGCACAAATATGCCTGGC [#]
<i>AP000320.7</i> <i>Ex9-10</i>	Forward	GTAATGGAACCCAGTCTACATAAAGCTG
	Reverse	GTGAGCAAATTACTTACTGTCCCTGG
	Probe	*TGACCCAGCACATGACGAAGAATCCAAAG [#]
<i>SLC5A3</i>	Forward	GCTGGCTTTAATCCTGAAAGCC
	Reverse	GGAGGAGAAATTCTGAGTTGCAGC
	Probe	*CAGCGGGGTTGGTACAGTAGGCTTCACTAG [#]
<i>MRPS6</i>	Forward	GGCTTTAATCCTGAAAGCCATGC
	Reverse	GCTCCTCTGTCCATCAGGGC
	Probe	*CGGCCAGAGACTGCTGCTACTTTGAAACG [#]
<i>KCNE1</i>	Forward	CTGTATCCAGAGGAAATAGCCAAGG
	Reverse	GCTGTGGTGTAGACAGGATCATCC
	Probe	*CAGAGGTGTGCCTGGGAAGTTTGAGCTG [#]
<i>KCNE2</i>	Forward	CTTTGAAAGGAATTCATCCTGC
	Reverse	CGGAAGACGTCTTCCAGCG
	Probe	*CCACACACTGCATAGCAGGAGGGAAGC [#]
<i>SMIM11</i>	Forward	GCTGGAGCCTCCGCTGC
	Reverse	GAATTCACCTCCTGATCAACACAAGC
	Probe	*CAATCAGACCTTCCAGCTGCCTCTCATGTACTTG [#]
<i>ATP50</i>	Forward	CTCTCTCCCACTCGGGTTT
	Reverse	TGACCACAGAGGTAAGCA
	Probe	*AGTGTCGGGCTCTCCCGGCAGGT [#]
<i>RCAN1</i>	Forward	CCTGTGTGGCAAACAGTGATATC
	Reverse	CATACGTCCTAAAGAGGGACTCAA
	Probe	*TCAGCGAAAGTGAAACCAGGGCCAA [#]

* FAM

TAMRA

Supplementary Table 4 | Oligonucleotides for PCR of transfected *E. coli*

Primer	Sequence in 5'-3'	Annealing temperature [°C]
M13 Forward	GTAAAACGACGGCCAG	60
M13 Reverse	CAGGAAACAGCTATGAC	60

Supplementary Table 5 | Oligonucleotides for Surveyor Assay

Analysed gRNA	Primer	Sequence in 5'-3'	Annealing temperature [°C]
5'-gRNA for Knockout 1 / 2	Forward	ATGCTGAGATGAACACTGTGTTTGG	62
	Reverse	TCAGTAGTTCGAGACCAGCCTGACC	62
3'-gRNA for Knockout 1	Forward	AAGCAGGATCAGACTATGTGAGAAGATG	62
	Reverse	TTGATTTGACGCTCCAGGCTG	62
3'-gRNA for Knockout 2	Forward	TCATGGCTTTGAGCCAGGGTAG	62
	Reverse	GCCACATAACATGCACCCAGTG	62

Supplementary Table 6 | Expected sizes of undigested PCR product and digested fragments

Tested gRNA	Size of PCR product	
	undigested	digested
5'-gRNA for Knockout 1 / 2	644 bp	246 bp + 398 bp
3'-gRNA for Knockout 1	625 bp	381 bp + 244 bp
3'-gRNA for Knockout 2	290 bp	221 bp + 69 bp

Supplementary Table 7 | Oligonucleotides for generation of protospacers for gRNAs of CRISPR/Cas9 system

gRNA	Oligonucleotide	Sequence in 5'-3'
5'-gRNA for Knockout 1 and 2	Oligonucleotide 1	CACCGTCCGCATGGGCACGACCAAC
	Oligonucleotide 2	AAACGTTGGTCGTGCCCATGCGGAC
3'-gRNA for Knockout 1	Oligonucleotide 1	CACCGAGCGGCATGATCCCGGTTCA
	Oligonucleotide 2	AAACTGAACCGGGATCATGCCGCTC
3'-gRNA for Knockout 2	Oligonucleotide 1	CACCGCTCATGAGGAATACGGCGGT
	Oligonucleotide 2	AAACACCGCCGTATTCCTCATGAGC

Supplementary Table 8 | Default traits for query in GWAS Catalog

Phenotype	Trait
ART	Cardiovascular disease risk factors
	Cardiovascular heart disease in diabetics
	Carotid intima media thickness
	Coronary artery calcification
	Coronary artery calcification (smoking interaction)
	Coronary artery disease
	Coronary artery disease or ischemic stroke
	Coronary artery disease or large artery stroke
	Coronary artery disease-related phenotypes
	Coronary heart disease
	Coronary heart disease in familial hypercholesterolemia
	Coronary restenosis
	Major CVD
	Myocardial infarction

	Myocardial infarction (early onset)
	Myocardial infarction in coronary artery disease
	Stroke (ischemic)
	Subclinical atherosclerosis traits (other)
	Sudden cardiac arrest
Elevated lipid levels	Apolipoprotein Levels
	Cholesterol
	Cholesterol and Triglycerides
	HDL cholesterol
	HDL cholesterol - Triglycerides (HDL-C-TG)
	HDL cholesterol (interaction)
	Hypertriglyceridemia
	LDL (oxidized)
	LDL cholesterol
	LDL cholesterol subfractions
	Lipid metabolism phenotypes
	Lipid traits
	Lipoprotein (a) - cholesterol levels
	Lp (a) levels
	Phytosterol levels
	Triglycerides
High blood pressure	Blood pressure
	Blood pressure (age interaction)
	Blood pressure (smoking interaction)
	Diastolic blood pressure
	Diastolic blood pressure (alcohol consumption interaction)
	Hypertension
	Hypertension (young onset)
	Pulse pressure
	Systolic blood pressure
Diabetes mellitus	Diabetes (incident)
	Diabetes related insulin traits
	Fasting glucose-related traits
	Fasting glucose-related traits (interaction with BMI)
	Fasting plasma glucose
	Type 1 diabetes
	Type 2 diabetes
	Type 2 diabetes (young onset) and obesity
	Type 2 diabetes and 6 quantitative traits
	Type 2 diabetes and other traits
Smoking	Nicotine dependence

	Nicotine use
	Smoking cessation
	Smoking initiation
	Smoking quantity

Supplementary Table 9 | Assessed genome-wide association studies

1	Aoki, A. et al. SNPs on chromosome 5p15.3 associated with myocardial infarction in Japanese population. <i>J Hum Genet</i> 56, 47-51, doi:10.1038/jhg.2010.141 (2011).
2	Aouizerat, B. E. et al. GWAS for discovery and replication of genetic loci associated with sudden cardiac arrest in patients with coronary artery disease. <i>BMC Cardiovasc Disord</i> 11, 29, doi:10.1186/1471-2261-11-29 (2011).
3	Arking, D. E. et al. Identification of a sudden cardiac death susceptibility locus at 2q24.2 through genome-wide association in European ancestry individuals. <i>PLoS Genet</i> 7, e1002158, doi:10.1371/journal.pgen.1002158 (2011).
4	Aulchenko, Y. S. et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. <i>Nat Genet</i> 41, 47-55, doi:10.1038/ng.269 (2009).
5	Barrett, J. C. et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. <i>Nat Genet</i> 41, 703-707, doi:10.1038/ng.381 (2009).
6	Bis, J. C. et al. Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. <i>Nat Genet</i> 43, 940-947, doi:10.1038/ng.920 (2011).
7	Bouatia-Naji, N. et al. A variant near MTNR1B is associated with increased fasting plasma glucose levels and type 2 diabetes risk. <i>Nat Genet</i> 41, 89-94, doi:10.1038/ng.277 (2009).
8	Bradfield, J. P. et al. A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. <i>PLoS Genet</i> 7, e1002293, doi:10.1371/journal.pgen.1002293 (2011).
9	Burkhardt, R. et al. Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. <i>Arterioscler Thromb Vasc Biol</i> 28, 2078-2084, doi:10.1161/ATVBAHA.108.172288 (2008).
10	Chambers, J. C. et al. Common genetic variation near melatonin receptor MTNR1B contributes to raised plasma glucose and increased risk of type 2 diabetes among Indian Asians and European Caucasians. <i>Diabetes</i> 58, 2703-2708, doi:10.2337/db08-1805 (2009).
11	Chasman, D. I. et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. <i>PLoS Genet</i> 5, e1000730, doi:10.1371/journal.pgen.1000730 (2009).
12	Cho, Y. S. et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. <i>Nat Genet</i> 44, 67-72, doi:10.1038/ng.1019 (2011).
13	Consortium, C. A. D. et al. Large-scale association analysis identifies new risk loci for coronary artery disease. <i>Nat Genet</i> 45, 25-33, doi:10.1038/ng.2480 (2013).
14	Consortium, S. T. D. et al. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. <i>Nature</i> 506, 97-101, doi:10.1038/nature12828 (2014).
15	Cooper, J. D. et al. Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. <i>Nat Genet</i> 40, 1399-1401, doi:10.1038/ng.249 (2008).
16	Coram, M. A. et al. Genome-wide characterization of shared and distinct genetic components that influence blood lipid levels in ethnically diverse human populations. <i>Am J Hum Genet</i> 92, 904-916, doi:10.1016/j.ajhg.2013.04.025 (2013).
17	Coronary Artery Disease Genetics, C. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. <i>Nat Genet</i> 43, 339-344, doi:10.1038/ng.782 (2011).
18	Cui, B. et al. A genome-wide association study confirms previously reported loci for type 2 diabetes in Han Chinese. <i>PLoS One</i> 6, e22353, doi:10.1371/journal.pone.0022353 (2011).

19	Davies, R. W. et al. A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. <i>Circ Cardiovasc Genet</i> 5, 217-225, doi:10.1161/CIRCGENETICS.111.961243 (2012).
20	Dichgans, M. et al. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. <i>Stroke</i> 45, 24-36, doi:10.1161/STROKEAHA.113.002707 (2014).
21	Dupuis, J. et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. <i>Nat Genet</i> 42, 105-116, doi:10.1038/ng.520 (2010).
22	Erdmann, J. et al. New susceptibility locus for coronary artery disease on chromosome 3q22.3. <i>Nat Genet</i> 41, 280-282, doi:10.1038/ng.307 (2009).
23	Erdmann, J. et al. Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. <i>Eur Heart J</i> 32, 158-168, doi:10.1093/eurheartj/ehq405 (2011).
24	Franceschini, N. et al. Genome-wide association analysis of blood-pressure traits in African-ancestry individuals reveals common associated genes in African and non-African populations. <i>Am J Hum Genet</i> 93, 545-554, doi:10.1016/j.ajhg.2013.07.010 (2013).
25	Grant, S. F. et al. Follow-up analysis of genome-wide association data identifies novel loci for type 1 diabetes. <i>Diabetes</i> 58, 290-295, doi:10.2337/db08-1022 (2009).
26	Hakonarson, H. et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. <i>Nature</i> 448, 591-594, doi:10.1038/nature06010 (2007).
27	Hakonarson, H. et al. A novel susceptibility locus for type 1 diabetes on Chr12q13 identified by a genome-wide association study. <i>Diabetes</i> 57, 1143-1146, doi:10.2337/db07-1305 (2008).
28	Hanson, R. L. et al. A genome-wide association study in American Indians implicates DNER as a susceptibility locus for type 2 diabetes. <i>Diabetes</i> 63, 369-376, doi:10.2337/db13-0416 (2014).
29	Hara, K. et al. Genome-wide association study identifies three novel loci for type 2 diabetes. <i>Hum Mol Genet</i> 23, 239-246, doi:10.1093/hmg/ddt399 (2014).
30	Heid, I. M. et al. Genome-wide association analysis of high-density lipoprotein cholesterol in the population-based KORA study sheds new light on intergenic regions. <i>Circ Cardiovasc Genet</i> 1, 10-20, doi:10.1161/CIRCGENETICS.108.776708 (2008).
31	Helgadottir, A. et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. <i>Science</i> 316, 1491-1493, doi:10.1126/science.1142842 (2007).
32	Hirokawa, M. et al. A genome-wide association study identifies PLCL2 and AP3D1-DOT1L-SF3A2 as new susceptibility loci for myocardial infarction in Japanese. <i>Eur J Hum Genet</i> 23, 374-380, doi:10.1038/ejhg.2014.110 (2015).
33	Howson, J. M. M. et al. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. <i>Nat Genet</i> 49, 1113-1119, doi:10.1038/ng.3874 (2017).
34	Huang, J., Ellinghaus, D., Franke, A., Howie, B. & Li, Y. 1000 Genomes-based imputation identifies novel and refined associations for the Wellcome Trust Case Control Consortium phase 1 Data. <i>Eur J Hum Genet</i> 20, 801-805, doi:10.1038/ejhg.2012.3 (2012).
35	Hwang, J. Y. et al. Genome-wide association meta-analysis identifies novel variants associated with fasting plasma glucose in East Asians. <i>Diabetes</i> 64, 291-298, doi:10.2337/db14-0563 (2015).
36	Imamura, M. et al. A single-nucleotide polymorphism in ANK1 is associated with susceptibility to type 2 diabetes in Japanese populations. <i>Hum Mol Genet</i> 21, 3042-3049, doi:10.1093/hmg/dds113 (2012).
37	International Consortium for Blood Pressure Genome-Wide Association, S. et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. <i>Nature</i> 478, 103-109, doi:10.1038/nature10405 (2011).
38	Kamatani, Y. et al. Genome-wide association study of hematological and biochemical traits in a Japanese population. <i>Nat Genet</i> 42, 210-215, doi:10.1038/ng.531 (2010).
39	Kathiresan, S. et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. <i>Nat Genet</i> 40, 189-197, doi:10.1038/ng.75 (2008).
40	Kathiresan, S. et al. Common variants at 30 loci contribute to polygenic dyslipidemia. <i>Nat Genet</i> 41, 56-65, doi:10.1038/ng.291 (2009).

41	Kato, N. et al. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. <i>Nat Genet</i> 43, 531-538, doi:10.1038/ng.834 (2011).
42	Kettunen, J. et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. <i>Nat Genet</i> 44, 269-276, doi:10.1038/ng.1073 (2012).
43	Kim, Y. J. et al. Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. <i>Nat Genet</i> 43, 990-995, doi:10.1038/ng.939 (2011).
44	Ko, A. et al. Amerindian-specific regions under positive selection harbour new lipid variants in Latinos. <i>Nat Commun</i> 5, 3983, doi:10.1038/ncomms4983 (2014).
45	Kooner, J. S. et al. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. <i>Nat Genet</i> 40, 149-151, doi:10.1038/ng.2007.61 (2008).
46	Kooner, J. S. et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. <i>Nat Genet</i> 43, 984-989, doi:10.1038/ng.921 (2011).
47	Kraja, A. T. et al. A bivariate genome-wide approach to metabolic syndrome: STAMPEED consortium. <i>Diabetes</i> 60, 1329-1339, doi:10.2337/db10-1011 (2011).
48	Lee, J. Y. et al. A genome-wide association study of a coronary artery disease risk variant. <i>J Hum Genet</i> 58, 120-126, doi:10.1038/jhg.2012.124 (2013).
49	Lettre, G. et al. Genome-wide association study of coronary heart disease and its risk factors in 8,090 African Americans: the NHLBI CARE Project. <i>PLoS Genet</i> 7, e1001300, doi:10.1371/journal.pgen.1001300 (2011).
50	Levy, D. et al. Genome-wide association study of blood pressure and hypertension. <i>Nat Genet</i> 41, 677-687, doi:10.1038/ng.384 (2009).
51	Levy, D. et al. Framingham Heart Study 100K Project: genome-wide associations for blood pressure and arterial stiffness. <i>BMC Med Genet</i> 8 Suppl 1, S3, doi:10.1186/1471-2350-8-S1-S3 (2007).
52	Li, H. et al. A genome-wide association study identifies GRK5 and RASGRP1 as type 2 diabetes loci in Chinese Hans. <i>Diabetes</i> 62, 291-298, doi:10.2337/db12-0454 (2013).
53	Lu, W. et al. Evidence for several independent genetic variants affecting lipoprotein (a) cholesterol levels. <i>Hum Mol Genet</i> 24, 2390-2400, doi:10.1093/hmg/ddu731 (2015).
54	Lu, X. et al. Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. <i>Nat Genet</i> 44, 890-894, doi:10.1038/ng.2337 (2012).
55	Lu, X. et al. Genome-wide association study in Chinese identifies novel loci for blood pressure and hypertension. <i>Hum Mol Genet</i> 24, 865-874, doi:10.1093/hmg/ddu478 (2015).
56	Ma, R. C. et al. Genome-wide association study in a Chinese population identifies a susceptibility locus for type 2 diabetes at 7q32 near PAX4. <i>Diabetologia</i> 56, 1291-1305, doi:10.1007/s00125-013-2874-4 (2013).
57	Manning, A. K. et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. <i>Nat Genet</i> 44, 659-669, doi:10.1038/ng.2274 (2012).
58	Meigs, J. B. et al. Genome-wide association with diabetes-related traits in the Framingham Heart Study. <i>BMC Med Genet</i> 8 Suppl 1, S16, doi:10.1186/1471-2350-8-S1-S16 (2007).
59	Middelberg, R. P. et al. Genetic variants in LPL, OASL and TOMM40/APOE-C1-C2-C4 genes are associated with multiple cardiovascular-related traits. <i>BMC Med Genet</i> 12, 123, doi:10.1186/1471-2350-12-123 (2011).
60	Myocardial Infarction Genetics, C. et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. <i>Nat Genet</i> 41, 334-341, doi:10.1038/ng.327 (2009).
61	Newton-Cheh, C. et al. Genome-wide association study identifies eight loci associated with blood pressure. <i>Nat Genet</i> 41, 666-676, doi:10.1038/ng.361 (2009).
62	Ng, M. C. et al. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. <i>PLoS Genet</i> 10, e1004517, doi:10.1371/journal.pgen.1004517 (2014).

63	Nikpay, M. et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. <i>Nat Genet</i> 47, 1121-1130, doi:10.1038/ng.3396 (2015).
64	Ober, C. et al. Genome-wide association study of plasma lipoprotein(a) levels identifies multiple genes on chromosome 6q. <i>J Lipid Res</i> 50, 798-806, doi:10.1194/jlr.M800515-JLR200 (2009).
65	Padmanabhan, S. et al. Genome-wide association study of blood pressure extremes identifies variant near UMOD associated with hypertension. <i>PLoS Genet</i> 6, e1001177, doi:10.1371/journal.pgen.1001177 (2010).
66	Palmer, N. D. et al. A genome-wide association search for type 2 diabetes genes in African Americans. <i>PLoS One</i> 7, e29202, doi:10.1371/journal.pone.0029202 (2012).
67	Parra, E. J. et al. Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County, Texas. <i>Diabetologia</i> 54, 2038-2046, doi:10.1007/s00125-011-2172-y (2011).
68	Perry, J. R. et al. Stratifying type 2 diabetes cases by BMI identifies genetic risk variants in LAMA1 and enrichment for risk variants in lean compared to obese cases. <i>PLoS Genet</i> 8, e1002741, doi:10.1371/journal.pgen.1002741 (2012).
69	Pollin, T. I. et al. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. <i>Science</i> 322, 1702-1705, doi:10.1126/science.1161524 (2008).
70	Qi, L. et al. Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. <i>Hum Mol Genet</i> 19, 2706-2715, doi:10.1093/hmg/ddq156 (2010).
71	Qi, L. et al. Association between a genetic variant related to glutamic acid metabolism and coronary heart disease in individuals with type 2 diabetes. <i>JAMA</i> 310, 821-828, doi:10.1001/jama.2013.276305 (2013).
72	Qi, Q., Workalemahu, T., Zhang, C., Hu, F. B. & Qi, L. Genetic variants, plasma lipoprotein(a) levels, and risk of cardiovascular morbidity and mortality among two prospective cohorts of type 2 diabetes. <i>Eur Heart J</i> 33, 325-334, doi:10.1093/eurheartj/ehr350 (2012).
73	Reilly, M. P. et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. <i>Lancet</i> 377, 383-392, doi:10.1016/S0140-6736(10)61996-4 (2011).
74	Replication, D. I. G. et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. <i>Nat Genet</i> 46, 234-244, doi:10.1038/ng.2897 (2014).
75	Rice, J. P. et al. CHRNB3 is more strongly associated with Fagerstrom test for cigarette dependence-based nicotine dependence than cigarettes per day: phenotype definition changes genome-wide association studies results. <i>Addiction</i> 107, 2019-2028, doi:10.1111/j.1360-0443.2012.03922.x (2012).
76	Ridker, P. M. et al. Polymorphism in the CETP gene region, HDL cholesterol, and risk of future myocardial infarction: Genomewide analysis among 18 245 initially healthy women from the Women's Genome Health Study. <i>Circ Cardiovasc Genet</i> 2, 26-33, doi:10.1161/CIRCGENETICS.108.817304 (2009).
77	Rung, J. et al. Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. <i>Nat Genet</i> 41, 1110-1115, doi:10.1038/ng.443 (2009).
78	Sabatti, C. et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. <i>Nat Genet</i> 41, 35-46, doi:10.1038/ng.271 (2009).
79	Samani, N. J. et al. Genomewide association analysis of coronary artery disease. <i>N Engl J Med</i> 357, 443-453, doi:10.1056/NEJMoa072366 (2007).
80	Sandhu, M. S. et al. LDL-cholesterol concentrations: a genome-wide association study. <i>Lancet</i> 371, 483-491, doi:10.1016/S0140-6736(08)60208-1 (2008).
81	Saxena, R. et al. Genome-wide association study identifies a novel locus contributing to type 2 diabetes susceptibility in Sikhs of Punjabi origin from India. <i>Diabetes</i> 62, 1746-1755, doi:10.2337/db12-1077 (2013).
82	Schunkert, H. et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. <i>Nat Genet</i> 43, 333-338, doi:10.1038/ng.784 (2011).

