Turun yliopisto
University of Turku

GENETIC DISSECTION OF THE PATHOPHYSIOLOGY OF CARDIOVASCULAR RISK FACTORS

Ari Ahola-Olli

TURUN YLIOPISTON JULKAISUJA - ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. D osa - tom. 1297 | Medica - Odontologica | Turku 2017

Turun yliopisto
University of Turku

GENETIC DISSECTION OF THE PATHOPHYSIOLOGY OF CARDIOVASCULAR RISK FACTORS

Ari Ahola-Olli

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. D osa - tom. 1297 | Medica - Odontologica | Turku 2017

University of Turku

Faculty of Medicine

Cardiology and Cardiovascular Medicine

University of Turku Doctoral Programme of Clinical Investigation

Research Centre of Applied and Preventive Cardiovascular Medicine

Supervised by

Olli Raitakari, MD PhD
Professor in Cardiovascular Medicine
University of Turku
Turku, Finland

Niina Pitkänen, PhD
Senior Scientist
University of Turku
Turku, Finland

Johannes Kettunen, PhD
Associate Professor
University of Oulu
Oulu, Finland

Reviewed by

William Hennah, PhD
Docent in Human Genetics
Institute for Molecular Medicine Finland FIMM
Helsinki, Finland

Jussi Naukkarinen, MD PhD
Research Scientist
Institute for Molecular Medicine Finland FIMM
Helsinki, Finland

Opponent

Juha Kere, MD PhD
Professor in Molecular Genetics
Karolinska Institutet
Stockholm, Sweden

ISBN 978-951-29-6913-5 (PRINT)
ISBN 978-951-29-6914-2 (PDF)
ISSN 0355-9483 (PRINT)
ISSN 2343-3213 (PDF)
Painosalama Oy – Turku, Finland 2017

To my family

ABSTRACT

Ari Ahola-Olli
Dissecting the pathophysiology of cardiovascular risk factors by means of genetics

University of Turku, Faculty of Medicine, Cardiology and Cardiovascular Medicine, University of Turku Doctoral Programme of Clinical Investigation, Research Centre of Applied and Preventive Cardiovascular Medicine.

Annales universitatis Turkuensis, Medica-Odontologica, Turku, Finland, 2017.

Background: Genome-wide association studies (GWAS) have uncovered a plethora of biological information behind complex diseases. Findings from the GWAS and subsequent functional studies can facilitate drug development by identifying novel drug targets and predicting side-effects.

Aims: This thesis investigates the ability of genome-wide data to reveal potential mechanisms and treatments for cardiovascular diseases using a gene-environment interaction study, combined use of mice and human data, and detailed molecular methods involving intermediate phenotype data.

Participants and methods: This thesis is based on data from The Cardiovascular Risk Factors in Young Finns Study, the Special Turku Coronary Risk Factor Intervention Project, and national FINRISK surveys.

Results: The gene-environment design was able to identify one potential drug target for lowering LDL-cholesterol levels. The effect of dietary fat on serum LDL-cholesterol was more pronounced in subjects with rs9364628-TT. Combined analyses of mice and human data were able to identify the melanocortin 1 receptor as a potential drug target for improving cardiovascular health. By using cytokines as detailed molecular intermediate phenotypes, we were able to identify 27 loci (17 novel) associated with concentrations of circulating cytokines.

Conclusions: Combined used of animal and human data and use of detailed molecular intermediate phenotypes can be efficiently used to identify drug targets. More research is needed to elucidate whether gene-environment studies can be utilized in drug discovery.

Keywords: Genome-wide association study, drug development, cytokine, genetic, melanocortin, cardiovascular, cholesterol, gene-environment, inflammation.

TIIVISTELMÄ

Ari Ahola-Olli
Sydän- ja verisuonitautien riskitekijöiden patofysiologia – genetiikka tautimekanismien avaajana

Turun yliopisto, Lääketieteellinen tiedekunta, Kardiologia ja kardiovaskulaarilääketiede, Turun yliopiston kliininen tohtoriohjelma, Sydäntutkimuskeskus.

Turun yliopiston julkaisuja, Turku, Finland, 2017

Tausta: Viimeaikaiset edistysaskeleet genetiikassa ja tietokonealalla ovat mahdollistaneet genominlaajuisten assosiaatiotutkimusten toteuttamisen. Tutkimustulosten odotetaan auttavan lääkekehitystä.

Tavoite: Väitöskirjatyön tavoitteena on arvioida, kuinka ihmisiin keskittyvä geneettinen tutkimus voisi auttaa osoittamaan uusia lääkevaikutuskohteita sydän- ja verisuonitautien ja niiden riskitekijöiden hoitamiseksi.

Menetelmät: Aineistona käytetään Lasten Sepelvaltimotaudin Riskitekijät (LASERI) –tutkimusta, Sepelvaltimotaudin Riskitekijöiden Interventio Projekti -tutkimusta (STRIP) ja FINRISK-tutkimuksia.

Tulokset: Geeni-ympäristö – interaktiotutkimuksessa havaittiin yksi lokus (PARK2), jossa sijaitseva variantti sääteli ravinnon rasvahappokoostumuksen vaikutusta verenkierrossa vallitsevaan LDL-kolesterolikonsentraatioon. Ravinnon rasvan laadun vaikutus oli korostunut henkilöillä, joilla oli rs9364628-TT alleeli. Toisessa osatyössä hyödynnettiin kliinistä ja eläintutkimusta. Työssä voitiin osoittaa, että MC1R-reseptorin toiminta vaikuttaa verisuoniterveyteen hiirillä ja ihmisillä. Kolmannessa osatyössä tunnistettiin 27 lokusta (17 uutta), jotka olivat yhteydessä vähintään yhteen tutkituista 41 sytokiinista tai kasvutekijästä. Tuloksia voidaan hyödyntää kehitettäessä hoitomuotoja Crohnin tautiin, MS-tautiin, Behcetin tautiin´ja keliakiaan.

Johtopäätökset: Yhdistämällä dataa kliinisistä tutkimuksista ja eläintutkimuksista sekä tutkimalla yksityiskohtaisia molekyylitason fenotyyppejä voidaan löytää potentiaalisia lääkevaikutuskohteita. Lisätutkimuksia tarvitaan geeni-ympäristö –interaktiotutkimusten merkityksen arvioimiseksi lääkekehityksessä.

Avainsanat: Genominlaajuinen assosiaatiotutkimus, lääkekehitys, sytokiini, genetiikka, melanokortiini, sydän- ja verisuonitauti, kolesteroli, geeniympäristö, tulehdus.

TABLE OF CONTENTS

ABBREVIATIONS

ACTH	Adrenocorticotropic hormone
BMI	Body-mass index
BMS	Bare metal stent
bNGF	Beta nerve growth factor
CETP	Cholesterol ester transferase protein
CHARGE	The Cohorts for Heart and Aging Research in Genomic Epidemiology
CRP	C-reactive protein
CTACK	Cutaneous T-cell attracting (CCL27)
DES	Drug eluting stent
DNA	Deoxyribonucleic acid
E %	Energy percent
Eotaxin	Eotaxin (CCL11)
eQTL	Expression quantitative trait locus
FGFBasic	Basic fibroblast growth factor
FMD	Flow-mediated dilatation
GCSF	Granulocyte colony-stimulating factor
GROa	Growth regulated oncogene-α (CXCL1)
GWAS	Genome-wide association study
HGF	Hepatocyte growth factor
IFNg	Interferon-gamma
IL-10	Interleukin-10
IL-12p70	Interleukin-12p70
IL-13	Interleukin-13
IL-16	Interleukin-16
IL-17	Interleukin-17
IL-18	Interleukin-18
IL-1b	Interleukin-1-beta
IL1ra	Interleukin-1 receptor antagonist
IL-2	Interleukin-2
IL2ra	Interleukin-2 receptor, alpha subunit
IL-4	Interleukin-4
IL-5	Interleukin-5
IL-6	Interleukin-6
IL-7	Interleukin-7
IL-8	Interleukin-8 (CXCL8)

IL-9	Interleukin-9
IMT	Intima-media thickness
IP10	Interferon gamma-induced protein 10 (CXCL10)
LD	Linkage disequilibrium
LDL	Low-density lipoprotein
Lp(a)	Lipoprotein(a)
MAF	Minor allele frequency
MC1R	Melanocortin 1 receptor
MCP1	Monocyte chemotactic protein-1 (CCL2)
MCP3	Monocyte specific chemokine 3 (CCL7)
MCSF	Macrophage colony-stimulating factor
MIF	Macrophage migration inhibitory factor (glycosylation-inhibiting factor)
MIG	Monokine induced by interferon-gamma (CXCL9)
MIP1a	Macrophage inflammatory protein-1α (CCL3)
MIP1b	Macrophage inflammatory protein-1β (CCL4)
MUFA	Monounsaturated fatty acid
NMR	Nuclear magnetic resonance
PDGFbb	Platelet derived growth factor BB
PUFA	Polyunsaturated fatty acid
RANTES	Regulated on Activation, Normal T Cell Expressed and Secreted (CCL5)
SFA	Saturated fatty acid
SCF	Stem cell factor
SCGFb	Stem cell growth factor beta
SD	Standard deviation
SDF1a	Stromal cell-derived factor-1 alpha (CXCL12)
SE	Standard error
SNP	Single nucleotide polymorphism
STRIP	Special Turku Coronary Risk Factor Intervention Project
T2DM	Type 2 diabetes mellitus
TNFa	Tumor necrosis factor-alpha
TNFb	Tumor necrosis factor-beta
TRAIL	TNF-related apoptosis inducing ligand
VEGF	Vascular endothelial growth factor
YEM	Young's elastic modulus
YFS	The Cardiovascular Risk in Young Finns
α-MSH	α-melanocyte stimulating hormone

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications which are referred to in the text by Roman numerals I-III.

I) **Ahola-Olli A,** Pitkänen N, Kettunen J, Oikonen MK, Mikkilä V, Lehtimäki T, Kähönen M, Pahkala K, Niinikoski H, Kangas AJ, Soininen P, Ala-Korpela M, Viikari JS, Rönnemaa T, Simell O, Raitakari OT. Interactions between genetic variants and dietary lipid composition: effects on circulating LDL cholesterol in children. *The American Journal of Clinical Nutrition* 2014;100:1569-77.

II) Rinne P, **Ahola-Olli AV**, Nuutinen S, Koskinen E, Kaipio K, Eerola K, Juonala M, Kähönen M, Lehtimäki T, Raitakari OT, Savontaus E. Deficiency in Melanocortin 1 Receptor Signaling Predisposes to Vascular Endothelial Dysfunction and Increased Arterial Stiffness in Mice and Humans. *Arteriosclerosis, Thrombosis and Vascular Biology* 2015; 35:1678-86.

III) **Ahola-Olli AV,** Würtz P, Havulinna A, Aalto K, Pitkänen N, Lehtimäki T, Kähönen M, Lyytikäinen LP, Raitoharju E, Seppälä I, Sarin AP, Ripatti S, Palotie A, Perola M, Viikari JS, Jalkanen S, Maksimow M, Salomaa V, Salmi M, Kettunen J*, Raitakari OT*. Genome-wide association study identifies 17 new loci influencing concentrations of circulating cytokines and growth factors. *The American Journal of Human Genetics* 2017; 100:40-50.