83	Scott, L. J. et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. <i>Science</i> 316, 1341-1345, doi:10.1126/science.1142382 (2007).
84	Shen, H. et al. Familial defective apolipoprotein B-100 and increased low-density lipoprotein cholesterol and coronary artery calcification in the old order amish. <i>Arch Intern Med</i> 170, 1850-1855, doi:10.1001/archinternmed.2010.384 (2010).
85	Shu, X. O. et al. Identification of new genetic risk variants for type 2 diabetes. <i>PLoS Genet</i> 6, e1001127, doi:10.1371/journal.pgen.1001127 (2010).
86	Sim, X. et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. <i>PLoS Genet</i> 7, e1001363, doi:10.1371/journal.pgen.1001363 (2011).
87	Simino, J. et al. Gene-age interactions in blood pressure regulation: a large-scale investigation with the CHARGE, Global BPgen, and ICBP Consortia. <i>Am J Hum Genet</i> 95, 24-38, doi:10.1016/j.ajhg.2014.05.010 (2014).
88	Simino, J., Sung, Y. J., Kume, R., Schwander, K. & Rao, D. C. Gene-alcohol interactions identify several novel blood pressure loci including a promising locus near SLC16A9. <i>Front Genet</i> 4, 277, doi:10.3389/fgene.2013.00277 (2013).
89	Slavin, T. P., Feng, T., Schnell, A., Zhu, X. & Elston, R. C. Two-marker association tests yield new disease associations for coronary artery disease and hypertension. <i>Hum Genet</i> 130, 725-733, doi:10.1007/s00439-011-1009-6 (2011).
90	Smith, E. N. et al. Longitudinal genome-wide association of cardiovascular disease risk factors in the Bogalusa heart study. <i>PLoS Genet</i> 6, e1001094, doi:10.1371/journal.pgen.1001094 (2010).
91	Sung, Y. J., de Las Fuentes, L., Schwander, K. L., Simino, J. & Rao, D. C. Gene-smoking interactions identify several novel blood pressure loci in the Framingham Heart Study. <i>Am J Hypertens</i> 28, 343-354, doi:10.1093/ajh/hpu149 (2015).
92	Surakka, I. et al. A genome-wide screen for interactions reveals a new locus on 4p15 modifying the effect of waist-to-hip ratio on total cholesterol. <i>PLoS Genet</i> 7, e1002333, doi:10.1371/journal.pgen.1002333 (2011).
93	Surakka, I. et al. A genome-wide association study of monozygotic twin-pairs suggests a locus related to variability of serum high-density lipoprotein cholesterol. <i>Twin Res Hum Genet</i> 15, 691-699, doi:10.1017/thg.2012.63 (2012).
94	Takeuchi, F. et al. Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. <i>Diabetes</i> 58, 1690-1699, doi:10.2337/db08-1494 (2009).
95	Takeuchi, F. et al. Genome-wide association study of coronary artery disease in the Japanese. <i>Eur J Hum Genet</i> 20, 333-340, doi:10.1038/ejhg.2011.184 (2012).
96	Teslovich, T. M. et al. Biological, clinical and population relevance of 95 loci for blood lipids. <i>Nature</i> 466, 707-713, doi:10.1038/nature09270 (2010).
97	Teupser, D. et al. Genetic regulation of serum phytosterol levels and risk of coronary artery disease. <i>Circ Cardiovasc Genet</i> 3, 331-339, doi:10.1161/CIRCGENETICS.109.907873 (2010).
98	Thorgeirsson, T. E. et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. <i>Nature</i> 452, 638-642, doi:10.1038/nature06846 (2008).
99	Timpson, N. J. et al. Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. <i>Diabetes</i> 58, 505-510, doi:10.2337/db08-0906 (2009).
100	Todd, J. A. et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. <i>Nat Genet</i> 39, 857-864, doi:10.1038/ng2068 (2007).
101	Traylor, M. et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. <i>Lancet Neurol</i> 11, 951-962, doi:10.1016/S1474-4422(12)70234-X (2012).
102	Tregouet, D. A. et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. <i>Nat Genet</i> 41, 283-285, doi:10.1038/ng.314 (2009).

103	Tsai, F. J. et al. A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. <i>PLoS Genet</i> 6, e1000847, doi:10.1371/journal.pgen.1000847 (2010).
104	Unoki, H. et al. SNPs in <i>KCNQ1</i> are associated with susceptibility to type 2 diabetes in East Asian and European populations. <i>Nat Genet</i> 40, 1098-1102, doi:10.1038/ng.208 (2008).
105	van Setten, J. et al. Genome-wide association study of coronary and aortic calcification implicates risk loci for coronary artery disease and myocardial infarction. <i>Atherosclerosis</i> 228, 400-405, doi:10.1016/j.atherosclerosis.2013.02.039 (2013).
106	Voight, B. F. et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. <i>Nat Genet</i> 42, 579-589, doi:10.1038/ng.609 (2010).
107	Wain, L. V. et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. <i>Nat Genet</i> 43, 1005-1011, doi:10.1038/ng.922 (2011).
108	Wallace, C. et al. The imprinted <i>DLK1-MEG3</i> gene region on chromosome 14q32.2 alters susceptibility to type 1 diabetes. <i>Nat Genet</i> 42, 68-71, doi:10.1038/ng.493 (2010).
109	Warren, H. R. et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. <i>Nat Genet</i> 49, 403-415, doi:10.1038/ng.3768 (2017).
110	Waterworth, D. M. et al. Genetic variants influencing circulating lipid levels and risk of coronary artery disease. <i>Arterioscler Thromb Vasc Biol</i> 30, 2264-2276, doi:10.1161/ATVBAHA.109.201020 (2010).
111	Webb, T. R. et al. Systematic Evaluation of Pleiotropy Identifies 6 Further Loci Associated With Coronary Artery Disease. <i>J Am Coll Cardiol</i> 69, 823-836, doi:10.1016/j.jacc.2016.11.056 (2017).
112	Weissglas-Volkov, D. et al. Genomic study in Mexicans identifies a new locus for triglycerides and refines European lipid loci. <i>J Med Genet</i> 50, 298-308, doi:10.1136/jmedgenet-2012-101461 (2013).
113	Wellcome Trust Case Control, C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. <i>Nature</i> 447, 661-678, doi:10.1038/nature05911 (2007).
114	Wild, P. S. et al. A genome-wide association study identifies <i>LIPA</i> as a susceptibility gene for coronary artery disease. <i>Circ Cardiovasc Genet</i> 4, 403-412, doi:10.1161/CIRCGENETICS.110.958728 (2011).
115	Willer, C. J. et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. <i>Nat Genet</i> 40, 161-169, doi:10.1038/ng.76 (2008).
116	Willer, C. J. et al. Discovery and refinement of loci associated with lipid levels. <i>Nat Genet</i> 45, 1274-1283, doi:10.1038/ng.2797 (2013).
117	Wu, Y. et al. Genetic association with lipids in Filipinos: waist circumference modifies an <i>APOA5</i> effect on triglyceride levels. <i>J Lipid Res</i> 54, 3198-3205, doi:10.1194/jlr.P042077 (2013).
118	Xing, C., Cohen, J. C. & Boerwinkle, E. A weighted false discovery rate control procedure reveals alleles at <i>FOXA2</i> that influence fasting glucose levels. <i>Am J Hum Genet</i> 86, 440-446, doi:10.1016/j.ajhg.2010.01.025 (2010).
119	Yamauchi, T. et al. A genome-wide association study in the Japanese population identifies susceptibility loci for type 2 diabetes at <i>UBE2E2</i> and <i>C2CD4A-C2CD4B</i> . <i>Nat Genet</i> 42, 864-868, doi:10.1038/ng.660 (2010).
120	Yang, H. C. et al. Genome-wide association study of young-onset hypertension in the Han Chinese population of Taiwan. <i>PLoS One</i> 4, e5459, doi:10.1371/journal.pone.0005459 (2009).
121	Yasuda, K. et al. Variants in <i>KCNQ1</i> are associated with susceptibility to type 2 diabetes mellitus. <i>Nat Genet</i> 40, 1092-1097, doi:10.1038/ng.207 (2008).
122	Zeggini, E. et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. <i>Nat Genet</i> 40, 638-645, doi:10.1038/ng.120 (2008).
123	Zeggini, E. et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. <i>Science</i> 316, 1336-1341, doi:10.1126/science.1142364 (2007).
124	Zhou, L. et al. A genome wide association study identifies common variants associated with lipid levels in the Chinese population. <i>PLoS One</i> 8, e82420, doi:10.1371/journal.pone.0082420 (2013).

Supplementary Table 10 | Atherosclerosis SNPs

Locus	lead SNP	Position	p-value
1p32.3	rs11206510	55.496.039	1.00x10 ⁻⁰⁸
1p32.2	rs17114036	56.962.821	4.00x10 ⁻¹⁹
1p32.2	rs9970807	56.965.664	5x10 ⁻¹⁴
1p13.3	rs7528419	109.817.192	1.97x10 ⁻²³
1p13.3	rs660240	109.817.838	2.00x10 ⁻²²
1p13.3	rs646776	109.818.530	8.00x10 ⁻¹²
1p13.3	rs602633	109.821.511	1.00x10 ⁻⁰⁸
1p13.3	rs599839	109.822.166	8.00x10 ⁻¹⁷
1q21.3	rs6689306	154.395.946	2.6x10 ⁻⁰⁹
1q21.3	rs4845625	154.422.067	3.93x10 ⁻⁰⁸
1q24.2	rs10919065	169.093.557	2.25x10 ⁻⁰⁸
1q24.2	rs1892094	169.094.459	1.57x10 ⁻⁰⁸
1q24.2	rs1200159	169.100.241	3.4x10 ⁻⁰⁸
1q25.3	rs10911021	182.081.960	2.00x10 ⁻⁰⁸
1q32.1	rs6700559	200.646.073	1.13x10 ⁻⁰⁸
1q32.1	rs2820315	201.872.264	7.7x10 ⁻¹⁰
1q32.1	rs2819348	201.884.952	1.77x10 ⁻⁰⁸
1q41	rs35700460	222.811.407	1.88x10 ⁻¹¹
1q41	rs17465637	222.823.529	1.00x10 ⁻⁰⁹
1q41	rs67180937	222.823.743	1.01x10 ⁻¹²
2p24.1	rs16986953	19.942.473	1.45x10 ⁻⁰⁸
2p24.1	rs2123536	19.945.577	6.83x10 ⁻¹¹
2p24.1	rs10199768	21.244.000	8.00x10 ⁻¹⁵
2p24.1	rs515135	21.286.057	3.09x10 ⁻⁰⁸
2p24.1	rs34908258	21.378.432	2.89x10 ⁻⁰⁸
2p23.3	rs1260326	27.730.940	2.00x10 ⁻⁰⁸
2p21	rs6544713	44.073.881	2.12x10 ⁻⁰⁹
2p21	rs4953024	44.074.128	2.6x10 ⁻⁰⁸
2p16.2	rs1559040	54.347.750	4.00x10 ⁻⁰⁸
2p11.2	rs10176176	85.762.048	2.88x10 ⁻¹⁰
2p11.2	rs7568458	85.788.175	3.62x10 ⁻¹⁰
2p11.2	rs1561198	85.809.989	6.37x10 ⁻¹⁰
2q22.3	rs17678683	145.286.559	3x10 ⁻⁰⁹
2q22.3	rs2252641	145.801.461	6.77x10 ⁻⁰⁹
2q24.2	rs4665058	160.190.209	2.00x10 ⁻¹⁰
2q31.3	rs16866933	180.566.678	6.00x10 ⁻¹⁴
2q33.2	rs6725887	203.745.885	1.00x10 ⁻⁰⁹
2q33.2	rs7582720	203.753.072	4.00x10 ⁻⁰⁹
2q33.2	rs577594671	203.828.797	2.15x10 ⁻¹⁸

Locus	lead SNP	Position	p-value
2q35	rs2571445	218.683.154	4.55x10 ⁻¹⁰
2q37.1	rs1801251	233.633.460	1.48x10 ⁻⁰⁹
3p24.3	rs4618210	17.124.384	2.60x10 ⁻⁰⁹
3q22.3	rs139016349	138.099.162	2.89x10 ⁻⁰⁹
3q22.3	rs2306374	138.119.952	3.34x10 ⁻⁰⁸
3q22.3	rs9818870	138.122.122	7.00x10 ⁻¹³
3q26.0	rs11924705	162.161.616	4.00x10 ⁻¹⁴
3q26.0	rs6789378	162.167.396	4.00x10 ⁻¹⁴
3q26.1	rs1803274	165.491.280	6x10 ⁻⁷⁶
4p16.2	rs7697839	5.367.985	2.00x10 ⁻¹¹
4p16.2	rs7673097	5.368.224	2.00x10 ⁻¹¹
4p16.1	rs7671266	10.056.376	9x10 ⁻⁵⁶
4p16.1	rs4698036	10.331.294	2.00x10 ⁻⁵²
4q12	rs17087335	57.838.583	4.6x10 ⁻⁰⁸
4q22.1	rs2199936	89.045.331	2.00x10 ⁻¹⁷
4q31.22	rs4593108	148.281.001	8.82x10 ⁻¹⁰
4q31.22	rs1878406	148.393.664	7.00x10 ⁻¹²
4q31.22	rs6841581	148.401.190	2.00x10 ⁻⁰⁸
4q32.1	rs1842896	156.511.459	1.00x10 ⁻¹¹
4q32.1	rs7692387	156.635.309	2.65x10 ⁻¹¹
4q32.1	rs72689147	156.639.888	6.07x10 ⁻⁰⁹
5p15.33	rs11748327	4.029.789	5.00x10 ⁻¹³
5q22.2	rs4621553	113.030.164	4.00x10 ⁻⁰⁸
5q31.1	rs273909	131.667.353	9.62x10 ⁻¹⁰
5q31.3	rs246600	142.516.897	1.29x10 ⁻⁰⁸
5q33.2	rs12189362	153.057.548	3.00x10 ⁻¹⁰
6p24.1	rs9369640	12.901.441	3.00x10 ⁻¹¹
6p24.1	rs4714955	12.903.435	2.00x10 ⁻¹¹
6p24.1	rs9349379	12.903.957	1.81x10 ⁻⁴²
6p24.1	rs12526453	12.927.544	1.00x10 ⁻⁰⁹
6p22.2	rs11754288	25.776.949	4.00x10 ⁻⁰⁹
6p22.2	rs1800562	26.093.141	5.00x10 ⁻¹²
6p21.33	rs3869109	31.184.196	1.00x10 ⁻⁰⁹
6p21.33	rs3130683	31.888.367	1.04x10 ⁻¹¹
6p21.32	rs9268402	32.341.353	2.77x10 ⁻¹⁵
6p21.31	rs17609940	35.034.800	1.00x10 ⁻⁰⁸
6p21.2	rs1544935	39.124.448	2.89x10 ⁻⁰⁸
6p21.2	rs56336142	39.134.099	1.85x10 ⁻⁰⁸
6p21.2	rs10947789	39.174.922	9.81x10 ⁻⁰⁹

Locus	lead SNP	Position	p-value
6q13	rs9351814	72.193.707	5.00x10 ⁻⁰⁸
6q23.2	rs12202017	134.173.151	1.98x10 ⁻¹¹
6q23.2	rs12190287	134.214.525	1.00x10 ⁻¹²
6q25.1	rs6922269	151.252.985	2.90x10 ⁻⁰⁸
6q25.1	rs2982694	152.285.687	7.00x10 ⁻¹⁰
6q25.3	rs2048327	160.863.532	4.00x10 ⁻¹⁵
6q25.3	rs3127599	160.907.134	4.00x10 ⁻¹⁵
6q25.3	rs3798220	160.961.137	3.00x10 ⁻¹¹
6q25.3	rs7767084	160.962.503	4.00x10 ⁻¹⁵
6q25.3	rs10755578	160.969.738	4.00x10 ⁻¹⁵
6q25.3	rs55730499	161.005.610	5.39x10 ⁻³⁹
6q25.3	rs10455872	161.010.118	9.00x10 ⁻¹⁴
6q26	rs2315065	161.108.144	2.88x10 ⁻³⁴
6q26	rs4252185	161.123.451	1.64x10 ⁻³²
6q26	rs4252120	161.143.608	4.88x10 ⁻¹⁰
7p21.1	rs2023938	19.036.775	4.94x10 ⁻⁰⁸
7p21.1	rs2107595	19.049.388	2.00x10 ⁻¹⁶
7q22.3	rs10953541	107.244.545	3.12x10 ⁻⁰⁸
7q32.2	rs11556924	129.663.496	5.34x10 ⁻¹¹
7q34	rs10237377	139.757.136	1.75x10 ⁻⁰⁸
7q36.1	rs3918226	150.690.176	1.7x10 ⁻⁰⁹
8p21.3	rs264	19.813.180	2.88x10 ⁻⁰⁹
8p21.3	rs17091905	19.849.757	5.00x10 ⁻¹⁵
8q24.13	rs2954029	126.490.972	4.75x10 ⁻⁰⁹
9p21.3	rs3217992	22.003.223	1.03x10 ⁻⁴²
9p21.3	rs7865618	22.031.005	2.00x10 ⁻²⁷
9p21.3	rs1537370	22.084.310	2.00x10 ⁻¹¹
9p21.3	rs10757274	22.096.055	8.00x10 ⁻⁴⁵
9p21.3	rs4977574	22.098.574	3x10 ⁻²⁸
9p21.3	rs2891168	22.098.619	2.29x10 ⁻⁹⁸
9p21.3	rs1333042	22.103.813	1.00x10 ⁻⁰⁹
9p21.3	rs944797	22.115.286	6.00x10 ⁻¹⁶
9p21.3	rs10757278	22.124.477	1.00x10 ⁻²⁰
9p21.3	rs1333048	22.125.347	7.00x10 ⁻¹⁴
9p21.3	rs1333049	22.125.503	1.00x10 ⁻⁵⁹
9q34.2	rs2519093	136.141.870	1.19x10 ⁻¹¹
9q34.2	rs514659	136.142.203	8x10 ⁻⁰⁹
9q34.2	rs532436	136.149.830	2.31x10 ⁻¹⁷
9q34.2	rs579459	136.154.168	4.00x10 ⁻¹⁴
10p11.23	rs3739998	30.316.072	1.00x10 ⁻¹¹