1 INTRODUCTION

Cardiovascular diseases are the leading cause of death worldwide and they have thus attracted epidemiological studies and drug development. Epidemiological studies have identified multiple risk factors associated with the incidence of cardiovascular diseases, such as smoking, high low-density lipoprotein (HDL) concentration, low high-density lipoprotein (LDL) concentration, obesity and hypertension (Kannel et al. 1961; Selvin et al. 2014). Subsequent clinical trials have evaluated interventions aimed at reducing cardiovascular risk by modifying the risk factor levels. In order for interventions to work, the risk factors they are modifying need to be causally related to cardiovascular disease. However, even a robust association and carefully designed animal studies do not guarantee this causal relationship in humans: Despite the association between low HDL-cholesterol concentration and increased cardiovascular risk with supporting data from animal models, drugs that increase HDL-cholesterol concentration have failed to reduce cardiovascular events (T. Gordon et al. 1977; Hime et al. 2008; Keene et al. 2014). One HDL-cholesterol increasing drug (CETP-inhibitor torcetrapib) even increased the cardiovascular risk (Barter et al. 2007). Since clinical trials are expensive to conduct, the drug candidate entering the clinical testing pipeline must be carefully selected. Drug targets with supporting data from genetic studies done in humans have double the success rate compared to non-supported drug targets (Nelson et al. 2015). Choosing potential drug targets that already have support from genetic studies in humans could be used to decrease the financial costs of drug development. Furthermore, human genetics can be used to predict long-term side effects of pharmacological interventions (Sattar et al. 2010; Swerdlow et al. 2015). To evaluate the role of genetic studies in humans in identifying novel therapeutic targets we applied three different study designs. Study I of this thesis exploited gene-environment interaction models to unravel why there exists interindividual differences in how dietary fat quality affects circulating LDL-cholesterol concentration. Identifying these pathways might help in developing new treatments for dyslipidemia. In Study II we combined genetic studies in both human and mice to assess whether melanocortins could be used to improve vascular health. In humans, variants disrupting the melanocortin signaling were used as a model for melanocortin 1 receptor knock-out in mice. In study III, we used circulating cytokines and growth factors as intermediate phenotype to shed light to molecular pathways responsible for the development of diseases with an inflammatory component such as coronary heart disease, inflammatory bowel diseases, and multiple sclerosis.

In conclusion, we used three different study designs to evaluate how genetic studies in humans could be exploited to discover previously unknown biology and to provide information for drug development. In study I, we identified a possible gene-environment interaction between *PARK2* variants and dietary lipid quality. The second study design used *MC1R* signaling disrupting variants to mimic gene knock-out in humans. In terms of discovering novel drug targets and previously unknown biology, the approach exploiting intermediate phenotypes was the most successful. In Study III, we identified a variant influencing both the circulating concentration of interleukin-2 receptor subunit α and the risk for Crohn's disease, thus pinpointing a potential new drug target.

2 REVIEW OF LITERATURE

2.1 GENOME-WIDE ASSOCIATION STUDIES

2.1.1 THEORY BEHIND GENOME-WIDE ASSOCIATION STUDIES

2.1.1.1 The structure of the human genome

The human genome contains instructions for cells on how to construct all the different kinds of tissues in the human body and how to regulate tissue function. The genome can be found in the nucleus of every cell capable of cell division. A small part of the genome is also located within the mitochondria. The genome within the cell nucleus is divided into 23 chromosome pairs. The DNA molecule is a double-helix which is supported by a sugar phosphate backbone and the genetic code is determined by the sequence of base pairs. The four nucleotides forming these base pairs are adenine (A), thymine (T), cytosine (C) and guanine (G). The four nucleotides form pairs by binding to each other non-covalently through hydrogen bonds: A pairs with T and G pairs with C. There are about 3.3 billion of these pairs in the human genome. Most of the base pairs are conserved across individuals and others can vary. These differences are generated through mutations the majority of which are single-nucleotide polymorphisms (SNPs) where one nucleotide changes to another. Most SNPs have no effect on the phenotype but some of these polymorphisms can predispose to diseases.

Locus is a term used to refer to a specific location in the genome. Allele refers to the alternative forms of a locus that an individual can carry within the genome. Haplotype means a specific combination of nearby alleles. Most SNPs are diallelic having a minor allele (the rarer base) and a major allele (the more common base) (Bernstein et al. 2012; Bochud 2012). The chromosomes of present humans are mosaics of ancestral chromosomes that have arisen via multiple recombination events in the past (Dawson et al. 2002). The recombination events take place during meiosis when homologous chromosomes pair up. Two nearby SNPs can be tightly linked by physical distance to each other and thus are infrequently separated from each other by recombination. This phenomenon creates correlation between genetic variants. The aforementioned nonrandom co-occurrence of genetic vari-

ants is called linkage disequilibrium (LD) (D. Gordon and Finch 2005). Recombination, which tends to take place in recombination hotspots, creates a block-like LD pattern in the genome. These haplotype blocks vary in size, with most of them being less than 5 kilobases (kb) in size, but some might extend to over 50 kb and more. Due to the LD structure it is not necessary to genotype each variant within the genome separately and test them all for association: By genotyping only carefully selected variants that can tag the haplotype block will result in cost effective analyses with only a little loss of statistical power (Gabriel 2002).

2.1.1.2 Genotyping and imputation

Nowadays, DNA microarrays are used to measure the alleles of hundreds of thousands of DNA polymorphisms from the human genome. Grunstein and Hogness can be considered to be the developers of the first DNA array. They cloned DNA randomly into *Escherichia Coli* plasmids and grew these bacteria in colonies on a petri dish covered with nitrocellulose filter. Radiolabeled DNA probes were hybridized into colonies that were then lyzed and fixed on a nitrocellulose filter in order to search for DNA that was complementary to the labeled probe. In 1986, Miller and Barnes from Washington University School of Medicine described a method which allowed the detection of single-base changes in regions of known sequence (Miller and Barnes 1986). In 2000, a group led by David Walt working at the Tufts University developed a method whereby DNA was synthesized on polystyrene beads. Different kinds of fluorophore labels were used to identify which DNA sequence was synthesized on which bead (Ferguson, Steemers, and Walt 2000; Walt 2000). The number of possible fluorophore combinations limited the diversity of beads that could be created. To overcome this obstacle, the fluorescently labeled oligonucleotides were fixed on the beads and these were used to distinguish the different kinds of beads from each other (Gunderson et al. 2004). The technology, developed by the David Walt's group was licensed to Illumina, who have further developed the original method (Bumgarner 2001). The method currently used by Illumina is depicted in **Figure 1**.