Locus	lead SNP	Position	p-value
10p11.23	rs2487928	30.323.892	4.41x10 ⁻¹¹
10p11.23	rs2505083	30.335.122	4.00x10 ⁻⁰⁸
10q11.21	rs1870634	44.480.811	5.55x10 ⁻¹⁵
10q11.21	rs2047009	44.539.913	2.75x10 ⁻¹¹
10q11.21	rs501120	44.753.867	1.39x10 ⁻¹¹
10q11.21	rs1746048	44.775.824	3.00x10 ⁻¹⁰
10q23.31	rs1412444	91.002.927	3.00x10 ⁻¹³
10q23.31	rs1332329	91.003.419	2.64x10 ⁻¹³
10q23.31	rs2246833	91.005.854	3.26x10 ⁻¹¹
10q24.32	rs11191416	104.604.916	4.65x10 ⁻⁰⁹
10q24.32	rs12413409	104.719.096	1.00x10 ⁻⁰⁹
11p15.4	rs10840293	9.751.196	1.3x10 ⁻⁰⁸
11p15.4	rs11042937	10.745.394	1.18x10 ⁻⁰⁸
11q13.1	rs12801636	65.391.317	9.71x10 ⁻⁰⁹
11q13.5	rs590121	75.274.150	1.54x10 ⁻⁰⁸
11q21	rs10765792	95.866.700	8.00x10 ⁻¹³
11q22.3	rs974819	103.660.567	2.00x10 ⁻⁰⁹
11q22.3	rs2019090	103.668.962	3.6x10 ⁻⁰⁹
11q22.3	rs2128739	103.673.277	7.05x10 ⁻¹¹
11q23.3	rs964184	116.648.917	1.00x10 ⁻¹⁷
11q23.3	rs508487	117.075.566	2.00x10 ⁻¹⁰
12q13.3	rs11172113	57.527.283	9.25x10 ⁻¹⁴
12q21.2	rs7307780	76.220.618	5.00x10 ⁻¹⁵
12q21.33	rs2681472	90.008.959	6.17x10 ⁻¹¹
12q21.33	rs7136259	90.081.188	6.00x10 ⁻¹⁰
12q23.3	rs1165668	104.317.996	3.05x10 ⁻⁰⁹
12q23.3	rs1165669	104.318.172	5.32x10 ⁻⁰⁹
12q24.11	rs3782889	111.350.655	4.00x10 ⁻¹⁴
12q24.12	rs3184504	111.884.608	1.03x10 ⁻⁰⁹
12q24.12	rs653178	112.007.756	2.8x10 ⁻¹¹
12q24.12	rs11065979	112.059.557	1.93x10 ⁻¹⁰
12q24.12	rs11065987	112.072.424	4.00x10 ⁻¹⁴
12q24.12	rs3782886	112.110.489	1.00x10 ⁻¹⁴
12q24.12	rs11066015	112.168.009	4.51x10 ⁻¹¹
12q24.12	rs671	112.241.766	2.00x10 ⁻³⁴
12q24.13	rs11066280	112.817.783	2.00x10 ⁻¹¹
12q24.23	rs11830157	118.265.441	2x10 ⁻⁰⁹
12q24.31	rs1169288	121.416.650	4.53x10 ⁻⁰⁸
12q24.31	rs2244608	121.416.988	1.57x10 ⁻⁰⁸
12q24.31	rs2258287	121.454.313	6x10 ⁻⁰⁹

Locus	lead SNP	Position	p-value
12q24.31	rs2708081	121.463.288	1.02x10 ⁻⁰⁸
12q24.31	rs3213545	121.471.337	4.00x10 ⁻¹⁵
12q24.31	rs11057830	125.307.053	5.65x10 ⁻⁰⁹
13q12.3	rs9319428	28.973.621	7.32x10 ⁻¹¹
13q22.1	rs12429889	74.742.322	5.00x10 ⁻²⁰
13q34	rs11617955	110.818.102	3.55x10 ⁻⁰⁸
13q34	rs4773144	110.960.712	4x10 ⁻⁰⁹
13q34	rs11838776	111.040.681	1x10 ⁻¹⁰
13q34	rs9515203	111.049.623	9.33x10 ⁻¹⁰
14q31.3	rs11624056	87.506.248	3.00x10 ⁻⁰⁸
14q32.2	rs2895811	100.133.942	1.00x10 ⁻¹⁰
14q32.2	rs10139550	100.145.710	1.38x10 ⁻⁰⁸
15q21.3	rs10468017	58.678.512	3.00x10 ⁻¹²
15q22.31	rs6494488	65.024.204	2.09x10 ⁻⁰⁸
15q22.33	rs56062135	67.455.630	4.5x10 ⁻⁰⁹
15q25.1	rs899997	79.019.578	2.00x10 ⁻⁰⁸
15q25.1	rs2219939	79.029.723	2.00x10 ⁻⁰⁹
15q25.1	rs1994016	79.080.234	5.00x10 ⁻¹³
15q25.1	rs3825807	79.089.111	1.00x10 ⁻¹²
15q25.1	rs4380028	79.111.093	4.00x10 ⁻⁰⁹
15q25.1	rs7165042	79.123.338	3.02x10 ⁻⁰⁹
15q25.1	rs4468572	79.124.475	4.44x10 ⁻¹⁶
15q25.1	rs7173743	79.141.784	5.55x10 ⁻¹⁶
15q26.1	rs8042271	89.574.218	3.7x10 ⁻⁰⁸
15q26.1	rs17514846	91.416.550	9.33x10 ⁻¹¹
16q13	rs173539	56.988.044	3x10 ⁻⁴⁹
16q13	rs247616	56.989.590	1.00x10 ⁻²³
16q13	rs1800775	56.995.236	9.83x10 ⁻⁰⁹
16q22.2	rs1050362	72.130.815	3.52x10 ⁻⁰⁸
16q22.2	rs2072142	72.132.713	4.26x10 ⁻⁰⁸
17p13.3	rs2281727	2.117.945	7.83x10 ⁻⁰⁹
17p13.3	rs1231206	2.125.605	9.00x10 ⁻¹⁰
17p13.3	rs216172	2.126.504	1.00x10 ⁻⁰⁹
17p11.2	rs12936587	17.543.722	2.00x10 ⁻¹⁰
17p11.2	rs12449964	17.544.704	2.00x10 ⁻⁰⁸
17q21.32	rs17608766	45.013.271	4.14x10 ⁻⁰⁸
17q21.32	rs46522	46.988.597	2.00x10 ⁻⁰⁸
17q23.2	rs7212798	59.013.488	1.9x10 ⁻⁰⁸
17q23.3	rs1867624	62.387.091	3.98x10 ⁻⁰⁸
17q23.3	rs9892152	62.401.965	5.00x10 ⁻⁰⁸

Locus	lead SNP	Position	p-value
17q24.3	rs17718586	68.644.189	2.00x10 ⁻⁰⁸
18p11.31	rs597503	6.939.947	2.00x10 ⁻⁰⁸
18q11.2	rs16942421	24.156.425	8.00x10 ⁻¹⁰
18q21.32	rs663129	57.838.401	3.2x10 ⁻⁰⁸
19p13.3	rs3803915	2.160.529	4.00x10 ⁻⁰⁹
19p13.2	rs1122608	11.163.601	3.00x10 ⁻¹²
19p13.2	rs55791371	11.188.153	2.95x10 ⁻⁰⁸
19p13.2	rs56289821	11.188.247	4.44x10 ⁻¹⁵
19p13.2	rs6511720	11.202.306	5x10 ⁻¹¹
19q13.11	rs12976411	32.882.020	1x10 ⁻¹⁴
19q13.32	rs2075650	45.395.619	2.00x10 ⁻¹⁴
19q13.32	rs445925	45.415.640	9.00x10 ⁻¹⁹
19q13.32	rs56131196	45.422.846	3.87x10 ⁻⁰⁸
19q13.32	rs4420638	45.422.946	7.07x10 ⁻¹¹
20q11.22	rs867186	33.764.554	2.7x10 ⁻⁰⁹
21q22.11	rs28451064	35.593.827	1.33x10 ⁻¹⁵
21q22.11	rs9982601	35.599.128	6.00x10 ⁻¹¹
22q11.23	rs180803	24.658.858	1.6x10 ⁻¹⁰

Supplementary Table 11 | Lipid SNPs

Locus	lead SNP	Position	p-value
1p36.11	rs12027135	25.775.733	2.00x10 ⁻¹⁴
1p36.11	rs12748152	27.138.393	1.00x10 ⁻¹⁵
1p34.3	rs4660293	40.028.180	3.00x10 ⁻¹⁸
1p32.3	rs11206510	55.496.039	4.00x10 ⁻¹¹
1p32.3	rs2479409	55.504.650	3.00x10 ⁻⁵⁰
1p32.3	rs11591147	55.505.647	2.00x10 ⁻⁴⁴
1p32.3	rs17111684	55.625.548	2.00x10 ⁻¹⁷
1p32.2	rs72669744	56.116.505	1.00x10 ⁻¹⁹
1p31.3	rs1167998	62.931.632	2.00x10 ⁻¹²
1p31.3	rs1168029	62.969.402	2.66x10 ⁻¹³
1p31.3	rs2131925	63.025.942	3.00x10 ⁻⁵⁸
1p31.3	rs11207995	63.049.551	6.00x10 ⁻¹⁰
1p31.3	rs1748195	63.049.593	2.00x10 ⁻¹⁰
1p31.3	rs10889353	63.118.196	1.00x10 ⁻¹⁴
1p31.3	rs12130333	63.191.777	2.00x10 ⁻⁰⁸
1p13.3	rs12740374	109.817.590	2.00x10 ⁻⁴²
1p13.3	rs660240	109.817.838	1.00x10 ⁻²⁶
1p13.3	rs629301	109.818.306	5.00x10 ⁻²⁴¹
1p13.3	rs646776	109.818.530	2.00x10 ⁻⁵³
1p13.3	rs599839	109.822.166	1.00x10 ⁻³³
1q21.3	rs267733	150.958.836	5.00x10 ⁻⁰⁹
1q23.1	rs12145743	156.700.651	2.00x10 ⁻⁰⁸
1q23.3	rs67418890	161.540.857	1.00x10 ⁻¹⁰
1q25.2	rs4650994	178.515.312	7.00x10 ⁻⁰⁹
1q25.3	rs1689800	182.168.885	5.00x10 ⁻²⁰
1q41	rs2642442	220.973.563	3.00x10 ⁻¹¹
1q42.13	rs2144300	230.294.916	3.00x10 ⁻¹⁴
1q42.13	rs4846914	230.295.691	4.00x10 ⁻⁴¹
1q42.13	rs10489615	230.304.988	4.00x10 ⁻⁰⁹
1q42.3	rs514230	234.858.597	9.00x10 ⁻¹²
2p24.1	rs4971516	20.903.015	2.00x10 ⁻⁵²
2p24.1	rs6754295	21.206.183	4.00x10 ⁻⁴⁷
2p24.1	rs7557067	21.208.211	9.00x10 ⁻¹²
2p24.1	rs1042034	21.225.281	1.00x10 ⁻⁴⁵
2p24.1	rs676210	21.231.524	4.00x10 ⁻⁶⁴
2p24.1	rs693	21.232.195	1.00x10 ⁻²¹
2p24.1	rs673548	21.237.544	5.00x10 ⁻⁴²
2p24.1	rs12713956	21.241.505	4.00x10 ⁻⁰⁸
2p24.1	rs1367117	21.263.900	1.00x10 ⁻¹⁸²

Locus	lead SNP	Position	p-value
2p24.1	rs515135	21.286.057	5.00x10 ⁻²⁹
2p24.1	rs562338	21.288.321	6.00x10 ⁻²²
2p24.1	rs312985	21.378.805	1.00x10 ⁻¹⁷
2p24.1	rs506585	21.397.182	1.00x10 ⁻²⁵
2p24.1	rs503662	21.414.142	2.50x10 ⁻⁰⁹
2p23.3	rs4665972	27.598.097	1.00x10 ⁻⁰⁸
2p23.3	rs1260326	27.730.940	2.00x10 ⁻²³⁹
2p23.3	rs780094	27.741.237	6.00x10 ⁻³²
2p23.3	rs780092	27.743.154	5.00x10 ⁻²⁷
2p23.3	rs1260333	27.748.624	2.00x10 ⁻¹⁹
2p21	rs6756629	44.065.090	3.00x10 ⁻¹⁰
2p21	rs4299376	44.072.576	4.00x10 ⁻⁷²
2p21	rs41360247	44.073.656	6.00x10 ⁻⁰⁹
2p21	rs6544713	44.073.881	2.00x10 ⁻²⁰
2p21	rs4245791	44.074.431	2.00x10 ⁻⁵⁰
2p15	rs2710642	63.149.557	6.00x10 ⁻⁰⁹
2q14.1	rs10490626	118.835.841	2.00x10 ⁻¹²
2q14.2	rs2030746	121.309.488	9.00x10 ⁻⁰⁹
2q24.3	rs10195252	165.513.091	2.00x10 ⁻¹⁰
2q24.3	rs12328675	165.540.800	2.00x10 ⁻¹⁵
2q34	rs1047891	211.540.507	9.00x10 ⁻¹⁰
2q35	rs1250229	216.304.384	3.00x10 ⁻⁰⁸
2q36.3	rs2972146	227.100.698	2.00x10 ⁻¹⁷
2q37.1	rs11563251	234.679.384	5.00x10 ⁻⁰⁸
3p25.3	rs2606736	11.400.249	5.00x10 ⁻⁰⁸
3p22.3	rs7640978	32.533.010	1.00x10 ⁻⁰⁸
3p21.31	rs2290547	47.061.183	4.00x10 ⁻⁰⁹
3p21.31	rs2013208	50.129.399	9.00x10 ⁻¹²
3p21.1	rs13326165	52.532.118	9.00x10 ⁻¹¹
3q13.33	rs6805251	119.560.606	1.00x10 ⁻⁰⁸
3q22.1	rs17404153	132.163.200	2.00x10 ⁻⁰⁹
3q22.3	rs645040	135.926.622	2.00x10 ⁻¹²
4p16.3	rs6831256	3.473.139	2.00x10 ⁻¹²
4p15.2	rs10019888	26.062.990	5.00x10 ⁻⁰⁸
4p15.1	rs6448771	31.397.618	5.00x10 ⁻⁰⁹
4q13.3	rs115136538	73.322.565	5.00x10 ⁻¹⁸
4q22.1	rs442177	88.030.261	1.00x10 ⁻¹⁸
4q22.1	rs3822072	89.741.269	4.00x10 ⁻¹²
4q23	rs2602836	100.014.805	5.00x10 ⁻⁰⁸

Locus	lead SNP	Position	p-value
4q24	rs13107325	103.188.709	1.00x10 ⁻¹⁵
5q11.2	rs6450176	53.298.025	7.00x10 ⁻¹⁰
5q11.2	rs9686661	55.861.786	3.00x10 ⁻¹⁶
5q13.3	rs7703051	74.625.487	1.00x10 ⁻⁰⁸
5q13.3	rs10045497	74.636.484	1.00x10 ⁻¹²
5q13.3	rs12654264	74.648.603	1.00x10 ⁻²⁰
5q13.3	rs3846662	74.651.084	2.00x10 ⁻¹¹
5q13.3	rs3846663	74.655.726	8.00x10 ⁻¹²
5q13.3	rs12916	74.656.539	8.00x10 ⁻⁷⁸
5q23.2	rs4530754	122.855.416	4.00x10 ⁻¹²
5q33.3	rs6882076	156.390.297	2.00x10 ⁻²²
5q33.3	rs1501908	156.398.169	1.00x10 ⁻¹¹
6p22.3	rs3757354	16.127.407	2.00x10 ⁻¹⁷
6p22.3	rs2142672	16.197.194	2.00x10 ⁻⁰⁸
6p22.2	rs1800562	26.093.141	8.00x10 ⁻¹⁴
6p22.1	rs6917603	30.017.071	3.00x10 ⁻¹³
6p21.33	rs2247056	31.265.490	2.00x10 ⁻¹⁵
6p21.32	rs3177928	32.412.435	3.00x10 ⁻¹⁷
6p21.32	rs2254287	33.143.948	5.00x10 ⁻⁰⁸
6p21.31	rs2814944	34.552.797	4.00x10 ⁻⁰⁹
6p21.1	rs998584	43.757.896	3.00x10 ⁻¹⁵
6q13	rs78536982	70.182.710	7.00x10 ⁻⁰⁹
6q22.1	rs9488822	116.312.893	1.70x10 ⁻¹⁰
6q22.33	rs1936800	127.436.064	3.00x10 ⁻¹⁰
6q24.1	rs605066	139.829.666	3.00x10 ⁻⁰⁸
6q25.3	rs17086702	156.150.733	4.00x10 ⁻⁰⁹
6q25.3	rs17087007	156.384.967	2.00x10 ⁻⁰⁸
6q25.3	rs9384478	156.916.766	7.28x10 ⁻⁰⁹
6q25.3	rs9356009	158.219.013	4.00x10 ⁻¹⁹
6q25.3	rs9346930	158.263.456	4.79x10 ⁻¹⁷
6q25.3	rs9347800	158.371.084	2.00x10 ⁻¹⁶
6q25.3	rs2475553	158.533.560	2.00x10 ⁻⁰⁸
6q25.3	rs9457516	159.335.251	5.72x10 ⁻⁰⁹
6q25.3	rs17830011	159.432.566	5.00x10 ⁻⁰⁹
6q25.3	rs926657	159.463.452	5.00x10 ⁻⁰⁹
6q25.3	rs294914	159.629.045	2.00x10 ⁻¹¹
6q25.3	rs1544167	159.801.024	4.00x10 ⁻¹⁰
6q25.3	rs1564348	160.578.860	3.00x10 ⁻²³
6q25.3	rs3120139	160.741.622	4.00x10 ⁻⁰⁹
6q25.3	rs12214416	160.910.517	5.00x10 ⁻⁰⁸

Locus	lead SNP	Position	p-value
6q25.3	rs6919346	160.960.359	4.00x10 ⁻¹¹
6q25.3	rs10455872	161.010.118	5.00x10 ⁻³⁹
6q26	rs1084651	161.089.817	3.00x10 ⁻⁰⁸
6q26	rs783147	161.137.990	3.00x10 ⁻¹⁷
6q26	rs1620921	161.197.087	5.00x10 ⁻⁰⁸
6q26	rs2293289	161.551.439	2.97x10 ⁻⁰⁸
6q27	rs62436827	167.548.547	7.00x10 ⁻⁰⁹
7p22.1	rs702485	6.449.272	6.00x10 ⁻¹²
7p21.1	rs4142995	17.919.258	9.00x10 ⁻¹²
7p15.3	rs12670798	21.607.352	5.00x10 ⁻¹⁴
7p15.2	rs4722551	25.991.826	4.00x10 ⁻¹⁴
7p13	rs2072183	44.579.180	7.00x10 ⁻¹⁶
7p12.2	rs4917014	50.305.863	1.00x10 ⁻⁰⁸
7q11.22	rs13238203	72.129.667	1.00x10 ⁻⁰⁹
7q11.23	rs2240466	72.856.269	1.00x10 ⁻¹²
7q11.23	rs1178979	72.856.430	2.00x10 ⁻¹²
7q11.23	rs714052	72.864.869	3.00x10 ⁻¹⁵
7q11.23	rs17145738	72.982.874	9.00x10 ⁻⁹⁹
7q11.23	rs2286276	72.987.354	1.00x10 ⁻¹⁵
7q11.23	rs13247874	73.010.442	8.00x10 ⁻¹⁴
7q11.23	rs3812316	73.020.337	1.00x10 ⁻¹⁰
7q21.11	rs2366858	80.340.622	6.00x10 ⁻¹⁰
7q31.2	rs38855	116.358.044	2.00x10 ⁻⁰⁸
7q32.2	rs4731702	130.433.384	5.00x10 ⁻¹⁷
7q36.1	rs17173637	150.529.449	2.00x10 ⁻⁰⁸
8p23.3	rs28680850	1.373.720	7.00x10 ⁻⁰⁹
8p23.1	rs9987289	9.183.358	2.00x10 ⁻⁴¹
8p23.1	rs6601299	9.184.691	5.62x10 ⁻⁰⁹
8p23.1	rs2126259	9.185.146	7.00x10 ⁻¹²
8p23.1	rs1461729	9.187.242	7.39x10 ⁻⁰⁹
8p23.1	rs11776767	10.683.929	3.00x10 ⁻¹¹
8p23.1	rs7819412	11.045.161	3.00x10 ⁻⁰⁸
8p22	rs1495741	18.272.881	4.00x10 ⁻¹⁴
8p21.3	rs325	19.819.328	8.00x10 ⁻²⁶
8p21.3	rs326	19.819.439	5.00x10 ⁻¹²
8p21.3	rs328	19.819.724	2.00x10 ⁻²⁸
8p21.3	rs331	19.820.405	1.00x10 ⁻¹⁷
8p21.3	rs13702	19.824.492	1.00x10 ⁻¹⁶
8p21.3	rs1059611	19.824.563	1.00x10 ⁻²⁰
8p21.3	rs10105606	19.827.848	4.00x10 ⁻²⁶

Locus	lead SNP	Position	p-value
8p21.3	rs10096633	19.830.921	2.00x10 ⁻¹⁸
8p21.3	rs17482753	19.832.646	3.00x10 ⁻¹¹
8p21.3	rs12678919	19.844.222	2.00x10 ⁻¹⁹⁹
8p21.3	rs10503669	19.847.690	8.00x10 ⁻⁴³
8p21.3	rs17410962	19.848.080	7.35x10 ⁻⁰⁹
8p21.3	rs79236614	19.860.460	3.79x10 ⁻⁰⁸
8p21.3	rs2083637	19.865.175	5.50x10 ⁻¹⁸
8p21.3	rs115849089	19.912.370	4.00x10 ⁻¹⁵
8p21.3	rs9644568	19.928.582	4.00x10 ⁻¹¹
8q11.23	rs10102164	55.421.614	4.00x10 ⁻¹¹
8q12.1	rs2081687	59.388.565	4.00x10 ⁻⁰⁹
8q23.3	rs2293889	116.599.199	4.00x10 ⁻¹⁷
8q24.13	rs2001945	126.477.978	1.00x10 ⁻²⁰
8q24.13	rs6982636	126.479.315	7.00x10 ⁻¹²
8q24.13	rs2954026	126.484.526	8.00x10 ⁻⁰⁹
8q24.13	rs17321515	126.486.409	4.00x10 ⁻¹⁷
8q24.13	rs2954029	126.490.972	1.00x10 ⁻¹⁰⁷
8q24.13	rs4360309	126.523.523	7.00x10 ⁻⁰⁹
8q24.3	rs11136341	145.043.543	4.00x10 ⁻¹³
9p24.2	rs3780181	2.640.759	2.00x10 ⁻⁰⁹
9p22.3	rs471364	15.289.578	3.00x10 ⁻¹⁰
9p22.3	rs643531	15.296.034	4.10x10 ⁻⁰⁸
9p22.3	rs581080	15.305.378	1.00x10 ⁻¹⁹
9q31.1	rs9282541	107.620.835	6.00x10 ⁻²⁶
9q31.1	rs4149268	107.647.220	1.00x10 ⁻¹⁰
9q31.1	rs3890182	107.647.655	3.00x10 ⁻¹⁰
9q31.1	rs12686004	107.653.426	2.00x10 ⁻¹⁸
9q31.1	rs3905000	107.657.070	9.00x10 ⁻¹³
9q31.1	rs1883025	107.664.301	2.00x10 ⁻⁶⁵
9q31.1	rs2575876	107.665.739	2.00x10 ⁻¹¹
9q34.2	rs657152	136.139.265	9.00x10 ⁻¹³
9q34.2	rs507666	136.149.399	2.00x10 ⁻¹¹
9q34.2	rs635634	136.155.000	8.00x10 ⁻²²
9q34.2	rs9411489	136.155.000	2.00x10 ⁻⁴¹
10p15.1	rs1832007	5.254.847	2.00x10 ⁻¹²
10q11.22	rs970548	46.013.277	2.00x10 ⁻¹⁰
10q21.3	rs10761731	65.027.610	3.00x10 ⁻¹²
10q23.33	rs2068888	94.839.642	2.00x10 ⁻¹¹
10q25.2	rs2255141	113.933.886	1.00x10 ⁻¹³
10q26.2	rs9733352	128.777.390	4.00x10 ⁻⁰⁸

Locus	lead SNP	Position	p-value
11p15.4	rs2923084	10.388.782	5.00x10 ⁻⁰⁸
11p11.2	rs3136441	46.743.247	7.00x10 ⁻²⁹
11p11.2	rs7120118	47.286.290	4.00x10 ⁻⁰⁸
11p11.2	rs7395662	48.518.893	6.00x10 ⁻¹¹
11q11	rs11246602	51.512.090	2.00x10 ⁻¹⁰
11q12.2	rs174546	61.569.830	2.00x10 ⁻³⁹
11q12.2	rs174547	61.570.783	8.00x10 ⁻²⁶²
11q12.2	rs174548	61.571.348	5.00x10 ⁻¹⁴
11q12.2	rs174570	61.597.212	4.00x10 ⁻¹³
11q13.1	rs12801636	65.391.317	3.00x10 ⁻⁰⁸
11q13.3	rs17610395	68.562.328	8.00x10 ⁻¹²
11q13.5	rs499974	75.455.021	1.00x10 ⁻⁰⁸
11q23.3	rs4938303	116.584.987	4.00x10 ⁻²¹
11q23.3	rs7350481	116.586.283	1.00x10 ⁻⁴⁹
11q23.3	rs12272004	116.603.724	5.00x10 ⁻¹³
11q23.3	rs1558861	116.607.437	2.00x10 ⁻²⁶
11q23.3	rs11216126	116.617.240	3.00x10 ⁻³⁴
11q23.3	rs28927680	116.619.073	2.00x10 ⁻¹⁷
11q23.3	rs10790162	116.639.104	2.80x10 ⁻¹⁵
11q23.3	rs2160669	116.647.607	3.00x10 ⁻¹¹²
11q23.3	rs964184	116.648.917	7.00x10 ⁻²²⁵
11q23.3	rs12286037	116.652.207	1.00x10 ⁻²⁶
11q23.3	rs6589566	116.652.423	5.00x10 ⁻¹⁴
11q23.3	rs2075290	116.653.296	2.00x10 ⁻¹⁴
11q23.3	rs603446	116.654.435	2.00x10 ⁻⁸⁶
11q23.3	rs2266788	116.660.686	5.00x10 ⁻¹³
11q23.3	rs651821	116.662.579	2.00x10 ⁻⁵⁹
11q23.3	rs662799	116.663.707	6.00x10 ⁻⁰⁹
11q23.3	rs2075292	116.732.512	5.00x10 ⁻⁰⁸
11q23.3	rs139961185	116.807.343	3.01x10 ⁻⁰⁹
11q23.3	rs11216230	116.884.789	3.00x10 ⁻¹⁰
11q23.3	rs10892151	117.531.731	3.00x10 ⁻¹³
11q23.3	rs12787909	117.546.482	5.00x10 ⁻⁰⁹
11q24.1	rs7941030	122.522.375	1.00x10 ⁻¹⁴
11q24.2	rs11220462	126.243.952	1.00x10 ⁻¹⁵
12p12.2	rs7134375	20.473.758	1.00x10 ⁻⁰⁸
12q13.3	rs11613352	57.792.580	9.00x10 ⁻¹⁴
12q24.11	rs2338104	109.895.168	1.00x10 ⁻¹⁰
12q24.11	rs7134594	110.000.193	7.00x10 ⁻¹⁵
12q24.11	rs12229654	111.414.461	3.00x10 ⁻²³

Locus	lead SNP	Position	p-value
12q24.12	rs11065987	112.072.424	1.00x10 ⁻¹¹
12q24.13	rs2074356	112.645.401	7.00x10 ⁻³⁷
12q24.31	rs2650000	121.388.962	2.00x10 ⁻⁰⁸
12q24.31	rs1169288	121.416.650	1.00x10 ⁻¹⁵
12q24.31	rs4759375	123.796.238	8.00x10 ⁻⁰⁹
12q24.31	rs4765127	124.460.167	3.00x10 ⁻¹⁰
12q24.31	rs838880	125.261.593	6.00x10 ⁻³²
13q13.1	rs4942486	32.953.388	2.00x10 ⁻¹¹
14q12	rs8017377	24.883.887	3.00x10 ⁻¹⁵
14q32.33	rs4983559	105.277.209	1.00x10 ⁻⁰⁸
15q15.1	rs2412710	42.683.787	2.00x10 ⁻¹¹
15q15.3	rs2929282	44.245.931	2.00x10 ⁻¹¹
15q21.3	rs4775041	58.674.695	3.00x10 ⁻²⁰
15q21.3	rs10468017	58.678.512	8.00x10 ⁻²³
15q21.3	rs35853021	58.680.643	7.00x10 ⁻⁷⁶
15q21.3	rs2043085	58.680.954	2.00x10 ⁻⁰⁸
15q21.3	rs1532085	58.683.366	1.00x10 ⁻¹⁸⁸
15q21.3	rs16940212	58.694.020	1.00x10 ⁻²⁴
15q21.3	rs1077835	58.723.426	2.00x10 ⁻¹⁴
15q21.3	rs1800588	58.723.675	2.00x10 ⁻³²
15q21.3	rs8034802	58.724.792	1.00x10 ⁻⁰⁹
15q21.3	rs261334	58.726.744	4.90x10 ⁻²²
15q21.3	rs588136	58.730.498	2.00x10 ⁻¹²
15q22.2	rs2652834	63.396.867	4.00x10 ⁻¹¹
16p13.11	rs3198697	15.129.940	2.00x10 ⁻⁰⁸
16p13.11	rs11075253	15.148.646	4.98x10 ⁻¹⁵
16p11.2	rs11649653	30.918.487	3.35x10 ⁻⁰⁸
16q12.2	rs1121980	53.809.247	7.00x10 ⁻⁰⁹
16q13	rs9989419	56.985.139	1.00x10 ⁻³²
16q13	rs173539	56.988.044	4.00x10 ⁻⁷⁵
16q13	rs247617	56.990.716	1.00x10 ⁻⁴⁴
16q13	rs3764261	56.993.324	1.00x10 ⁻⁷⁶⁹
16q13	rs1800775	56.995.236	4.00x10 ⁻⁹³
16q13	rs1864163	56.997.233	7.00x10 ⁻³⁹
16q13	rs1532624	57.005.479	9.00x10 ⁻⁷⁹
16q13	rs7499892	57.006.590	1.00x10 ⁻²⁰
16q13	rs12708980	57.012.379	2.00x10 ⁻²⁸
16q13	rs5880	57.015.091	2.00x10 ⁻¹⁶
16q22.1	rs3729639	67.225.501	2.00x10 ⁻¹¹
16q22.1	rs2271293	67.902.070	8.00x10 ⁻¹⁶

Locus	lead SNP	Position	p-value
16q22.1	rs16942887	67.928.042	8.00x10 ⁻⁵⁴
16q22.1	rs255049	68.013.471	3.06x10 ⁻⁰⁸
16q22.2	rs2000999	72.108.093	4.00x10 ⁻⁴¹
16q23.2	rs2925979	81.534.790	1.00x10 ⁻¹⁹
17p13.1	rs314253	7.091.650	3.00x10 ⁻¹⁰
17q12	rs11869286	37.813.856	3.00x10 ⁻¹⁷
17q21.31	rs8077889	41.878.166	1.00x10 ⁻⁰⁸
17q21.32	rs7206971	45.425.115	4.00x10 ⁻⁰⁹
17q24.2	rs1801689	64.210.580	1.00x10 ⁻¹¹
17q24.2	rs4148008	66.875.294	1.00x10 ⁻¹²
17q25.3	rs4129767	76.403.984	2.00x10 ⁻¹¹
17q25.3	rs72925845	76.439.361	7.00x10 ⁻⁰⁹
18q11.2	rs9949617	20.879.217	2.00x10 ⁻⁰⁸
18q21.1	rs7240405	47.159.090	5.00x10 ⁻¹⁰
18q21.1	rs7228085	47.160.814	6.70x10 ⁻¹¹
18q21.1	rs7241918	47.160.953	1.00x10 ⁻⁴⁴
18q21.1	rs4939883	47.167.214	7.00x10 ⁻¹⁶
18q21.1	rs2156552	47.181.668	2.00x10 ⁻¹²
18q21.32	rs12967135	57.849.023	7.00x10 ⁻⁰⁹
19p13.2	rs7248104	7.224.431	5.00x10 ⁻¹⁰
19p13.2	rs7255436	8.433.196	2.00x10 ⁻⁰⁸
19p13.2	rs2967605	8.469.738	1.00x10 ⁻⁰⁸
19p13.2	rs11669133	11.092.139	1.00x10 ⁻⁰⁸
19p13.2	rs55791371	11.188.153	8.00x10 ⁻¹⁷
19p13.2	rs17249141	11.200.008	2.00x10 ⁻¹⁷
19p13.2	rs6511720	11.202.306	4.00x10 ⁻²⁶²
19p13.2	rs2228671	11.210.912	4.00x10 ⁻¹⁴
19p13.2	rs12979813	11.342.703	2.00x10 ⁻⁰⁹
19p13.2	rs737337	11.347.493	3.10x10 ⁻⁰⁹
19p13.2	rs2278426	11.350.488	3.44x10 ⁻⁰⁹
19p13.11	rs10401969	19.407.718	1.00x10 ⁻⁶⁹
19p13.11	rs16996148	19.658.472	3.00x10 ⁻⁰⁹
19p13.11	rs17216525	19.662.220	4.00x10 ⁻¹¹
19p13.11	rs2304130	19.789.528	4.00x10 ⁻⁰⁸
19q13.11	rs731839	33.899.065	3.00x10 ⁻⁰⁹
19q13.32	rs4803750	45.247.627	3.00x10 ⁻⁵¹
19q13.32	rs519113	45.376.284	8.00x10 ⁻¹¹
19q13.32	rs157580	45.395.266	2.00x10 ⁻¹⁹
19q13.32	rs1160985	45.403.412	2.00x10 ⁻²¹
19q13.32	rs7412	45.412.079	3.00x10 ⁻⁵⁸

Locus	lead SNP	Position	p-value
19q13.32	rs439401	45.414.451	1.00x10 ⁻³⁰
19q13.32	rs445925	45.415.640	2.00x10 ⁻¹⁴
19q13.32	rs12721054	45.422.587	3.00x10 ⁻¹⁹
19q13.32	rs4420638	45.422.946	2.00x10 ⁻¹⁷⁸
19q13.41	rs17695224	52.324.216	2.00x10 ⁻¹³
19q13.42	rs386000	54.792.761	3.00x10 ⁻²³
20p12.1	rs364585	12.962.718	4.00x10 ⁻¹⁰
20p12.1	rs2328223	17.845.921	6.00x10 ⁻⁰⁹
20q12	rs2902940	39.091.487	2.00x10 ⁻¹¹
20q12	rs6102059	39.228.784	4.00x10 ⁻⁰⁹
20q12	rs6029526	39.672.618	3.00x10 ⁻¹⁹
20q13.12	rs1800961	43.042.364	2.00x10 ⁻³⁴
20q13.12	rs6065904	44.534.651	4.00x10 ⁻⁴⁰
20q13.12	rs6065906	44.554.015	1.00x10 ⁻⁴⁵
20q13.12	rs7679	44.576.502	7.00x10 ⁻¹¹
22q11.21	rs181362	21.932.068	4.00x10 ⁻¹⁸
22q12.2	rs5763662	30.378.703	1.00x10 ⁻⁰⁸
22q13.1	rs5756931	38.546.033	3.00x10 ⁻⁰⁸
22q13.31	rs4253772	46.627.603	1.00x10 ⁻⁰⁸

Supplementary Table 12 | Blood pressure SNPs

Locus	lead SNP	Position	p-value
1p36.33	rs139385870	1.685.922	1.30x10 ⁻¹²
1p36.22	rs9662255	9.441.949	1.90x10 ⁻¹⁰
1p36.22	rs880315	10.796.866	3.00x10 ⁻¹⁰
1p36.22	rs17367504	11.862.778	2.00x10 ⁻¹⁶
1p36.21	rs3820068	15.798.197	1.10x10 ⁻¹²
1p36.11	rs6686889	25.030.470	3.60x10 ⁻⁰⁹
1p34.3	rs4360494	38.455.891	3.70x10 ⁻¹⁶
1p32.2	rs112557609	56.576.924	6.80x10 ⁻¹²
1p32.1	rs3889199	59.653.742	1.80x10 ⁻²⁴
1p22.2	rs10922502	89.360.158	2.20x10 ⁻¹⁵
1p13.2	rs10745332	113.189.053	2.70x10 ⁻⁰⁹
1p13.2	rs17030613	113.190.807	1.00x10 ⁻⁰⁸
1p13.2	rs2932538	113.216.543	1.00x10 ⁻⁰⁹
1q24.3	rs12405515	172.357.441	1.40x10 ⁻⁰⁹
1q41	rs12408022	217.718.789	2.40x10 ⁻¹⁰
1q42.13	rs10916082	227.252.626	8.40x10 ⁻⁰⁹
1q42.13	rs2760061	228.191.075	2.10x10 ⁻¹⁶
1q43	rs953492	243.471.192	7.40x10 ⁻¹⁶

Locus	lead SNP	Position	p-value
2p24.1	rs2289081	20.881.840	5.50x10 ⁻¹²
2p23.3	rs55701159	25.139.596	7.20x10 ⁻¹¹
2p23.2	rs7562	28.635.740	1.90x10 ⁻⁰⁸
2p22.3	rs9308945	34.284.852	3.00x10 ⁻¹⁰
2p22.3	rs6711736	34.284.984	3.00x10 ⁻¹⁰
2p22.3	rs6729869	34.285.194	3.00x10 ⁻¹⁰
2p22.3	rs10495809	34.305.257	3.00x10 ⁻¹⁰
2p22.2	rs13420463	37.517.566	7.00x10 ⁻¹¹
2p22.1	rs4952611	40.567.743	4.00x10 ⁻⁰⁸
2p21	rs76326501	43.167.878	3.60x10 ⁻¹⁸
2p21	rs11690961	46.363.336	3.90x10 ⁻¹²
2p14	rs74181299	65.283.972	9.60x10 ⁻¹³
2p12	rs10496288	83.292.565	2.00x10 ⁻⁰⁹
2p12	rs10496289	83.293.380	2.00x10 ⁻⁰⁹
2p11.2	rs11689667	85.491.365	1.70x10 ⁻⁰⁸
2q11.2	rs2579519	96.675.166	4.80x10 ⁻¹²
2q21.2	rs13420028	133.188.106	1.00x10 ⁻¹⁰
2q21.2	rs10188442	133.189.239	1.00x10 ⁻¹⁰