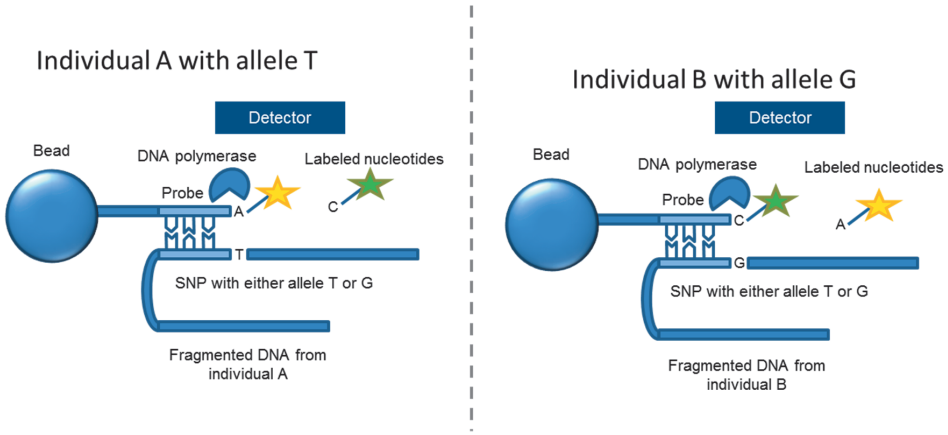


Figure 1 Schematic presentation of the Illumina BeadChip. Probes are attached on the surface of beads. The probes hybridize with target sequences and are subsequently elongated by DNA polymerase. The identification of the nucleotide added to the probe is based on a fluorescent label.

A typical modern DNA array contains probes for over 500 000 single-nucleotide changes, or polymorphisms. The entire human genome consists of about 3 289 000 000 (3.3 billion) nucleotides (Lander et al. 2001). Thus, only 0.015 % of the nucleotides within the genome are captured by the arrays. Genotype imputation is a term used to describe a computational method that predicts genotypes that are not directly detected in a study sample. The imputation process requires a haplotype reference panel which is used to predict the haplotypes of an individual. The first complete haplotype reference panel was produced by the HapMap Project and it included populations with ancestries from parts of Africa, Asia and Europe (International HapMap Consortium 2003, 2005). The HapMap Project was an extension of the Human Genome Project (Collins and McKusick 2001; International HapMap Consortium 2003). The Human Genome Project aimed to identify sequences that are common across individuals whereas one of the goals of HapMap Project was to identify variation within the genome across individuals. The HapMap Project focused on variants which have minor allele frequency over 5 %. The 1000 Genomes Project extended the work started by the HapMap Project by aiming to identify 95 % of all common variation with a minor allele frequency over 1 %. In addition, the 1000 Genomes Project catalogued even lower minor allele frequencies from coding regions (down towards 0.1 %) (The 1000 Genomes Project Consortium 2010). The 1000 Genomes Project produced a larger and more accurate reference panel which can be used to detect new associations from datasets already analyzed with HapMap reference panels (The 1000 Genomes Project Consortium 2010; Wood et al. 2013).

Because imputed genotypes are predicted, there is always some degree of uncertainty linked to them. Four measures that range from 0 to 1 have been described to quantify this uncertainty. The degree of uncertainty is further reflected in the association statistics. The imputation process is depicted in **Figure 2**. Imputing almost all common variants into the data by utilizing 1000 Genomes as a reference panel might reveal additional candidate causal variants that underlie each association (Wood et al. 2013). For example, *ORMDL3* have been previously linked to asthma, Crohn's disease, type 1 diabetes and rheumatoid arthritis. Another variant in the splice site of *GSDMB* is in the strong linkage disequilibrium with the *ORMDL3* variant; this is not captured by commercial DNA arrays and would, without imputation, remain undetected (The 1000 Genomes Project Consortium 2010). A similar example demonstrating the value of genotype imputation was the discovery that *APOL1* is a more likely causal gene for kidney disease than *MYH9,* which was originally considered as mediating the association signal in 22q12.3 (Genovese et al. 2010; Oleksyk et al. 2010).

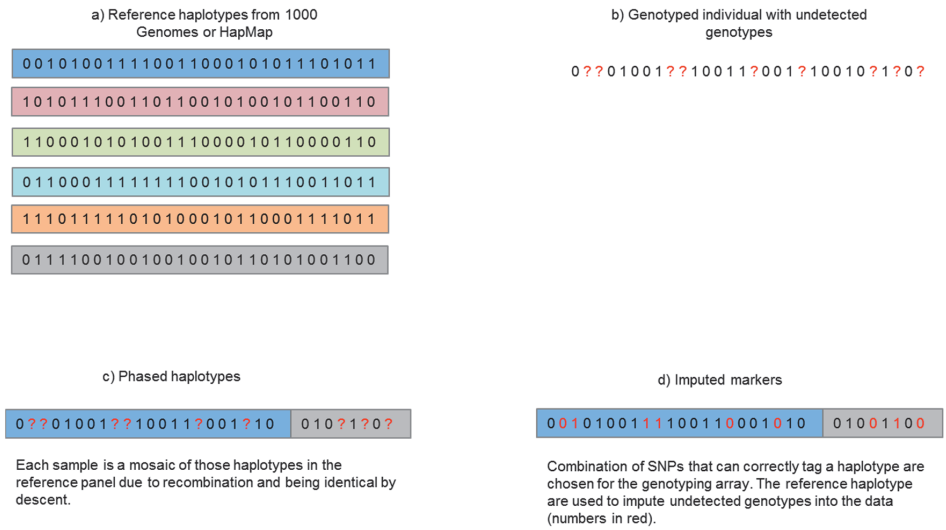


Figure 2 1000 Genomes project produced a haplotype map from a human genome which can be used to impute variants that are not directly genotyped with genotyping chip (a-b). Haplotypes are phased prior to imputation (c-d). Modified from Marchini et al. 2010 (Nature Reviews Genetics).