Locus	lead SNP	Position	p-value
2q22.3	rs1438896	145.646.072	2.00x10 ⁻¹⁵
2q24.3	rs16849225	164.906.820	4.00x10 ⁻¹¹
2q24.3	rs13002573	164.915.208	2.00x10 ⁻⁰⁸
2q24.3	rs1446468	164.963.486	6.00x10 ⁻¹²
2q31.1	rs2268365	170.094.169	5.00x10 ⁻⁰⁸
2q31.2	rs79146658	179.786.068	2.40x10 ⁻¹⁰
2q32.2	rs7592578	191.439.591	9.50x10 ⁻¹²
2q33.3	rs55780018	208.526.140	5.90x10 ⁻¹⁶
2q35	rs1250259	216.300.482	8.70x10 ⁻¹⁹
2q35	rs1063281	218.668.732	1.30x10 ⁻¹²
3p24.1	rs13082711	27.537.909	4.00x10 ⁻⁰⁹
3p24.1	rs820430	27.548.900	1.00x10 ⁻¹²
3p22.1	rs9815354	41.912.651	2.54x10 ⁻⁰⁹
3p22.1	rs1717027	41.987.920	5.00x10 ⁻¹³
3p21.31	rs319690	47.927.484	2.70x10 ⁻⁰⁸
3p21.31	rs36022378	49.913.705	4.70x10 ⁻⁰⁹
3p21.31	rs743757	50.476.378	2.40x10 ⁻¹⁰
3p21.1	rs9810888	53.635.595	4.00x10 ⁻¹²
3p14.3	rs9827472	56.726.646	4.30x10 ⁻¹⁰
3q21.3	rs62270945	128.201.889	1.80x10 ⁻⁰⁹
3q22.2	rs9859176	134.000.025	1.30x10 ⁻¹¹
3q25.2	rs143112823	154.707.967	1.40x10 ⁻¹⁴
3q26.1	rs16833934	163.737.250	1.00x10 ⁻⁰⁸
3q26.2	rs419076	169.100.886	2.00x10 ⁻¹³
3q27.2	rs12374077	185.317.674	9.20x10 ⁻⁰⁹
4q12	rs871606	54.799.245	1.00x10 ⁻⁰⁸
4q21.21	rs1902859	81.157.703	2.00x10 ⁻²²
4q21.21	rs1458038	81.164.723	9.00x10 ⁻²⁵
4q21.21	rs11099098	81.169.912	2.00x10 ⁻¹¹
4q21.21	rs16998073	81.184.341	1.00x10 ⁻²¹
4q24	rs13107325	103.188.709	2.00x10 ⁻¹⁷
4q25	rs6825911	111.381.638	9.00x10 ⁻⁰⁹
4q26	rs66887589	120.509.279	3.40x10 ⁻¹⁵
4q32.1	rs13139571	156.645.513	2.00x10 ⁻¹⁰
4q32.3	rs1566497	169.717.148	1.90x10 ⁻¹³
4q34.1	rs17059668	174.584.663	2.80x10 ⁻⁰⁸
5p13.3	rs1173766	32.804.528	2.00x10 ⁻⁰⁸
5p13.3	rs1173771	32.815.028	2.00x10 ⁻¹⁶
5p13.2	rs7735940	36.423.931	5.00x10 ⁻¹³
5p13.2	rs12522034	36.425.593	5.00x10 ⁻¹³

Locus	lead SNP	Position	p-value
5q13.3	rs10078021	75.038.431	1.30x10 ⁻⁰⁸
5q14.1	rs10057188	77.837.789	6.70x10 ⁻¹¹
5q14.3	rs10059921	87.514.515	4.00x10 ⁻⁰⁹
5q23.3	rs6595838	127.868.199	7.60x10 ⁻¹²
5q33.3	rs9313772	157.804.457	1.00x10 ⁻¹¹
5q33.3	rs11953630	157.845.402	4.00x10 ⁻¹³
5q35.2	rs72812846	173.377.636	2.20x10 ⁻¹¹
6p22.3	rs6911827	22.130.601	2.00x10 ⁻¹⁰
6p22.2	rs1799945	26.091.179	2.00x10 ⁻¹⁵
6p22.2	rs198846	26.107.463	2.00x10 ⁻¹²
6p21.33	rs805303	31.616.366	2.00x10 ⁻¹¹
6p21.33	rs2021783	32.044.851	2.00x10 ⁻¹²
6p12.3	rs78648104	50.683.009	1.30x10 ⁻⁰⁸
6p12.2	rs13205180	51.832.494	7.00x10 ⁻¹⁰
6q14.1	rs3798440	76.556.619	3.00x10 ⁻¹⁰
6q14.1	rs9350602	76.560.498	3.00x10 ⁻¹⁰
6q22.31	rs9372498	118.572.486	1.80x10 ⁻¹¹
6q22.31	rs11154027	121.781.390	1.10x10 ⁻¹⁰
6q22.32	rs13209747	127.115.454	2.00x10 ⁻¹¹
6q25.1	rs17080102	151.004.770	2.00x10 ⁻¹¹
6q25.1	rs36083386	152.397.913	1.50x10 ⁻¹⁸
6q25.3	rs449789	159.699.125	2.40x10 ⁻¹⁵
6q27	rs147212971	166.178.451	1.60x10 ⁻⁰⁹
6q27	rs1322639	169.587.103	4.80x10 ⁻¹⁷
7p15.2	rs17428471	27.337.867	2.00x10 ⁻¹²
7p14.1	rs76206723	40.447.971	7.40x10 ⁻¹²
7q22.1	rs7801190	100.458.093	3.40x10 ⁻⁰⁸
7q22.3	rs17477177	106.411.858	2.00x10 ⁻¹³
7q32.3	rs13238550	131.059.056	1.90x10 ⁻¹²
7q34	rs1011018	139.463.264	1.50x10 ⁻⁰⁸
8p11.21	rs2978456	42.324.765	1.20x10 ⁻⁰⁸
8q22.3	rs2978098	101.676.675	1.50x10 ⁻⁰⁹
8q24.12	rs2469997	120.353.267	3.00x10 ⁻¹⁶
8q24.12	rs6469823	120.353.984	3.00x10 ⁻¹⁶
8q24.12	rs2071518	120.435.812	4.00x10 ⁻⁰⁹
8q24.22	rs7827545	135.566.567	2.00x10 ⁻⁴⁴
8q24.22	rs1372662	135.567.046	2.00x10 ⁻⁴⁴
8q24.22	rs894344	135.612.745	3.20x10 ⁻⁰⁸
8q24.3	rs4454254	141.060.027	5.10x10 ⁻¹⁶
8q24.3	rs62524579	144.060.955	3.80x10 ⁻⁰⁹

Locus	lead SNP	Position	p-value
9p21.3	rs4364717	21.801.530	1.30x10 ⁻¹⁰
9q33.3	rs72765298	127.900.996	2.70x10 ⁻¹⁴
10p12.33	rs4373814	18.419.972	5.00x10 ⁻¹¹
10p12.31	rs11014166	18.708.798	1.00x10 ⁻⁰⁸
10p12.31	rs12258967	18.727.959	2.00x10 ⁻¹⁶
10p11.23	rs9337951	30.317.073	2.50x10 ⁻¹⁵
10p11.22	rs10826995	32.082.658	1.10x10 ⁻⁰⁹
10q21.2	rs4590817	63.467.553	2.00x10 ⁻¹⁸
10q21.2	rs1530440	63.524.591	1.00x10 ⁻⁰⁹
10q23.33	rs9663362	95.895.177	5.00x10 ⁻⁰⁹
10q23.33	rs932764	95.895.940	9.40x10 ⁻⁰⁹
10q24.31	rs112184198	102.604.514	3.60x10 ⁻¹⁸
10q24.32	rs1004467	104.594.507	1.00x10 ⁻¹⁰
10q24.32	rs4409766	104.616.663	6.00x10 ⁻¹⁷
10q24.32	rs12416687	104.629.011	4.00x10 ⁻⁰⁹
10q24.32	rs11191548	104.846.178	7.00x10 ⁻²⁴
10q24.33	rs11191593	104.939.215	1.00x10 ⁻¹⁵
10q25.3	rs2782980	115.781.527	2.00x10 ⁻⁰⁹
11p15.4	rs7129220	10.350.538	3.00x10 ⁻¹²
11p15.2	rs1401454	16.250.183	5.00x10 ⁻¹⁰
11p15.2	rs4757391	16.302.939	5.00x10 ⁻⁰⁹
11p15.2	rs381815	16.902.268	2.00x10 ⁻⁰⁹
11p14.1	rs11030119	27.728.102	2.90x10 ⁻⁰⁸
11p11.2	rs11442819	45.208.142	7.10x10 ⁻⁰⁹
11q13.3	rs67330701	69.079.707	2.10x10 ⁻¹²
11q14.3	rs2289125	89.224.453	9.10x10 ⁻²²
11q22.1	rs633185	100.593.538	1.00x10 ⁻¹⁷
11q23.3	rs8258	117.283.676	2.90x10 ⁻¹³
11q24.3	rs11222084	130.273.230	2.00x10 ⁻¹¹
12q21.33	rs4842666	89.941.549	4.00x10 ⁻⁰⁸
12q21.33	rs2681472	90.008.959	2.00x10 ⁻¹¹
12q21.33	rs2681492	90.013.089	4.00x10 ⁻¹¹
12q21.33	rs17249754	90.060.586	8.00x10 ⁻²⁰
12q22	rs139236208	94.880.742	1.60x10 ⁻¹⁰
12q24.12	rs3184504	111.884.608	4.00x10 ⁻²⁵
12q24.12	rs653178	112.007.756	7.00x10 ⁻²⁰
12q24.13	rs11066280	112.817.783	1.00x10 ⁻³⁵
12q24.21	rs2384550	115.352.731	4.00x10 ⁻⁰⁸
12q24.21	rs35444	115.552.437	1.00x10 ⁻¹⁰
12q24.21	rs11067763	116.198.341	2.00x10 ⁻¹⁸

Locus	lead SNP	Position	p-value
13q22.3	rs9318552	79.025.787	5.00x10 ⁻⁰⁸
13q34	rs9549328	113.636.156	1.50x10 ⁻⁰⁸
14q22.1	rs9888615	53.377.540	3.50x10 ⁻¹⁰
14q23.2	rs8016306	63.928.546	3.70x10 ⁻⁰⁹
14q32.2	rs9323988	98.587.630	4.10x10 ⁻¹¹
15q22.31	rs7178615	66.869.072	2.60x10 ⁻¹⁰
15q24.1	rs1378942	75.077.367	3.00x10 ⁻²⁶
15q24.1	rs6495122	75.125.645	2.00x10 ⁻¹⁰
15q25.1	rs62012628	79.070.000	5.10x10 ⁻¹²
15q25.1	rs35199222	81.013.037	5.20x10 ⁻¹²
15q26.1	rs2521501	91.437.388	5.00x10 ⁻¹⁹
15q26.2	rs12906962	95.312.071	5.60x10 ⁻¹⁴
16p13.3	rs12921187	4.943.019	2.50x10 ⁻¹⁰
16p12.3	rs13333226	20.365.654	4.00x10 ⁻¹¹
16p11.2	rs72799341	30.936.743	5.80x10 ⁻⁰⁹
16q22.1	rs12149862	69.497.203	4.00x10 ⁻⁰⁹
16q22.1	rs117006983	70.755.610	4.10x10 ⁻¹²
16q23.1	rs11643209	75.331.044	1.80x10 ⁻¹²
16q23.2	rs8059962	81.574.197	1.30x10 ⁻⁰⁹
16q23.3	rs7500448	83.045.790	1.10x10 ⁻¹⁹
16q23.3	rs3096277	83.764.204	1.00x10 ⁻⁰⁹
17p13.3	rs12941318	1.333.598	2.50x10 ⁻⁰⁸
17p13.1	rs7226020	6.473.828	2.30x10 ⁻¹⁴
17p13.1	rs78378222	7.571.752	1.80x10 ⁻¹⁰
17q21.2	rs79089478	40.317.241	3.10x10 ⁻⁰⁹
17q21.31	rs62080325	42.060.631	4.00x10 ⁻⁰⁸
17q21.31	rs12946454	43.208.121	1.00x10 ⁻⁰⁸
17q21.32	rs17608766	45.013.271	6.00x10 ⁻¹⁵
17q21.33	rs16948048	47.440.466	5.00x10 ⁻⁰⁹
17q23.2	rs740698	60.767.151	3.10x10 ⁻¹²
17q23.3	rs4308	61.559.625	6.80x10 ⁻¹⁴
17q25.1	rs2467099	73.949.045	3.30x10 ⁻⁰⁸
18q12.3	rs7236548	43.097.750	2.00x10 ⁻¹⁸
18q21.1	rs745821	48.142.854	1.40x10 ⁻⁰⁹
18q23	rs4573996	76.494.592	8.77x10 ⁻⁰⁹
19q12	rs62104477	30.294.991	1.20x10 ⁻⁰⁹
20p12.3	rs6108168	8.626.271	1.10x10 ⁻¹¹
20p12.2	rs1887320	10.965.998	1.00x10 ⁻⁰⁸
20p12.2	rs1327235	10.969.030	1.00x10 ⁻¹⁵
20p12.1	rs200752	15.599.531	6.65x10 ⁻⁰⁹

Locus	lead SNP	Position	p-value
20p12.1	rs200759	15.606.421	6.65x10 ⁻⁰⁹
20p11.23	rs6081613	19.465.907	1.60x10 ⁻¹³
20q13.32	rs6015450	57.751.117	4.00x10 ⁻²³
22q11.21	rs12628032	19.967.980	5.50x10 ⁻¹²
22q12.1	rs133980	28.022.728	1.00x10 ⁻⁰⁸
22q13.2	rs73161324	42.038.786	2.80x10 ⁻¹¹

Supplementary Table 13 | Diabetes mellitus SNPs

Locus	lead SNP	Position	p-value
1p32.3	rs17106184	50.909.985	4.00x10 ⁻⁰⁹
1p13.2	rs6679677	114.303.808	1.00x10 ⁻⁴⁰
1p13.2	rs2476601	114.377.568	9.00x10 ⁻⁸⁵
1p12	rs10923931	120.517.959	4.00x10 ⁻⁰⁸
1q32.1	rs3024505	206.939.904	2.00x10 ⁻⁰⁹
1q32.3	rs340874	214.159.256	7.00x10 ⁻¹²
2p23.3	rs478222	25.301.755	4.00x10 ⁻⁰⁹
2p23.3	rs1371614	27.152.874	2.00x10 ⁻¹²
2p23.3	rs780094	27.741.237	6.00x10 ⁻³⁸
2p23.2	rs3736594	27.995.781	1.00x10 ⁻¹⁵
2p21	rs7578597	43.732.823	1.00x10 ⁻⁰⁹
2p21	rs895636	45.188.353	2.52x10 ⁻¹³
2p16.1	rs243021	60.584.819	2.90x10 ⁻¹⁵
2q23.3	rs7560163	151.637.936	7.00x10 ⁻⁰⁹
2q24.2	rs7593730	161.171.454	4.00x10 ⁻⁰⁸
2q24.2	rs1990760	163.124.051	2.00x10 ⁻¹¹
2q24.3	rs3923113	165.501.849	1.00x10 ⁻⁰⁸
2q24.3	rs13387347	169.754.846	2.00x10 ⁻³⁶
2q31.1	rs560887	169.763.148	9.00x10 ⁻²⁰³
2q31.1	rs733331	173.546.313	6.98x10 ⁻¹¹
2q33.2	rs3087243	204.738.919	1.00x10 ⁻¹⁵
2q36.3	rs7578326	227.020.653	5.00x10 ⁻²⁰
2q36.3	rs2943640	227.093.585	7.20x10 ⁻⁰⁹
2q36.3	rs2943641	227.093.745	9.00x10 ⁻¹²
3p25.2	rs1801282	12.393.125	6.00x10 ⁻¹⁰
3p24.3	rs7612463	23.336.450	6.70x10 ⁻⁰⁹
3p14.1	rs831571	64.048.297	8.00x10 ⁻¹¹
3p14.1	rs4607103	64.711.904	1.00x10 ⁻⁰⁸
3q21.1	rs11708067	123.065.778	7.00x10 ⁻²²
3q21.1	rs11717195	123.082.398	2.00x10 ⁻⁰⁸
3q26.2	rs11920090	170.717.521	8.00x10 ⁻¹³

Locus	lead SNP	Position	p-value
3q27.2	rs4402960	185.511.687	1.00x10 ⁻¹⁷
3q27.2	rs1470579	185.529.080	2.00x10 ⁻¹⁹
3q27.2	rs6769511	185.530.290	1.00x10 ⁻⁰⁹
3q27.3	rs16861329	186.666.461	3.00x10 ⁻⁰⁸
3q27.3	rs6808574	187.740.523	5.80x10 ⁻⁰⁹
4p16.3	rs7656416	1.254.535	1.00x10 ⁻⁰⁸
4p16.3	rs6815464	1.309.901	2.00x10 ⁻²⁰
4p16.1	rs4689388	6.270.056	1.00x10 ⁻⁰⁸
4p16.1	rs4458523	6.289.986	2.00x10 ⁻⁰⁹
4p16.1	rs1801214	6.303.022	3.16x10 ⁻⁰⁸
4p15.2	rs10517086	26.085.511	5.00x10 ⁻¹⁰
4q27	rs4505848	123.132.492	5.00x10 ⁻¹³
4q31.3	rs6813195	153.520.475	4.00x10 ⁻¹⁴
5q11.2	rs702634	53.271.420	6.90x10 ⁻⁰⁹
5q13.3	rs4457053	76.424.949	3.00x10 ⁻¹²
5q15	rs13179048	95.542.726	2.00x10 ⁻¹⁰
6p24.3	rs9502570	7.258.617	1.00x10 ⁻⁰⁹
6p22.3	rs4712523	20.657.564	7.00x10 ⁻²⁰
6p22.3	rs4712524	20.657.865	3.00x10 ⁻¹⁰
6p22.3	rs10946398	20.661.034	1.00x10 ⁻⁰⁸
6p22.3	rs7754840	20.661.250	2.00x10 ⁻¹³
6p22.3	rs7756992	20.679.709	2.00x10 ⁻²⁶
6p22.3	rs9356744	20.685.486	9.00x10 ⁻¹⁰
6p22.3	rs7766070	20.686.573	6.00x10 ⁻¹¹
6p22.3	rs10440833	20.688.121	2.00x10 ⁻²²
6p22.3	rs6931514	20.703.952	1.00x10 ⁻¹¹
6p21.33	rs3132524	31.136.714	4.00x10 ⁻⁰⁹
6p21.33	rs2244020	31.347.451	6.57x10 ⁻⁰⁹
6p21.32	rs9268645	32.408.527	1.00x10 ⁻¹⁰⁰
6p21.32	rs9272346	32.604.372	5.00x10 ⁻¹³⁴
6p21.32	rs2647044	32.667.910	1.00x10 ⁻¹⁶

Locus	lead SNP	Position	p-value
6p21.2	rs9470794	38.106.844	2.00x10 ⁻¹⁰
6p21.2	rs1535500	39.284.050	2.00x10 ⁻⁰⁸
6q13	rs1048886	71.289.189	3.48x10 ⁻⁰⁸
6q15	rs11755527	90.958.231	5.00x10 ⁻¹²
6q22.32	rs9388489	126.698.719	4.00x10 ⁻¹³
6q27	rs924043	170.379.025	8.00x10 ⁻⁰⁹
7p21.2	rs2191349	15.064.309	3.00x10 ⁻²⁸
7p21.2	rs1974620	15.065.467	3.00x10 ⁻¹¹
7p15.2	rs7804356	26.891.665	5.00x10 ⁻⁰⁹
7p15.1	rs864745	28.180.556	5.00x10 ⁻¹⁴
7p15.1	rs849134	28.196.222	2.80x10 ⁻⁰⁹
7p15.1	rs849135	28.196.413	2.00x10 ⁻⁰⁹
7p13	rs730497	44.223.721	8.00x10 ⁻²⁷
7p13	rs4607517	44.235.668	7.00x10 ⁻⁹²
7p12.1	rs4948088	51.027.194	4.00x10 ⁻⁰⁸
7q32.1	rs6467136	127.164.958	4.96x10 ⁻¹¹
7q32.1	rs10229583	127.246.903	2.00x10 ⁻¹⁰
7q32.1	rs791595	127.862.802	2.55x10 ⁻¹³
7q32.2	rs972283	130.466.854	2.00x10 ⁻¹⁰
8p23.1	rs4841132	9.183.596	8.00x10 ⁻⁰⁹
8p11.21	rs2722425	40.484.239	2.00x10 ⁻⁰⁸
8p11.21	rs515071	41.519.462	1.00x10 ⁻⁰⁸
8q22.1	rs896854	95.960.511	1.00x10 ⁻⁰⁹
8q24.11	rs13266634	118.184.783	2.00x10 ⁻¹⁴
8q24.11	rs3802177	118.185.025	2.00x10 ⁻¹⁸
8q24.11	rs11558471	118.185.733	3.00x10 ⁻²⁰
9p24.3	rs10815355	622.523	1.00x10 ⁻⁰⁹
9p24.2	rs7041847	4.287.466	2.00x10 ⁻¹⁴
9p24.2	rs7034200	4.289.050	1.00x10 ⁻¹³
9p24.2	rs4237150	4.290.085	4.00x10 ⁻⁰⁹
9p24.2	rs7020673	4.291.747	5.00x10 ⁻¹²
9p24.2	rs10814916	4.293.150	6.00x10 ⁻¹²
9p24.1	rs17584499	8.879.118	9.00x10 ⁻¹⁰
9p21.3	rs2383208	22.132.076	2.00x10 ⁻²⁹
9p21.3	rs10965250	22.133.284	1.00x10 ⁻¹⁰
9p21.3	rs10811661	22.134.094	1.00x10 ⁻²⁷
9p21.3	rs1333051	22.136.489	6.00x10 ⁻¹⁰
9p21.3	rs7018475	22.137.685	2.50x10 ⁻⁰⁸
9q21.31	rs17791513	81.905.590	3.20x10 ⁻⁰⁸
9q21.31	rs13292136	81.952.128	2.80x10 ⁻⁰⁸

Locus	lead SNP	Position	p-value
9q34.3	rs11787792	139.252.148	2.00x10 ⁻¹⁰
10p15.1	rs61839660	6.094.697	5.00x10 ⁻⁰⁹
10p15.1	rs12251307	6.123.495	1.00x10 ⁻¹³
10p15.1	rs947474	6.390.450	4.00x10 ⁻⁰⁹
10p13	rs11257655	12.307.894	6.56x10 ⁻⁰⁹
10p13	rs10906115	12.314.997	1.00x10 ⁻⁰⁸
10p13	rs12779790	12.328.010	1.00x10 ⁻¹⁰
10q22.1	rs1802295	70.931.474	4.00x10 ⁻⁰⁸
10q22.3	rs12571751	80.942.631	2.00x10 ⁻¹⁰
10q23.31	rs10509540	90.023.033	1.00x10 ⁻²⁸
10q23.33	rs1111875	94.462.882	3.00x10 ⁻¹⁹
10q23.33	rs5015480	94.465.559	1.00x10 ⁻¹⁵
10q25.2	rs10885122	113.042.093	3.00x10 ⁻¹⁶
10q25.2	rs7901695	114.754.088	1.00x10 ⁻⁴⁸
10q25.2	rs4506565	114.756.041	5.00x10 ⁻¹²
10q25.2	rs7903146	114.758.349	4.00x10 ⁻⁹⁴
10q26.11	rs10886471	121.149.403	7.10x10 ⁻⁰⁹
11p15.5	rs1004446	2.170.143	4.00x10 ⁻⁰⁹
11p15.5	rs3842770	2.178.670	3.00x10 ⁻⁰⁸
11p15.5	rs7111341	2.213.166	4.00x10 ⁻⁴⁸
11p15.5	rs231362	2.691.471	2.80x10 ⁻¹³
11p15.5	rs231356	2.705.343	4.00x10 ⁻⁰⁸
11p15.4	rs8181588	2.831.541	5.00x10 ⁻⁰⁹
11p15.4	rs2237892	2.839.751	2.00x10 ⁻⁴²
11p15.4	rs163182	2.844.216	2.00x10 ⁻¹⁷
11p15.4	rs163184	2.847.069	2.00x10 ⁻¹⁴
11p15.4	rs2283228	2.849.530	5.00x10 ⁻¹³
11p15.4	rs2237895	2.857.194	1.00x10 ⁻⁰⁹
11p15.4	rs2237897	2.858.546	1.00x10 ⁻¹⁶
11p15.1	rs5215	17.408.630	3.00x10 ⁻¹¹
11p15.1	rs5219	17.409.572	7.00x10 ⁻¹¹
11p11.2	rs11605924	45.873.091	4.00x10 ⁻¹⁵
11p11.2	rs7944584	47.336.320	2.00x10 ⁻¹⁸
11p11.2	rs1483121	48.333.360	2.00x10 ⁻⁰⁸
11q12.2	rs174550	61.571.478	2.00x10 ⁻¹⁵
11q13.4	rs11603334	72.432.985	2.00x10 ⁻¹⁴
11q13.4	rs1552224	72.433.098	1.00x10 ⁻²²
11q14.3	rs3847554	92.668.826	2.00x10 ⁻¹¹
11q14.3	rs1387153	92.673.828	2.00x10 ⁻³⁶
11q14.3	rs2166706	92.691.532	2.00x10 ⁻⁰⁹

Locus	lead SNP	Position	p-value
11q14.3	rs10830963	92.708.710	6.00x10 ⁻¹⁷⁵
12p13.31	rs3764021	9.833.628	5.00x10 ⁻⁰⁸
12p13.31	rs4763879	9.910.164	2.00x10 ⁻¹¹
12q13.2	rs1701704	56.412.487	9.00x10 ⁻¹⁰
12q13.2	rs11171739	56.470.625	1.00x10 ⁻¹¹
12q13.2	rs2292239	56.482.180	2.00x10 ⁻²⁵
12q14.3	rs1531343	66.174.894	4.00x10 ⁻⁰⁹
12q14.3	rs2261181	66.212.318	4.00x10 ⁻⁰⁸
12q14.3	rs343092	66.250.940	3.00x10 ⁻¹²
12q21.1	rs7961581	71.663.102	1.00x10 ⁻⁰⁹
12q23.2	rs35767	102.875.569	2.00x10 ⁻⁰⁹
12q24.11	rs1265564	111.708.458	1.00x10 ⁻¹⁶
12q24.12	rs3184504	111.884.608	3.00x10 ⁻²⁷
12q24.13	rs17696736	112.486.818	6.00x10 ⁻¹⁸
12q24.31	rs7305618	121.402.932	2.00x10 ⁻⁰⁸
12q24.31	rs7957197	121.460.686	2.00x10 ⁻⁰⁸
12q24.31	rs1727313	123.640.853	1.00x10 ⁻⁰⁸
13q12.12	rs9552911	23.864.657	2.00x10 ⁻⁰⁸
13q12.2	rs2293941	28.491.198	5.00x10 ⁻¹⁰
13q22.2	rs539514	76.326.282	6.00x10 ⁻¹¹
13q31.1	rs1359790	80.717.156	6.49x10 ⁻⁰⁹
14q24.1	rs1465788	69.263.599	2.00x10 ⁻¹²
14q32.2	rs4900384	98.498.951	4.00x10 ⁻⁰⁹
14q32.2	rs941576	101.306.045	1.00x10 ⁻¹⁰
15q14	rs7403531	38.822.905	4.00x10 ⁻⁰⁹
15q22.2	rs7172432	62.396.389	9.00x10 ⁻¹⁴
15q22.2	rs7173964	62.396.942	6.00x10 ⁻¹²
15q22.2	rs11071657	62.433.962	4.00x10 ⁻⁰⁸
15q24.3	rs7178572	77.747.190	2.00x10 ⁻¹¹
15q25.1	rs3825932	79.235.446	3.17x10 ⁻¹⁵
15q25.1	rs11634397	80.432.222	2.00x10 ⁻⁰⁹
15q26.1	rs2028299	90.374.257	2.00x10 ⁻¹¹
15q26.1	rs8042680	91.521.337	2.00x10 ⁻¹⁰
15q26.3	rs2018860	99.258.710	2.99x10 ⁻⁰⁸
16p13.13	rs12708716	11.179.873	3.00x10 ⁻¹⁸
16p13.13	rs2903692	11.238.783	1.03x10 ⁻¹⁰
16p11.2	rs4788084	28.539.848	2.60x10 ⁻¹³
16q12.2	rs8050136	53.816.275	2.00x10 ⁻¹⁷
16q12.2	rs9936385	53.819.169	1.00x10 ⁻¹²
16q12.2	rs9939609	53.820.527	1.00x10 ⁻²⁰

Locus	lead SNP	Position	p-value
16q12.2	rs11642841	53.845.487	3.40x10 ⁻⁰⁸
16q23.1	rs7202877	75.247.245	3.10x10 ⁻¹⁵
17p13.3	rs391300	2.216.258	3.06x10 ⁻⁰⁹
17p13.1	rs312457	6.940.393	8.00x10 ⁻¹³
17p13.1	rs75493593	6.945.087	5.00x10 ⁻¹⁵
17q12	rs4430796	36.098.040	2.00x10 ⁻¹¹
17q21.1	rs2290400	38.066.240	5.50x10 ⁻¹³
17q21.2	rs7221109	38.770.286	1.00x10 ⁻⁰⁹
18p11.31	rs8090011	7.068.462	8.00x10 ⁻⁰⁹
18p11.21	rs2542151	12.779.947	1.00x10 ⁻¹⁴
18p11.21	rs1893217	12.809.340	4.00x10 ⁻¹⁵
18q21.32	rs12970134	57.884.750	2.60x10 ⁻⁰⁸
18q22.2	rs763361	67.531.642	1.00x10 ⁻⁰⁸
19p13.2	rs2304256	10.475.652	4.00x10 ⁻⁰⁹
19q13.11	rs3786897	33.893.008	1.00x10 ⁻⁰⁸
19q13.32	rs425105	47.208.481	3.00x10 ⁻¹¹
20p13	rs2281808	1.610.551	1.00x10 ⁻¹¹
20p11.21	rs6048205	22.559.601	2.00x10 ⁻¹²
20p11.21	rs1209523	22.567.942	2.00x10 ⁻¹¹
20p11.21	rs6048216	22.581.268	2.00x10 ⁻¹²
20q13.12	rs6017317	42.946.966	1.00x10 ⁻¹¹
20q13.12	rs4812829	42.989.267	3.00x10 ⁻¹⁰
21q22.3	rs11203203	43.836.186	2.00x10 ⁻⁰⁹
21q22.3	rs9976767	43.836.390	2.00x10 ⁻⁰⁸
22q12.2	rs5753037	30.581.722	3.00x10 ⁻¹⁶
22q12.3	rs229541	37.591.318	2.00x10 ⁻⁰⁸
Xq28	rs12010175	152.862.638	2.00x10 ⁻⁰⁹
Xq28	rs5945326	152.899.922	7.00x10 ⁻¹⁶
Xq28	rs2664170	153.945.602	8.00x10 ⁻⁰⁹

Supplementary Table 14 | Smoking SNPs

Locus	lead SNP	Position	p-value
8p11.21	rs1451240	42.546.711	6.70x10 ⁻¹⁶
15q25.1	rs1051730	78.894.339	6.00x10 ⁻²⁰

Supplementary Table 15 | Atherosclerosis haplotype blocks

Locus	lead SNPs	5' SNP	3' SNP	5' position	3' position	length (bp)
1p32.3	rs11206510	rs2864123	rs17192725	55.481.104	55.496.131	15.027
1p32.2	rs17114036, rs9970807	rs76380660	rs17416285	56.899.575	57.020.650	121.075
1p13.3	rs7528419, rs660240, rs646776, rs602633, rs599839	rs57166416	rs14000	109.739.555	109.822.509	82.954
1q21.3	rs6689306, rs4845625	rs4601580	rs61811372	154.394.417	154.543.215	148.798
1q24.2	rs10919065, rs1892094, rs1200159	rs754551	rs34525133	169.087.787	169.440.938	353.151
1q25.3	rs10911021	rs73064624	rs12730154	182.047.972	182.297.222	249.250
1q32.1.1	rs6700559	rs72743086	rs2808234	200.469.229	200.666.887	197.658
1q32.1.2	rs2820315, rs2819348	rs704633	rs17432675	201.760.981	201.887.721	126.740
1q41	rs35700460, rs17465637, rs67180937	rs17163191	rs4846778	222.719.586	222.971.968	252.382
2p24.1.0	rs16986953	rs3887882	rs74173283	19.916.846	19.943.923	27.077
2p24.1.00	rs2123536	rs4666560	rs62110853	19.945.418	19.979.502	34.084
2p24.1.1	rs10199768	rs11673889	rs531819	21.104.631	21.263.639	159.008
2p24.1.2	rs515135, rs34908258	rs531819	rs7605009	21.263.639	21.540.399	276.760
2p23.3	rs1260326	rs4665378	rs6753534	27.548.038	27.752.871	204.833
2p21	rs4953024, rs6544713	rs55680582	rs34198326	44.049.795	44.101.229	51.434
2p16.2	rs1559040	rs7584996	rs7566894	53.861.607	54.386.344	524.737
2p11.2	rs10176176, rs7568458, rs1561198	rs72838574	rs2576402	85.689.845	85.893.340	203.495
2q22.3	rs17678683	rs13403907	rs13018460	145.222.038	145.303.576	81.538
2q22.3.1	rs2252641	rs17479267	rs6755330	145.630.762	145.969.587	338.825
2q24.2	rs4665058	rs12991748	rs113353693	160.148.532	160.201.480	52.948

Locus	lead SNPs	5' SNP	3' SNP	5' position	3' position	length (bp)
2q31.3	rs16866933	rs16866933	rs78200327	180.566.678	180.662.102	95.424
2q33.2	rs6725887, rs7582720, rs577594671	rs72932707	rs7572405	203.639.395	204.444.331	804.936
2q35	rs2571445	rs4674215	rs1122987	218.666.889	218.704.894	38.005
2q37.1	rs1801251	rs6738529	rs58660411	233.550.961	233.807.585	256.624
3p24.3	rs4618210	rs55683442	rs78991648	16.665.458	17.152.071	486.613
3q22.3	rs139016349, rs2306374, rs9818870	rs4678407	rs13080413	138.050.685	138.129.143	78.458
3q26.0	rs11924705, rs6789378	rs2897218	rs74423902	162.068.721	162.260.451	191.730
3q26.1	rs1803274	rs35143009	rs13072970	165.436.928	165.782.436	345.508
4p16.2	rs7697839, rs7673097	rs73211129	rs17457282	5.255.563	5.398.808	143.245
4p16.1	rs7671266, rs4698036	rs12649245	rs28733364	9.915.173	10.416.739	501.566
4q12	rs17087335	rs71657222	rs62308298	57.696.697	57.971.951	275.254
4q22.1	rs2199936	rs12505410	rs7672821	89.030.841	89.099.150	68.309
4q31.22.1	rs4593108	rs28627143	rs1560231	148.148.844	148.362.953	214.109
4q31.22.2	rs6841581, rs1878406	rs72955609	rs6537484	148.363.840	148.427.893	64.053
4q32.1.1	rs1842896	rs80115713	rs7657725	156.131.360	156.518.696	387.336
4q32.1.2	rs72689147, rs7692387	rs17388219	rs34922326	156.608.338	156.686.754	78.416
5p15.33	rs11748327	rs593677	rs34367558	3.896.734	4.179.913	283.179
5q22.2	rs4621553	rs115139014	rs72807696	113.013.296	113.117.822	104.526
5q31.1	rs273909	rs270610	rs2073644	131.627.840	131.723.289	95.449
5q31.3	rs246600	rs3822388	rs3822384	142.464.329	142.520.586	56.257

Locus	lead SNPs	5' SNP	3' SNP	5' position	3' position	length (bp)
5q33.2	rs12189362	rs55903771	rs815622	153.052.642	153.494.463	441.821
6p24.1	rs9369640, rs4714955, rs9349379, rs12526453	rs9472484	rs112239423	12.758.212	12.950.520	192.308
6p22.2	rs11754288, rs1800562	rs80154106	rs806970	25.684.802	26.141.375	456.573
6p21.33	rs3869109	rs1265156	rs113390008	31.142.297	31.207.351	65.054
6p21.33.1	rs3130683	rs62395851	rs616634	31.881.223	31.948.448	67.225
6p21.32	rs9268402	rs1967688	rs368404259	32.340.068	32.454.575	114.507
6p21.31	rs17609940	rs57844307	rs760782	34.980.945	35.448.189	467.244
6p21.2	rs1544935, rs56336142, rs10947789	rs9296275	rs3892126	39.024.619	39.189.959	165.340
6q13	rs9351814	rs828656	rs489755	72.057.634	72.205.797	148.163
6q23.2	rs12202017, rs12190287	rs1208240	rs162189	134.097.764	134.243.918	146.154
6q25.1.1	rs6922269	rs67008194	rs9478863	151.249.527	151.265.927	16.400
6q25.1.2	rs2982694	rs9322336	rs2144025	152.200.430	152.307.706	107.276
6q25.3	rs3798220, rs2048327, rs3127599, rs7767084, rs10755578	rs3127578	rs12210186	160.698.155	160.978.997	280.842
6q25.3.26	rs55730499, rs10455872, rs2315065, rs4252185	rs9364559	rs9458010	160.976.148	161.125.419	149.271
6q26	rs4252120	rs4252066	rs76031264	161.127.125	161.312.655	185.530
7p21.1	rs2107595, rs2023938	rs11984041	rs2717327	19.031.935	19.064.300	32.365
7q22.3	rs10953541	rs111354917	rs2712227	106.665.984	107.305.115	639.131
7q32.2	rs11556924	rs1045896	rs191086346	129.592.517	129.707.897	115.380
7q34	rs10237377	rs10234812	rs7801310	139.716.145	139.761.248	45.103
7q36.1	rs3918226	rs57293996	rs2435608	150.495.187	150.890.806	395.619

Locus	lead SNPs	5' SNP	3' SNP	5' position	3' position	length (bp)
8p21.3.0	rs264	rs1441778	rs278	19.727.047	19.814.523	87.476
8p21.3	rs17091905	rs301	rs12676079	19.816.934	19.943.364	126.430
8q24.13	rs2954029	rs72647352	rs6989280	126.452.509	126.508.746	56.237
9p21.3	rs3217992, rs7865618, rs1537370, rs10757274, rs4977574, rs2891168, rs1333042, rs944797, rs10757278, rs1333048, rs1333049	rs62556523	rs1333049	21.937.687	22.125.503	187.816
9q34.2	rs2519093, rs514659, rs532436, rs579459	rs11789171	rs3124767	136.122.488	136.308.542	186.054
10p11.23	rs3739998, rs2487928, rs2505083	rs12764479	rs2948899	30.286.263	30.566.692	280.429
10q11.21	rs1870634, rs2047009, rs501120, rs1746048	rs12572433	rs1159931	44.432.455	44.806.042	373.587
10q23.31	rs1412444, rs1332329, rs2246833	rs2071509	rs12258120	90.988.250	91.014.351	26.101
10q24.32	rs11191416, rs12413409	rs76085620	rs11191612	104.219.908	104.969.578	749.670
11p15.4	rs10840293	rs7102407	rs11042750	9.556.117	10.372.959	816.842
11p15.3	rs11042937	rs7929879	rs7115390	10.702.348	10.757.993	55.645
11q13.1	rs12801636	rs7116712	rs72922800	65.372.518	65.570.754	198.236
11q13.5	rs590121	rs590121	rs1144150	75.274.150	75.285.109	10.959
11q21	rs10765792	rs7945959	rs79686966	95.861.651	95.979.334	117.683
11q22.3	rs974819, rs2019090, rs2128739	rs10488758	rs561801	103.493.096	103.763.585	270.489
11q23.3.1	rs964184	rs2212428	rs1729409	116.516.587	116.673.768	157.181
11q23.3.2	rs508487	rs5142	rs12793361	116.701.850	117.101.877	400.027
12q13.3	rs11172113	rs703816	rs1799737	57.497.005	57.534.912	37.907
12q21.2	rs7307780	rs12315478	rs11180748	76.116.463	76.303.759	187.296
12q21.33	rs2681472, rs7136259	rs4842666	rs139338748	89.941.549	90.228.941	287.392