2.1.1.3 What can be found and cannot be found?

The GWASs based on HapMap imputed data and population samples relied on the so called common disease common variant hypothesis, which proposes that most

of the genetic risk is common: A person has a high risk for a complex disease because she or he has several disease predisposing common variants of small effect size in disease-associated loci. These studies were unlikely to detect variants with a minor allele frequency less than 5 % (Bodmer and Bonilla 2008). However, 1000 Genomes and UK10K projects have brought this detection limit down by introducing more accurate imputation panels. Using the same phenotype data as in HapMap studies, but with a new imputation reference, new low-frequency variants have been identified both from novel loci and from known disease loci (Walter et al. 2015; Wood et al. 2013). These variants are shared across the population due to being identical by descent and although each individual variant makes only a small contribution to the phenotype the combined effect of multiple risk variants is cumulative (**Figure 3** and **Figure 4**).

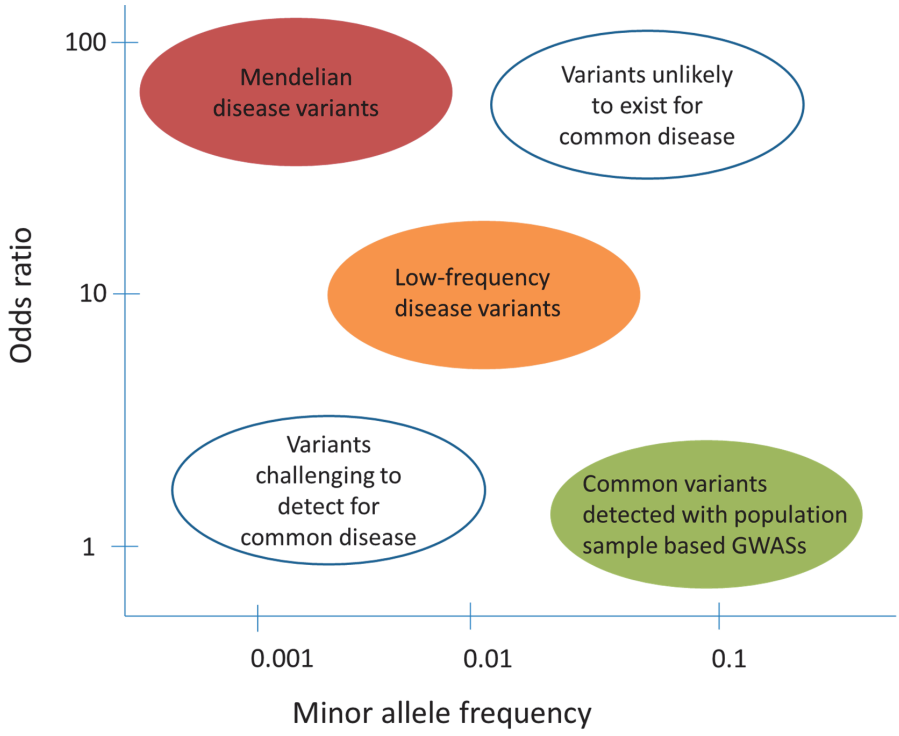


Figure 3 Genome-wide association studies are able to capture common and low-frequency variants. Common variants with large effect size are unlikely to exist for common diseases although they could be easily captured by genome-wide association studies. Currently, population sample based GWASs are not a cost-effective method to study Mendelian diseases. Modified from McCarthy et al. (Nature Reviews Genetics 2008).

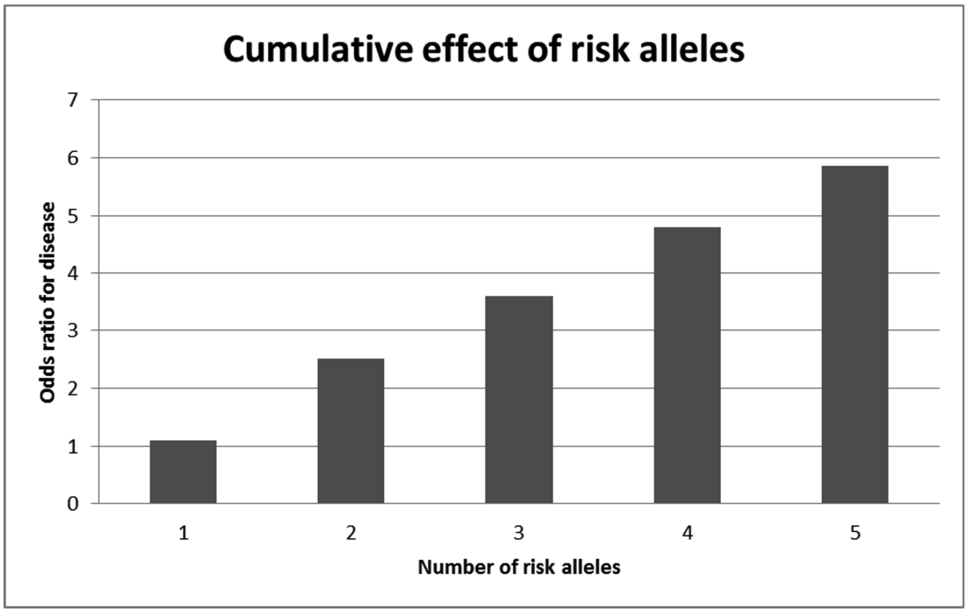


Figure 4 Schematic presentation of the additive effects of multiple disease predisposing variants on disease risk. Each variant has an individual effect size. Although individual variants confer a minimal effect on disease risk the combined effect of all disease associated variants can be substantial.