Locus	lead SNPs	5' SNP	3' SNP	5' position	3' position	length (bp)
12q23.3	rs1165668, rs1165669	rs1165592	rs1905849	104.232.180	104.454.527	222.347
12q24.11	rs3782889	rs2339635	rs2040571	111.284.872	111.357.727	72.855
12q24.12.13	rs3184504, rs653178, rs11065979, rs11065987, rs3782886, rs11066015, rs671, rs11066280	rs10849941	rs112505310	111.818.487	113.027.651	1.209.164
12q24.23	rs11830157	rs479143	rs2592295	118.257.673	118.285.904	28.231
12q24.31	rs3213545, rs2258287, rs2708081, rs2244608, rs1169288	rs67279017	rs45555142	121.041.224	121.490.498	449.274
12q24.31.1	rs11057830	rs11057820	rs12581963	125.296.964	125.317.125	20.161
13q12.3	rs9319428	rs9508016	rs6490304	28.895.069	29.025.116	130.047
13q22.1	rs12429889	rs9530283	rs966081	74.715.428	74.865.005	149.577
13q34.1	rs11617955	rs7988738	rs1838557	110.785.284	110.821.834	36.550
13q34.2	rs4773144	rs74124096	rs9559766	110.953.710	110.962.689	8.979
13q34.3	rs11838776, rs9515203	rs9521715	rs7334986	111.015.333	111.080.609	65.276
14q31.3	rs11624056	rs1638719	rs7160581	87.503.864	87.633.901	130.037
14q32.2	rs2895811, rs10139550	rs4900436	rs76927709	100.063.084	100.191.710	128.626
15q21.3	rs10468017	rs117486534	rs12914524	58.669.703	58.716.647	46.944
15q22.31	rs6494488	rs332271	rs4777602	64.339.748	65.104.826	765.078
15q22.33	rs56062135	rs2053295	rs112146149	67.391.739	67.468.715	76.976
15q25.1.1	rs899997, rs2219939, rs1994016, rs3825807	rs12915539	rs12907065	78.940.795	79.107.355	166.560
15q25.1.2	rs4380028, rs7165042, rs4468572, rs7173743	rs12907065	rs57533456	79.107.355	79.258.685	151.330
15q26.1	rs8042271	rs2280213	rs28701165	89.452.763	89.581.388	128.625
15q26.1.2	rs17514846	rs12917588	rs10744971	91.390.651	91.443.059	52.408

Locus	lead SNPs	5' SNP	3' SNP	5' position	3' position	length (bp)
16q13	rs173539, rs247616, rs1800775	rs1122390	rs112039804	56.986.045	57.018.856	32.811
16q22.2	rs1050362, rs2072142	rs2052579	rs1006436	72.044.005	72.232.316	188.311
17p13.3	rs1231206, rs216172, rs2281727	rs72815303	rs12945198	1.999.941	2.220.815	220.874
17p11.2	rs12936587, rs12449964	rs3785499	rs11656699	17.415.217	17.694.761	279.544
17q21.32.1	rs17608766	rs12600562	rs9901977	44.977.040	45.048.564	71.524
17q21.32	rs46522	rs869492	rs12602595	46.835.629	47.159.335	323.706
17q23.2	rs7212798	rs12149972	rs150043692	58.813.056	59.312.396	499.340
17q23.3	rs1867624, rs9892152	rs7221302	rs1987032	62.364.965	62.410.447	45.482
17q24.3	rs17718586	rs17780119	rs16976211	68.516.253	68.738.756	222.503
18p11.31	rs597503	rs71360054	rs949215	6.930.792	6.955.676	24.884
18q11.2	rs16942421	rs73944615	rs62085477	24.120.614	24.157.155	36.541
18q21.32	rs663129	rs4121765	rs61194872	57.730.096	57.915.448	185.352
19p13.3	rs3803915	rs9967630	rs8112524	2.099.820	2.250.528	150.708
19p13.2.1	rs1122608	rs11085749	rs11085758	10.961.273	11.185.014	223.741
19p13.2.2	rs55791371, rs6511720, rs56289821	rs143020224	rs2228671	11.187.324	11.210.912	23.588
19q13.11	rs12976411	rs28607751	rs358539	32.747.302	32.996.606	249.304
19q13.32	rs2075650, rs445925, rs56131196, rs4420638	rs3745150	rs7254133	45.385.759	45.438.554	52.795
20q11.22	rs867186	rs35291022	rs6142329	33.537.157	33.795.623	258.466
21q22.11	rs28451064, rs9982601	rs28451064	rs4817645	35.593.827	35.724.134	130.307
22q11.23	rs180803	rs144405277	rs11704820	24.613.907	24.912.248	298.341

B

References

- 1 Disease, G. B. D., Injury, I. & Prevalence, C. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789-1858, doi:10.1016/S0140-6736(18)32279-7 (2018).
- 2 DALYs, G. B. D. & Collaborators, H. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1859-1922, doi:10.1016/S0140-6736(18)32335-3 (2018).
- 3 Collaborators, G. B. D. C. o. D. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1736-1788, doi:10.1016/S0140-6736(18)32203-7 (2018).
- 4 Ross, R. Atherosclerosis--an inflammatory disease. *N Engl J Med* **340**, 115-126, doi:10.1056/NEJM199901143400207 (1999).
- 5 Napoli, C. *et al.* Fatty streak formation occurs in human fetal aortas and is greatly enhanced by maternal hypercholesterolemia. Intimal accumulation of low density lipoprotein and its oxidation precede monocyte recruitment into early atherosclerotic lesions. *J Clin Invest* **100**, 2680-2690, doi:10.1172/JCI119813 (1997).
- 6 Quinn, M. T., Parthasarathy, S., Fong, L. G. & Steinberg, D. Oxidatively modified low density lipoproteins: a potential role in recruitment and retention of monocyte/macrophages during atherogenesis. *Proc Natl Acad Sci U S A* **84**, 2995-2998, doi:10.1073/pnas.84.9.2995 (1987).
- 7 Reidy, M. A. & Bowyer, D. E. Scanning electron microscopy of arteries. The morphology of aortic endothelium in haemodynamically stressed areas associated with branches. *Atherosclerosis* **26**, 181-194, doi:10.1016/0021-9150(77)90101-0 (1977).
- 8 Levesque, M. J., Liepsch, D., Moravec, S. & Nerem, R. M. Correlation of endothelial cell shape and wall shear stress in a stenosed dog aorta. *Arteriosclerosis* **6**, 220-229, doi:10.1161/01.atv.6.2.220 (1986).
- 9 Lusis, A. J. Atherosclerosis. *Nature* **407**, 233-241, doi:10.1038/35025203 (2000).
- 10 Gimbrone, M. A., Jr. Vascular endothelium, hemodynamic forces, and atherogenesis. *Am J Pathol* **155**, 1-5, doi:10.1016/S0002-9440(10)65090-0 (1999).

- 11 BabaeV, V. R., Patel, M. B., Semenkovich, C. F., Fazio, S. & Linton, M. F. Macrophage lipoprotein lipase promotes foam cell formation and atherosclerosis in low density lipoprotein receptor-deficient mice. *J Biol Chem* **275**, 26293-26299, doi:10.1074/jbc.M002423200 (2000).
- 12 Bentzon, J. F., Otsuka, F., Virmani, R. & Falk, E. Mechanisms of plaque formation and rupture. *Circ Res* **114**, 1852-1866, doi:10.1161/CIRCRESAHA.114.302721 (2014).
- 13 Zdravkovic, S. *et al.* Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J Intern Med* **252**, 247-254 (2002).
- 14 Wienke, A., Holm, N. V., Skytthe, A. & Yashin, A. I. The heritability of mortality due to heart diseases: a correlated frailty model applied to Danish twins. *Twin Res* **4**, 266-274 (2001).
- 15 Lloyd-Jones, D. M. *et al.* Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA* **291**, 2204-2211, doi:10.1001/jama.291.18.2204 (2004).
- 16 O'Donnell, C. J. & Nabel, E. G. Genomics of cardiovascular disease. *N Engl J Med* **365**, 2098-2109, doi:10.1056/NEJMra1105239 (2011).
- 17 Lehrman, M. A. *et al.* Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science* **227**, 140-146 (1985).
- 18 Garcia, C. K. *et al.* Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. *Science* **292**, 1394-1398, doi:10.1126/science.1060458 (2001).
- 19 Brown, M. S. & Goldstein, J. L. A receptor-mediated pathway for cholesterol homeostasis. *Science* **232**, 34-47 (1986).
- 20 Myers, R. H., Kiely, D. K., Cupples, L. A. & Kannel, W. B. Parental history is an independent risk factor for coronary artery disease: the Framingham Study. *Am Heart J* **120**, 963-969, doi:10.1016/0002-8703(90)90216-k (1990).
- 21 Peden, J. F. & Farrall, M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. *Hum Mol Genet* **20**, R198-205, doi:10.1093/hmg/ddr384 (2011).
- 22 Nikpay, M. *et al.* A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* **47**, 1121-1130, doi:10.1038/ng.3396 (2015).
- 23 Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74, doi:10.1038/nature15393 (2015).
- 24 Bush, W. S. & Moore, J. H. Chapter 11: Genome-wide association studies. *PLoS Comput Biol* **8**, e1002822, doi:10.1371/journal.pcbi.1002822 (2012).
- 25 Sudmant, P. H. *et al.* An integrated map of structural variation in 2,504 human genomes. *Nature* **526**, 75-81, doi:10.1038/nature15394 (2015).

- 26 International HapMap, C. A haplotype map of the human genome. *Nature* **437**, 1299-1320, doi:10.1038/nature04226 (2005).
- 27 Devlin, B. & Risch, N. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **29**, 311-322, doi:10.1006/geno.1995.9003 (1995).
- 28 Distefano, J. K. & Taverna, D. M. Technological issues and experimental design of gene association studies. *Methods Mol Biol* **700**, 3-16, doi:10.1007/978-1-61737-954-3_1 (2011).
- 29 van den Oord, E. J. Controlling false discoveries in genetic studies. *Am J Med Genet B Neuropsychiatr Genet* **147B**, 637-644, doi:10.1002/ajmg.b.30650 (2008).
- 30 Hirschhorn, J. N. & Daly, M. J. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* **6**, 95-108, doi:10.1038/nrg1521 (2005).
- 31 Samani, N. J. *et al.* Genomewide association analysis of coronary artery disease. *N Engl J Med* **357**, 443-453, doi:10.1056/NEJMoa072366 (2007).
- 32 Wang, F. *et al.* Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. *Nat Genet* **43**, 345-349, doi:10.1038/ng.783 (2011).
- 33 Coronary Artery Disease Genetics, C. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet* **43**, 339-344, doi:10.1038/ng.782 (2011).
- 34 Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* **43**, 333-338, doi:10.1038/ng.784 (2011).
- 35 Consortium, C. A. D. *et al.* Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet* **45**, 25-33, doi:10.1038/ng.2480 (2013).
- 36 Holdt, L. M. *et al.* Alu elements in ANRIL non-coding RNA at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks. *PLoS Genet* **9**, e1003588, doi:10.1371/journal.pgen.1003588 (2013).
- 37 Holdt, L. M. *et al.* Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat Commun* **7**, 12429, doi:10.1038/ncomms12429 (2016).
- 38 Holdt, L. M. & Teupser, D. Long Noncoding RNA ANRIL: Lnc-ing Genetic Variation at the Chromosome 9p21 Locus to Molecular Mechanisms of Atherosclerosis. *Front Cardiovasc Med* **5**, 145, doi:10.3389/fcvm.2018.00145 (2018).
- 39 Maurano, M. T. *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190-1195, doi:10.1126/science.1222794 (2012).
- 40 Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304-1351, doi:10.1126/science.1058040 (2001).
- 41 Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921, doi:10.1038/35057062 (2001).

- 42 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).
- 43 Harrow, J. *et al.* GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* **22**, 1760-1774, doi:10.1101/gr.135350.111 (2012).
- 44 Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* **22**, 1775-1789, doi:10.1101/gr.132159.111 (2012).
- 45 Jarroux, J., Morillon, A. & Pinskaya, M. History, Discovery, and Classification of lncRNAs. *Adv Exp Med Biol* **1008**, 1-46, doi:10.1007/978-981-10-5203-3_1 (2017).
- 46 Rinn, J. L. & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annu Rev Biochem* **81**, 145-166, doi:10.1146/annurev-biochem-051410-092902 (2012).
- 47 Akhade, V. S., Pal, D. & Kanduri, C. Long Noncoding RNA: Genome Organization and Mechanism of Action. *Adv Exp Med Biol* **1008**, 47-74, doi:10.1007/978-981-10-5203-3_2 (2017).
- 48 Beltran, M. *et al.* A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev* **22**, 756-769, doi:10.1101/gad.455708 (2008).
- 49 Gong, C. & Maquat, L. E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **470**, 284-288, doi:10.1038/nature09701 (2011).
- 50 Wang, H. *et al.* Dendritic BC1 RNA in translational control mechanisms. *J Cell Biol* **171**, 811-821, doi:10.1083/jcb.200506006 (2005).
- 51 Kindler, S., Wang, H., Richter, D. & Tiedge, H. RNA transport and local control of translation. *Annu Rev Cell Dev Biol* **21**, 223-245, doi:10.1146/annurev.cellbio.21.122303.120653 (2005).
- 52 Yoon, J. H. *et al.* LincRNA-p21 suppresses target mRNA translation. *Mol Cell* **47**, 648-655, doi:10.1016/j.molcel.2012.06.027 (2012).
- 53 Carrieri, C. *et al.* Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **491**, 454-457, doi:10.1038/nature11508 (2012).
- 54 Cesana, M. *et al.* A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **147**, 358-369, doi:10.1016/j.cell.2011.09.028 (2011).
- 55 Zhang, Z., Salisbury, D. & Sallam, T. Long Noncoding RNAs in Atherosclerosis: JACC Review Topic of the Week. *J Am Coll Cardiol* **72**, 2380-2390, doi:10.1016/j.jacc.2018.08.2161 (2018).
- 56 Wang, Y., Song, X., Li, Z. & Liu, B. Long non-coding RNAs in coronary atherosclerosis. *Life Sci* **211**, 189-197, doi:10.1016/j.lfs.2018.08.072 (2018).
- 57 Halley, P. *et al.* Regulation of the apolipoprotein gene cluster by a long noncoding RNA. *Cell Rep* **6**, 222-230, doi:10.1016/j.celrep.2013.12.015 (2014).

- 58 Liu, G. *et al.* Long non-coding RNAs expression profile in HepG2 cells reveals the potential role of long non-coding RNAs in the cholesterol metabolism. *Chin Med J (Engl)* **128**, 91-97, doi:10.4103/0366-6999.147824 (2015).
- 59 Cui, M. *et al.* Long noncoding RNA HULC modulates abnormal lipid metabolism in hepatoma cells through an miR-9-mediated RXRA signaling pathway. *Cancer Res* **75**, 846-857, doi:10.1158/0008-5472.CAN-14-1192 (2015).
- 60 Li, P. *et al.* A liver-enriched long non-coding RNA, lncLSTR, regulates systemic lipid metabolism in mice. *Cell Metab* **21**, 455-467, doi:10.1016/j.cmet.2015.02.004 (2015).
- 61 Powell, W. T. *et al.* A Prader-Willi locus lncRNA cloud modulates diurnal genes and energy expenditure. *Hum Mol Genet* **22**, 4318-4328, doi:10.1093/hmg/ddt281 (2013).
- 62 Liu, S. *et al.* SRA gene knockout protects against diet-induced obesity and improves glucose tolerance. *J Biol Chem* **289**, 13000-13009, doi:10.1074/jbc.M114.564658 (2014).
- 63 Gao, Y. *et al.* The H19/let-7 double-negative feedback loop contributes to glucose metabolism in muscle cells. *Nucleic Acids Res* **42**, 13799-13811, doi:10.1093/nar/gku1160 (2014).
- 64 Trajkovski, M. *et al.* MicroRNAs 103 and 107 regulate insulin sensitivity. *Nature* **474**, 649-653, doi:10.1038/nature10112 (2011).
- 65 Wessel, J. *et al.* Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility. *Nat Commun* **6**, 5897, doi:10.1038/ncomms6897 (2015).
- 66 Wang, Y. N. *et al.* Long Noncoding RNA-GAS5: A Novel Regulator of Hypertension-Induced Vascular Remodeling. *Hypertension* **68**, 736-748, doi:10.1161/HYPERTENSIONAHA.116.07259 (2016).
- 67 Leung, A. *et al.* Novel long noncoding RNAs are regulated by angiotensin II in vascular smooth muscle cells. *Circ Res* **113**, 266-278, doi:10.1161/CIRCRESAHA.112.300849 (2013).
- 68 Bayoglu, B., Yuksel, H., Cakmak, H. A., Dirican, A. & Cengiz, M. Polymorphisms in the long non-coding RNA CDKN2B-AS1 may contribute to higher systolic blood pressure levels in hypertensive patients. *Clin Biochem* **49**, 821-827, doi:10.1016/j.clinbiochem.2016.02.012 (2016).
- 69 Nurnberg, S. T. *et al.* From Loci to Biology: Functional Genomics of Genome-Wide Association for Coronary Disease. *Circ Res* **118**, 586-606, doi:10.1161/CIRCRESAHA.115.306464 (2016).
- 70 Consortium, G. T. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-660, doi:10.1126/science.1262110 (2015).
- 71 Rockman, M. V. & Kruglyak, L. Genetics of global gene expression. *Nat Rev Genet* **7**, 862-872, doi:10.1038/nrg1964 (2006).

- 72 Albert, F. W. & Kruglyak, L. The role of regulatory variation in complex traits and disease. *Nat Rev Genet* **16**, 197-212, doi:10.1038/nrg3891 (2015).
- 73 Foroughi Asl, H. *et al.* Expression quantitative trait Loci acting across multiple tissues are enriched in inherited risk for coronary artery disease. *Circ Cardiovasc Genet* **8**, 305-315, doi:10.1161/CIRCGENETICS.114.000640 (2015).
- 74 Pickrell, J. K. *et al.* Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* **464**, 768-772, doi:10.1038/nature08872 (2010).
- 75 Lappalainen, T. *et al.* Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* **501**, 506-511, doi:10.1038/nature12531 (2013).
- 76 Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821, doi:10.1126/science.1225829 (2012).
- 77 Strong, A. & Musunuru, K. Genome editing in cardiovascular diseases. *Nat Rev Cardiol* **14**, 11-20, doi:10.1038/nrcardio.2016.139 (2017).
- 78 Patsch, C. *et al.* Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. *Nat Cell Biol* **17**, 994-1003, doi:10.1038/ncb3205 (2015).
- 79 Ding, Q. *et al.* Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* **12**, 393-394, doi:10.1016/j.stem.2013.03.006 (2013).
- 80 Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826, doi:10.1126/science.1232033 (2013).
- 81 R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria, 2019).
- 82 Zhang, S. & Cui, W. Sox2, a key factor in the regulation of pluripotency and neural differentiation. *World J Stem Cells* **6**, 305-311, doi:10.4252/wjsc.v6.i3.305 (2014).
- 83 Zhang, L. *et al.* Mesodermal Nkx2.5 is necessary and sufficient for early second heart field development. *Dev Biol* **390**, 68-79, doi:10.1016/j.ydbio.2014.02.023 (2014).
- 84 Yang, R. *et al.* Amnion signals are essential for mesoderm formation in primates. *Nat Commun* **12**, 5126, doi:10.1038/s41467-021-25186-2 (2021).
- 85 Prummel, K. D., Nieuwenhuize, S. & Mosimann, C. The lateral plate mesoderm. *Development* **147**, doi:10.1242/dev.175059 (2020).
- 86 Goncharov, N. V., Nadeev, A. D., Jenkins, R. O. & Avdonin, P. V. Markers and Biomarkers of Endothelium: When Something Is Rotten in the State. *Oxid Med Cell Longev* **2017**, 9759735, doi:10.1155/2017/9759735 (2017).
- 87 Bennett, M. R., Sinha, S. & Owens, G. K. Vascular Smooth Muscle Cells in Atherosclerosis. *Circ Res* **118**, 692-702, doi:10.1161/CIRCRESAHA.115.306361 (2016).

- 88 Jae, N., Heumuller, A. W., Fouani, Y. & Dimmeler, S. Long non-coding RNAs in vascular biology and disease. *Vascul Pharmacol* **114**, 13-22, doi:10.1016/j.vph.2018.03.003 (2019).
- 89 Ross, R. & Glomset, J. A. Atherosclerosis and the arterial smooth muscle cell: Proliferation of smooth muscle is a key event in the genesis of the lesions of atherosclerosis. *Science* **180**, 1332-1339, doi:10.1126/science.180.4093.1332 (1973).
- 90 O'Brien, E. R. *et al.* Proliferation in primary and restenotic coronary atherectomy tissue. Implications for antiproliferative therapy. *Circ Res* **73**, 223-231, doi:10.1161/01.res.73.2.223 (1993).
- 91 Doran, A. C., Meller, N. & McNamara, C. A. Role of smooth muscle cells in the initiation and early progression of atherosclerosis. *Arterioscler Thromb Vasc Biol* **28**, 812-819, doi:10.1161/ATVBAHA.107.159327 (2008).
- 92 Allahverdian, S., Chaabane, C., Boukais, K., Francis, G. A. & Bochaton-Piallat, M. L. Smooth muscle cell fate and plasticity in atherosclerosis. *Cardiovasc Res* **114**, 540-550, doi:10.1093/cvr/cvy022 (2018).
- 93 Liang, C. C., Park, A. Y. & Guan, J. L. In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. *Nat Protoc* **2**, 329-333, doi:10.1038/nprot.2007.30 (2007).
- 94 Mestas, J. & Ley, K. Monocyte-endothelial cell interactions in the development of atherosclerosis. *Trends Cardiovasc Med* **18**, 228-232, doi:10.1016/j.tcm.2008.11.004 (2008).
- 95 Felsenfeld, G., Boyes, J., Chung, J., Clark, D. & Studitsky, V. Chromatin structure and gene expression. *Proc Natl Acad Sci U S A* **93**, 9384-9388, doi:10.1073/pnas.93.18.9384 (1996).
- 96 Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349-352, doi:10.1038/38664 (1997).
- 97 Sproul, D., Gilbert, N. & Bickmore, W. A. The role of chromatin structure in regulating the expression of clustered genes. *Nat Rev Genet* **6**, 775-781, doi:10.1038/nrg1688 (2005).
- 98 Westholm, J. O., Xu, F., Ronne, H. & Komorowski, J. Genome-scale study of the importance of binding site context for transcription factor binding and gene regulation. *BMC Bioinformatics* **9**, 484, doi:10.1186/1471-2105-9-484 (2008).
- 99 Doniger, S. W. & Fay, J. C. Frequent gain and loss of functional transcription factor binding sites. *PLoS Comput Biol* **3**, e99, doi:10.1371/journal.pcbi.0030099 (2007).
- 100 Carrasco Pro, S., Bulekova, K., Gregor, B., Labadorf, A. & Fuxman Bass, J. I. Prediction of genome-wide effects of single nucleotide variants on transcription factor binding. *Sci Rep* **10**, 17632, doi:10.1038/s41598-020-74793-4 (2020).
- 101 Johnston, A. D., Simoes-Pires, C. A., Thompson, T. V., Suzuki, M. & Grealley, J. M. Functional genetic variants can mediate their regulatory effects through alteration of transcription factor binding. *Nat Commun* **10**, 3472, doi:10.1038/s41467-019-11412-5 (2019).

- 102 Kockx, M. M. & Herman, A. G. Apoptosis in atherosclerosis: beneficial or detrimental? *Cardiovasc Res* **45**, 736-746, doi:10.1016/s0008-6363(99)00235-7 (2000).
- 103 Dimmeler, S., Hermann, C. & Zeiher, A. M. Apoptosis of endothelial cells. Contribution to the pathophysiology of atherosclerosis? *Eur Cytokine Netw* **9**, 697-698 (1998).
- 104 Tricot, O. *et al.* Relation between endothelial cell apoptosis and blood flow direction in human atherosclerotic plaques. *Circulation* **101**, 2450-2453, doi:10.1161/01.cir.101.21.2450 (2000).
- 105 Freyberg, M. A., Kaiser, D., Graf, R., Buttenbender, J. & Friedl, P. Proatherogenic flow conditions initiate endothelial apoptosis via thrombospondin-1 and the integrin-associated protein. *Biochem Biophys Res Commun* **286**, 141-149, doi:10.1006/bbrc.2001.5314 (2001).
- 106 Rekhter, M. D. Collagen synthesis in atherosclerosis: too much and not enough. *Cardiovasc Res* **41**, 376-384, doi:10.1016/s0008-6363(98)00321-6 (1999).
- 107 Eken, S. M. *et al.* MicroRNA-210 Enhances Fibrous Cap Stability in Advanced Atherosclerotic Lesions. *Circ Res* **120**, 633-644, doi:10.1161/CIRCRESAHA.116.309318 (2017).
- 108 Bennett, M. R., Evan, G. I. & Schwartz, S. M. Apoptosis of human vascular smooth muscle cells derived from normal vessels and coronary atherosclerotic plaques. *J Clin Invest* **95**, 2266-2274, doi:10.1172/JCI117917 (1995).
- 109 Clarke, M. C. *et al.* Apoptosis of vascular smooth muscle cells induces features of plaque vulnerability in atherosclerosis. *Nat Med* **12**, 1075-1080, doi:10.1038/nm1459 (2006).
- 110 Clarke, M. C. *et al.* Chronic apoptosis of vascular smooth muscle cells accelerates atherosclerosis and promotes calcification and medial degeneration. *Circ Res* **102**, 1529-1538, doi:10.1161/CIRCRESAHA.108.175976 (2008).
- 111 Littlewood, T. D. & Bennett, M. R. Apoptotic cell death in atherosclerosis. *Curr Opin Lipidol* **14**, 469-475, doi:10.1097/00041433-200310000-00007 (2003).
- 112 Patel, V. A. *et al.* Defect in insulin-like growth factor-1 survival mechanism in atherosclerotic plaque-derived vascular smooth muscle cells is mediated by reduced surface binding and signaling. *Circ Res* **88**, 895-902, doi:10.1161/hh0901.090305 (2001).
- 113 Gautier, E. L. *et al.* Macrophage apoptosis exerts divergent effects on atherogenesis as a function of lesion stage. *Circulation* **119**, 1795-1804, doi:10.1161/CIRCULATIONAHA.108.806158 (2009).
- 114 Arai, S. *et al.* A role for the apoptosis inhibitory factor AIM/Spalpha/Api6 in atherosclerosis development. *Cell Metab* **1**, 201-213, doi:10.1016/j.cmet.2005.02.002 (2005).
- 115 Linton, M. F. *et al.* Macrophage Apoptosis and Efferocytosis in the Pathogenesis of Atherosclerosis. *Circ J* **80**, 2259-2268, doi:10.1253/circj.CJ-16-0924 (2016).

- 116 Van Vre, E. A., Ait-Oufella, H., Tedgui, A. & Mallat, Z. Apoptotic cell death and efferocytosis in atherosclerosis. *Arterioscler Thromb Vasc Biol* **32**, 887-893, doi:10.1161/ATVBAHA.111.224873 (2012).
- 117 Tabas, I. Apoptosis and efferocytosis in mouse models of atherosclerosis. *Curr Drug Targets* **8**, 1288-1296, doi:10.2174/138945007783220623 (2007).
- 118 Thorp, E., Cui, D., Schrijvers, D. M., Kuriakose, G. & Tabas, I. Mertk receptor mutation reduces efferocytosis efficiency and promotes apoptotic cell accumulation and plaque necrosis in atherosclerotic lesions of apoe^{-/-} mice. *Arterioscler Thromb Vasc Biol* **28**, 1421-1428, doi:10.1161/ATVBAHA.108.167197 (2008).
- 119 Lutgens, E. *et al.* Biphasic pattern of cell turnover characterizes the progression from fatty streaks to ruptured human atherosclerotic plaques. *Cardiovasc Res* **41**, 473-479, doi:10.1016/s0008-6363(98)00311-3 (1999).
- 120 Libby, P., Ridker, P. M. & Hansson, G. K. Progress and challenges in translating the biology of atherosclerosis. *Nature* **473**, 317-325, doi:10.1038/nature10146 (2011).
- 121 Schober, A. *et al.* MicroRNA-126-5p promotes endothelial proliferation and limits atherosclerosis by suppressing Dlk1. *Nat Med* **20**, 368-376, doi:10.1038/nm.3487 (2014).
- 122 Obikane, H. *et al.* Effect of endothelial cell proliferation on atherogenesis: a role of p21(Sdi/Cip/Waf1) in monocyte adhesion to endothelial cells. *Atherosclerosis* **212**, 116-122, doi:10.1016/j.atherosclerosis.2010.05.029 (2010).
- 123 Rudijanto, A. The role of vascular smooth muscle cells on the pathogenesis of atherosclerosis. *Acta Med Indones* **39**, 86-93 (2007).
- 124 Kraemer, R. Regulation of cell migration in atherosclerosis. *Curr Atheroscler Rep* **2**, 445-452, doi:10.1007/s11883-000-0084-2 (2000).
- 125 Ferns, G. A. *et al.* Inhibition of neointimal smooth muscle accumulation after angioplasty by an antibody to PDGF. *Science* **253**, 1129-1132, doi:10.1126/science.1653454 (1991).
- 126 Fingerle, J., Johnson, R., Clowes, A. W., Majesky, M. W. & Reidy, M. A. Role of platelets in smooth muscle cell proliferation and migration after vascular injury in rat carotid artery. *Proc Natl Acad Sci U S A* **86**, 8412-8416, doi:10.1073/pnas.86.21.8412 (1989).
- 127 Newby, A. C. & Zaltsman, A. B. Fibrous cap formation or destruction--the critical importance of vascular smooth muscle cell proliferation, migration and matrix formation. *Cardiovasc Res* **41**, 345-360 (1999).
- 128 Newby, A. C. & Zaltsman, A. B. Molecular mechanisms in intimal hyperplasia. *J Pathol* **190**, 300-309, doi:10.1002/(SICI)1096-9896(200002)190:3<300::AID-PATH596>3.0.CO;2-I (2000).
- 129 Schwartz, S. M. Perspectives series: cell adhesion in vascular biology. Smooth muscle migration in atherosclerosis and restenosis. *J Clin Invest* **99**, 2814-2816, doi:10.1172/JCI119472 (1997).

- 130 Liao, J. K. Linking endothelial dysfunction with endothelial cell activation. *J Clin Invest* **123**, 540-541, doi:10.1172/JCI66843 (2013).
- 131 Galkina, E. & Ley, K. Vascular adhesion molecules in atherosclerosis. *Arterioscler Thromb Vasc Biol* **27**, 2292-2301, doi:10.1161/ATVBAHA.107.149179 (2007).
- 132 Beaney, K. E. *et al.* Functional Analysis of the Coronary Heart Disease Risk Locus on Chromosome 21q22. *Dis Markers* **2017**, 1096916, doi:10.1155/2017/1096916 (2017).
- 133 Wang, Y. & Tabas, I. Emerging roles of mitochondria ROS in atherosclerotic lesions: causation or association? *J Atheroscler Thromb* **21**, 381-390, doi:10.5551/jat.23929 (2014).
- 134 Li, J. *et al.* The oncogenic potentials and diagnostic significance of long non-coding RNA LINC00310 in breast cancer. *J Cell Mol Med* **22**, 4486-4495, doi:10.1111/jcmm.13750 (2018).
- 135 Xia, S. *et al.* Dynamic changes in HCN2, HCN4, KCNE1, and KCNE2 expression in ventricular cells from acute myocardial infarction rat hearts. *Biochem Biophys Res Commun* **395**, 330-335, doi:10.1016/j.bbrc.2010.04.003 (2010).
- 136 Abbott, G. W. *et al.* MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell* **97**, 175-187, doi:10.1016/s0092-8674(00)80728-x (1999).
- 137 Thomas, G. *et al.* Mechanisms of ventricular arrhythmogenesis in mice following targeted disruption of KCNE1 modelling long QT syndrome 5. *J Physiol* **578**, 99-114, doi:10.1113/jphysiol.2006.118133 (2007).
- 138 Lee, S. M., Nguyen, D., Hu, Z. & Abbott, G. W. Kcne2 deletion promotes atherosclerosis and diet-dependent sudden death. *J Mol Cell Cardiol* **87**, 148-151, doi:10.1016/j.yjmcc.2015.08.013 (2015).
- 139 Hu, Z., Crump, S. M., Zhang, P. & Abbott, G. W. Kcne2 deletion attenuates acute post-ischaemia/reperfusion myocardial infarction. *Cardiovasc Res* **110**, 227-237, doi:10.1093/cvr/cvw048 (2016).
- 140 Mendez-Barbero, N. *et al.* A major role for RCAN1 in atherosclerosis progression. *EMBO Mol Med* **5**, 1901-1917, doi:10.1002/emmm.201302842 (2013).
- 141 Torac, E., Gaman, L. & Atanasiu, V. The regulator of calcineurin (RCAN1) an important factor involved in atherosclerosis and cardiovascular diseases development. *J Med Life* **7**, 481-487 (2014).
- 142 Myocardial Infarction Genetics, C. *et al.* Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet* **41**, 334-341, doi:10.1038/ng.327 (2009).
- 143 Niland, C. N., Merry, C. R. & Khalil, A. M. Emerging Roles for Long Non-Coding RNAs in Cancer and Neurological Disorders. *Front Genet* **3**, 25, doi:10.3389/fgene.2012.00025 (2012).

- 144 Spizzo, R., Almeida, M. I., Colombatti, A. & Calin, G. A. Long non-coding RNAs and cancer: a new frontier of translational research? *Oncogene* **31**, 4577-4587, doi:10.1038/onc.2011.621 (2012).
- 145 Hauptman, N. & Glavac, D. Long non-coding RNA in cancer. *Int J Mol Sci* **14**, 4655-4669, doi:10.3390/ijms14034655 (2013).
- 146 Zhang, H. *et al.* Long non-coding RNA: a new player in cancer. *J Hematol Oncol* **6**, 37, doi:10.1186/1756-8722-6-37 (2013).
- 147 Chiesa, N. *et al.* The KCNQ1OT1 imprinting control region and non-coding RNA: new properties derived from the study of Beckwith-Wiedemann syndrome and Silver-Russell syndrome cases. *Hum Mol Genet* **21**, 10-25, doi:10.1093/hmg/ddr419 (2012).
- 148 Cabianca, D. S., Casa, V. & Gabellini, D. A novel molecular mechanism in human genetic disease: a DNA repeat-derived lncRNA. *RNA Biol* **9**, 1211-1217, doi:10.4161/rna.21922 (2012).
- 149 Cai, B., Song, X. Q., Cai, J. P. & Zhang, S. HOTAIR: a cancer-related long non-coding RNA. *Neoplasma* **61**, 379-391, doi:10.4149/neo_2014_075 (2014).
- 150 Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071-1076, doi:10.1038/nature08975 (2010).
- 151 Ji, P. *et al.* MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **22**, 8031-8041, doi:10.1038/sj.onc.1206928 (2003).
- 152 Li, L. *et al.* Long Noncoding RNA MALAT1 Promotes Aggressive Pancreatic Cancer Proliferation and Metastasis via the Stimulation of Autophagy. *Mol Cancer Ther* **15**, 2232-2243, doi:10.1158/1535-7163.MCT-16-0008 (2016).
- 153 Zhang, Y., Dai, Q., Zeng, F. & Liu, H. MALAT1 Promotes the Proliferation and Metastasis of Osteosarcoma Cells By Activating the Rac1/JNK Pathway Via Targeting MiR-509. *Oncol Res*, doi:10.3727/096504017X14957939026111 (2018).
- 154 Zhou, Y. *et al.* The Long Noncoding RNA MALAT-1 Is Highly Expressed in Ovarian Cancer and Induces Cell Growth and Migration. *PLoS One* **11**, e0155250, doi:10.1371/journal.pone.0155250 (2016).
- 155 Nie, F. Q. *et al.* Long noncoding RNA ANRIL promotes non-small cell lung cancer cell proliferation and inhibits apoptosis by silencing KLF2 and P21 expression. *Mol Cancer Ther* **14**, 268-277, doi:10.1158/1535-7163.MCT-14-0492 (2015).
- 156 Zhang, E. B. *et al.* Long noncoding RNA ANRIL indicates a poor prognosis of gastric cancer and promotes tumor growth by epigenetically silencing of miR-99a/miR-449a. *Oncotarget* **5**, 2276-2292, doi:10.18632/oncotarget.1902 (2014).
- 157 Huarte, M. *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **142**, 409-419, doi:10.1016/j.cell.2010.06.040 (2010).
- 158 Hung, T. *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat Genet* **43**, 621-629, doi:10.1038/ng.848 (2011).

- 159 Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Mol Cell* **43**, 904-914, doi:10.1016/j.molcel.2011.08.018 (2011).
- 160 Faghihi, M. A. *et al.* Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med* **14**, 723-730, doi:10.1038/nm1784 (2008).
- 161 Hampel, H. & Shen, Y. Beta-site amyloid precursor protein cleaving enzyme 1 (BACE1) as a biological candidate marker of Alzheimer's disease. *Scand J Clin Lab Invest* **69**, 8-12, doi:10.1080/00365510701864610 (2009).
- 162 Tilgner, H. *et al.* Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res* **22**, 1616-1625, doi:10.1101/gr.134445.111 (2012).
- 163 Qureshi, I. A., Mattick, J. S. & Mehler, M. F. Long non-coding RNAs in nervous system function and disease. *Brain Res* **1338**, 20-35, doi:10.1016/j.brainres.2010.03.110 (2010).
- 164 Carrieri, C. *et al.* Expression analysis of the long non-coding RNA antisense to Uchl1 (AS Uchl1) during dopaminergic cells' differentiation in vitro and in neurochemical models of Parkinson's disease. *Front Cell Neurosci* **9**, 114, doi:10.3389/fncel.2015.00114 (2015).
- 165 Nishimoto, Y. *et al.* The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. *Mol Brain* **6**, 31, doi:10.1186/1756-6606-6-31 (2013).
- 166 Sunwoo, J. S. *et al.* Altered Expression of the Long Noncoding RNA NEAT1 in Huntington's Disease. *Mol Neurobiol* **54**, 1577-1586, doi:10.1007/s12035-016-9928-9 (2017).
- 167 Holdt, L. M. *et al.* ANRIL expression is associated with atherosclerosis risk at chromosome 9p21. *Arterioscler Thromb Vasc Biol* **30**, 620-627, doi:10.1161/ATVBAHA.109.196832 (2010).
- 168 Ishii, N. *et al.* Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. *J Hum Genet* **51**, 1087-1099, doi:10.1007/s10038-006-0070-9 (2006).
- 169 Michalik, K. M. *et al.* Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. *Circ Res* **114**, 1389-1397, doi:10.1161/CIRCRESAHA.114.303265 (2014).
- 170 Cremer, S. *et al.* Hematopoietic Deficiency of the Long Noncoding RNA MALAT1 Promotes Atherosclerosis and Plaque Inflammation. *Circulation* **139**, 1320-1334, doi:10.1161/CIRCULATIONAHA.117.029015 (2019).
- 171 Lekka, E. & Hall, J. Noncoding RNAs in disease. *FEBS Lett* **592**, 2884-2900, doi:10.1002/1873-3468.13182 (2018).
- 172 Kazemzadeh, M., Safaralizadeh, R. & Orang, A. V. LncRNAs: emerging players in gene regulation and disease pathogenesis. *J Genet* **94**, 771-784, doi:10.1007/s12041-015-0561-6 (2015).

- 173 Crooke, S. T., Witztum, J. L., Bennett, C. F. & Baker, B. F. RNA-Targeted Therapeutics. *Cell Metab* **27**, 714-739, doi:10.1016/j.cmet.2018.03.004 (2018).
- 174 Kaczmarek, J. C., Kowalski, P. S. & Anderson, D. G. Advances in the delivery of RNA therapeutics: from concept to clinical reality. *Genome Med* **9**, 60, doi:10.1186/s13073-017-0450-0 (2017).
- 175 Poller, W. *et al.* Non-coding RNAs in cardiovascular diseases: diagnostic and therapeutic perspectives. *Eur Heart J* **39**, 2704-2716, doi:10.1093/eurheartj/ehx165 (2018).
- 176 Arun, G. *et al.* Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. *Genes Dev* **30**, 34-51, doi:10.1101/gad.270959.115 (2016).
- 177 Gutschner, T. *et al.* The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res* **73**, 1180-1189, doi:10.1158/0008-5472.CAN-12-2850 (2013).
- 178 Khorkova, O., Hsiao, J. & Wahlestedt, C. Basic biology and therapeutic implications of lncRNA. *Adv Drug Deliv Rev* **87**, 15-24, doi:10.1016/j.addr.2015.05.012 (2015).
- 179 Matsui, M. & Corey, D. R. Non-coding RNAs as drug targets. *Nat Rev Drug Discov* **16**, 167-179, doi:10.1038/nrd.2016.117 (2017).