2.1.1.4 Effects of population structure and migration history

The distinct demographic history of Finland makes the Finnish population exceptionally suitable for genetic studies. Two separate founder populations have settled in Finland in two waves. The first wave arrived in inner Finland from the Lake Ladoga region 4000 years ago and the other came across the Gulf of Finland 2000 years ago (Kittles et al. 1998). Population bottle necks have also shaped the LD structure in Finland. The latest population bottle neck was experienced between 1690 and 1730 when the population declined from 600 000 to 250 000 (de la Chapelle 1993). These bottle necks increased the LD which then decayed again during further generations through meioses. The increased LD structure facilitates the detecting of disease-associated variants with genotyping chips. However, the genetic bottle necks experienced might also have caused Finns to lack some variants which are present elsewhere. For example, we lack variants causing phenyl ketonuria, a Mendelian disease that is more prevalent in other parts of Europe (Guldberg et al. 1995). On the contrary, the population history has created a group of hereditary diseases which are much more prevalent in Finland than in rest of the world. This group of 36 diseases is known as the Finnish disease heritage first

described in 1972 (Perheentupa 1972) and recently updated (Polvi et al. 2013). The incidence of these diseases in Finland ranges from 1 in 8000 to 1 in 100 000 (Polvi et al. 2013). Causal mutations for all of the 36 diseases have been successfully characterized. The first causal mutation was identified in 1988 (for gyrate atrophy of choroid and retina with or without ornithinemia; GACR, GA) and the latest in 2017 (for progressive encephalopathy with edema, hypsarrhythmia, and optic atrophy; PEHO) (Anttonen et al. 2017; Ramesh et al. 1988). In addition to certain Mendelian disease variants, the Finnish population is also enriched for low-frequency loss-of-function variants which can be exploited as a model for human knock-out (Chheda et al. 2017; Lim et al. 2014). However, these kinds of differences in variant distributions across populations might cause difficulties in replication of the GWAS findings.

2.1.1.5 Genetics of complex diseases

Coronary heart disease, hypertension, T2DM and dyslipidemia are all examples of complex diseases. In genetics, a complex disease means that there are multiple genetic and environmental factors contributing to a disease risk. Each individual variant makes a small contribution to the overall disease risk, but the combined effects of all the disease associated variants may explain a significant proportion of the disease risk in the population level. Although a single variant only has a minor effect on the phenotype identification of these individual variants, its location can reveal therapeutic pathways for pharmacological modification that can produce clinically meaningful changes in disease risk (Musunuru et al. 2010). Additionally, sequencing of GWAS loci harboring variants with small effect size can reveal novel rare variants with large effect size; these rare variants have a major impact on an individual level although their population attributable risk is small. For example, by sequencing GWAS loci the CHARGE consortium identified rare variants that were associated with 120 ms lengthening (+4.3 standard deviations) of PR interval (the time interval between the beginning of the atrial contraction and the beginning of the ventricular contraction) as measured from electrocardiograms (Magnani et al. 2014).

2.1.1.6 Genetics of coronary artery disease and major cardiovascular risk factors

Global Lipids Consortium have published the largest GWAS concerning circulating lipid levels to date; the study included a total of 188 577 subjects (Willer et al.

2013). The study confirmed the associations of 95 lipid loci and identified 62 new loci contributing to circulating concentrations of total cholesterol, LDL-cholesterol, HDL-cholesterol or triglycerides (Teslovich et al. 2010). Thus, the number of loci known to contribute to circulating lipid concentrations is now 157. The total variance (including both environmental and genetic variation) explained by all the known lipid loci is 15.0 % for total cholesterol, 14.6 % for LDL-cholesterol, 13.7 % for HDL-cholesterol and 11.7 % for triglycerides (Willer et al. 2013).

Obesity predisposes an individual to a variety of complex diseases including coronary heart disease and T2DM. A total of 97 loci have been associated with body-mass index (BMI) with a genome-wide significance (Locke et al. 2015). The total BMI variance explained by these loci is 2.7 % which is considerably less than the lipid loci explained by lipid variance, even though the study (Locke et al. 2015) included up to 339 224 subjects. The authors demonstrate that the SNPs with suggestive association with BMI ($p < 5 \times 10^{-4}$) are enriched in genes highly expressed in the nervous system. The second highest enrichment is seen in the hemic and immune system although the enrichment seen is not statistically significant.

T2DM predicts the incidence of new myocardial infarction as strongly as the history of prior myocardial infarction (Schramm et al. 2008). A total of 75 loci have been linked to T2DM in various populations (Cho et al. 2012; Kooner et al. 2011; Scott et al. 2007; Voight et al. 2010; Zeggini et al. 2008). Together these loci explain approximately 10 % of familial clustering of T2DM. Variants in *FTO*-locus are associated with both T2DM and BMI and the effect of these variants on T2DM susceptibility is mediated through BMI (Voight et al. 2010).

Hypertension is a highly heritable trait with heritability estimates being 0.57 for long-term systolic blood pressure and 0.56 for long-term diastolic blood pressure (Levy et al. 2000). SNPs from 28 loci have been linked to blood pressure in GWASs (Ehret et al. 2011). The 29 SNPs which form these 28 loci explain 0.9 % of the variation in systolic and diastolic blood pressure. The authors estimate that there are 116 (95 % CI 57-174) independent blood pressure variants that could explain 2.2 % of the total blood pressure variance. Thus a considerable amount of heritability still remains undetected. The genetic risk score calculated based on the identified variants is associated with stroke, coronary artery disease and left ventricular wall thickness, thus suggesting that hypertension is a causal factor for these cardiovascular end points (Ehret et al. 2011).

Coronary artery disease is a common cause of death in industrialized nations. A total of 58 loci have been linked to this disease explaining approximately 6 % of genetic variance in coronary artery disease susceptibility. Twelve of these loci

were associated with a lipid trait at p-value threshold < 1×10^{-4}. Among the lipid traits, the strongest association was detected with LDL-cholesterol concentration in 8 of the 12 loci (*APOB*, *ABCG5-ABCG8*, *PCSK9*, *SORT1*, *ABO*, *LDLR*, *APOE* and *LPA*) highlighting the important causal role of LDL-cholesterol in the development of coronary artery disease. Two of the 12 loci were associated with triglyceride concentration (*TRIB1* and *APOA5* cluster). Similarly, two of the loci were associated with T2DM (*MFGE8* and *SWAP70*) (Deloukas et al. 2013; Nikpay et al. 2015).

2.1.1.7 Mendelian randomization studies

The success of epidemiological research can be viewed in terms of specific individual health problems, such as lung cancer, or in a broader sense how it has altered our methods of dealing with non-communicable diseases (Kannel et al. 1961; Müller 1940). However, the inability of epidemiological research to make causal inference is a shortcoming: Despite careful epidemiological evaluations, some randomized controlled clinical trials have failed to demonstrate any benefit from therapies expected to work based on prior epidemiological knowledge (Hennekens et al. 1996). Some additional disadvantages of epidemiological studies include reverse causality (exposure X is associated with outcome Y but instead of X causing Y, the Y predisposes to X), selective loss to follow-up, and unknown confounding factors (Groenwold et al. 2016; Preston and Stokes 2014). These drawbacks can be partly overcome in Mendelian randomization studies. In Mendelian randomization, genetics variants are used as exposure indicators that are not subject to reverse causation and are less suspect to confounding factors (Davey Smith and Hemani 2014). Most importantly, Mendelian randomization allows causal inference if a genetic variant (or genetic variants), chosen as the instrument for exposure A, meets the following conditions:

1) The instrument is reliably associated with exposure A
2) The instrument is associated with the outcome only via exposure A
3) The instrument is independent of unobserved confounders that influence exposure A and the outcome after conditioning by observed confounders.

Since chromosomes segregate independently during meiosis, Mendelian randomization can be regarded as a lifelong randomized trial conducted by nature. Furthermore, individuals are unaware of their genotype which is analogous to blinding

in randomized clinical trials. In Mendelian randomization, the exposure can be a biomarker, such as a C-reactive protein, and an outcome can be, for example, a cardiovascular disease event. To avoid violating the aforementioned conditions it is necessary to select cis-acting variants near the *CRP* gene as instruments instead of those near the *APOC1* locus, which is also associated with the circulating CRP concentration. There is high suspicion that the *APOC1* locus might mediate its effects via lipids (violation of condition 2) (Dehghan et al. 2011). The validity of the whole method is based on careful selection of the genetic instrument. In addition, large sample sizes are needed since individual genetic variants typically explain less than 1 % of variation observed in exposure. By constructing a genetic score from multiple exposure-associated variants, it is possible to increase the power of the study considerably (S. Burgess 2014).

2.1.1.8 Genetics and drug discovery

Reliable information from drug development costs is difficult to obtain but cost estimates of 569.3 million US dollars have been suggested for a new cardiovascular disease targeting drug (DiMasi, Hansen, and Grabowski 2003; Light and Warburton 2005; Morgan et al. 2011). During the drug development process, 56 % of new drug candidates fail in Phase II or Phase III trials due to a lack of efficacy; this is after a considerable amount of money has already been invested in the drug development (Arrowsmith and Miller 2013). In order to make the drug development process more effective the drug developers have now started to search for new sources of information to complement conventional preclinical model systems. Genetic studies conducted in humans can be used as this kind of complementary information. The drugs approved for marketing in the United States and in the European Union are enriched with drugs targeting products of genes which have been associated with variation in human traits. Furthermore, from the total number of drugs under development, the proportion of drugs with supporting human genetic information has increased in the drug development pipeline. Thus, supporting genetic information can be considered as predictive of successful drug development. The lack of supporting genetic information has the greatest impact early on in the drug development process (Nelson et al. 2015). One specific example of a drug discovery guided by genetics is the invention of PCSK9 inhibitors (Stein et al. 2012). A locus causing autosomal familial hypercholesterolemia was first mapped to 1p34.1−p32 (Hunt et al. 2000) and later mutations in *PCSK9* were shown to cause the disease within this locus (Abifadel et al. 2003). The development of PCSK9 inhibitors for the prevention of cardiovascular disease was further

boosted by the discovery that *PCSK9* variants are associated with LDL-cholesterol concentration also in healthy population samples, and the *PCSK9* variants which lower LDL-cholesterol concentration are also inversely associated with coronary artery disease; this strongly suggested that drugs inhibiting PCSK9 could also prevent coronary artery disease (Willer et al. 2008). The ability of PCSK9 inhibitors to lower cardiovascular disease risk was later confirmed in a randomized controlled trial (Sabatine et al. 2017). Another example of the role of genetics in drug development is the Lp(a). By using circulating metabolites as an intermediate phenotype the research group was able to construct a genetic risk score that explained 45 % of variation in the circulating Lp(a). This genetic risk score was associated with coronary heart disease in hospital discharge registry based data. Furthermore, no other diseases were associated with the Lp(a) genetic risk score suggesting that therapies targeted to lower Lp(a) would not have major side effects (Kettunen et al. 2016).

2.2 GENE-ENVIRONMENT INTERACTIONS

The presence of gene-environment interaction means that the effect size of environmental exposure differs by genotype. After publication of the first large GWASs, concerns started to emerge due to the small proportion of heritability explained by the identified variants. It was suggested that gene-environment interactions might explain some of this missing heritability (Cole, Nikpay, and McPherson 2015). In addition to discovering missing heritability, gene-environment interaction studies might help to dissect the pathophysiology of complex exposures and help identify biological pathways mediating the effect of such exposures. For example, *NAT2* gene polymorphisms associated with slow detoxification of arylamines are associated with a higher risk of bladder cancer in smokers than in non-smokers (interaction between smoking status and *NAT2* polymorphism). Since arylamines are known carcinogens in tobacco smoke and are metabolized by *NAT2*, these results suggest that the increased risk of bladder cancer in smokers is mediated by arylamines rather than by many other carcinogens in tobacco smoke (for example polycyclic aromatic hydrocarbons) (García-Closas et al. 2005; Vineis and Pirastu 1997). In addition, variants altering function of proteins involved in elimination of xenobiotics in tobacco smoke modify the relationship between smoking and adverse birth outcomes, hence identifying a group of mothers whose children were particularly prone for adverse outcomes if their mothers smoke (X. Wang et al. 2002). Despite of the aforementioned examples, few gene-environment interactions have been robustly replicated to date (Qi et al. 2012, 2014). This might partly be due to sample size requirements which

are even more difficult to meet in gene-environment interaction studies than in GWASs (Thomas 2010).

2.3 GENETIC STUDIES IN ANIMALS AND HUMANS COMPLEMENT EACH OTHER

Knock-out animal models have been widely used in cardiovascular and other research to study the effects of specific signaling pathways on disease processes under interest (Boring et al. 1998; Cao et al. 2017; Rinne et al. 2017). However, even after careful planning of animal studies the results may not always be transferable to humans (Barter et al. 2007; Morehouse et al. 2007). This problem may be circumvented by using humans carrying loss-of-function alleles of gene under interest as a proxy for animal knock-out models (Valverde et al. 1995). Animal studies are still needed to obtain, for example, histological samples that cannot be obtained from humans (Rinne et al. 2015).

2.4 NUTRIGENOMICS IN CARDIOVACULAR DISEASE

As the causal role of LDL-cholesterol concentration in the development of coronary artery disease has been established, lifestyle habits lowering LDL-cholesterol concentration are of importance in maintaining population health. Dietary intake of fats has been shown to affect blood lipid concentrations (Niinikoski et al. 2012). Observational studies have also linked a higher intake of unsaturated fats to a lower risk of coronary artery disease (Oh et al. 2005). A recent systematic Cochrane review of randomized controlled trials showed that reducing dietary saturated fat intake reduces the risk of cardiovascular events by 17 %. A subgroup of studies, in this same review, demonstrated that the cardiovascular event reduction was seen in studies replacing saturated fats with polyunsaturated fats and no reduction was seen in studies replacing saturated fats with carbohydrates. It remains unclear whether replacing monounsaturated fats with polyunsaturated fats reduces cardiovascular mortality since the data is still sparse. The reduction in cardiovascular events was related to the degree of reduction in serum total cholesterol concentration (Hooper et al. 2015). This relationship between dietary fats and serum lipid concentration can be modified by genotype. In carriers of *APOE* ε2 allele, the higher intake of saturated fatty acids and cholesterol might paradoxically decrease LDL-cholesterol concentration (Lehtimäki et al. 1995).The effect of the *APOE* allele has also been extensively studied in diet intervention trials, and these have demonstrated that the diet has a greater

effect on circulating cholesterol concentration in subjects homozygous for the *APOE* ε4 allele (Tikkanen et al.). However, it remains unclear whether there exist other loci in the genome which modify the relationship between dietary fat intake and serum lipid concentration. The heritability of LDL-cholesterol concentration has been estimated to be around 0.534 and the known LDL-associated variants explain 14.6 % of total variation (Namboodiri et al. 1985). The previously described interactions between genetic variants and environmental exposures (dietary intake of lipids) might account for part of this missing heritability. Loci mediating this missing heritability might have implications for drug development.

2.5 MELANOCORTIN SIGNALING AND CARDIOVASCULAR DISEASE

Pharmacological targeting of melanocortin signaling has potential in preventing coronary heart disease (Rinne, Nordlund, et al. 2013; Rinne, Penttinen, et al. 2013). Melanocortins are a family of peptides that are cleaved from the common precursor, pro-opiomelanocortin. The family includes α-γ –melanocyte stimulating hormones (MSH) and adrenocorticotropic hormone (ACTH). The melanocortins can exert their biological actions through five different types of melanocortin receptors (MC1R-MC5R) (Mountjoy et al. 1992). The MC2R primarily binds only ACTH peptides, but the four other melanocortin receptors all bind α-MSH (Hruby et al. 1995). Melanotan II is a long-acting analogue of α-MSH that binds the same melanocortin receptors as α-MSH (Grieco et al. 2003; Hruby et al. 1995). Melanotan II treatment has been shown to reduce atherosclerotic plaque inflammation and to ameliorate endothelial dysfunction in LDL receptor-deficient, high-fat diet fed mice (Rinne et al. 2014). Since Melanotan II has an affinity to several types of melanocortin receptors it is currently unknown through which type of receptor the beneficial effects of melanocortins on vascular function are mediated. In addition, since the data is mainly derived from animal studies it is unknown whether the melanocortin signaling also contributes to vascular health in humans. Since skin pigmentation is mediated through MC1R, the genetic polymorphisms in *MC1R* have been associated with melanoma susceptibility (Barrett et al. 2011). In addition, *MC4R* variants have been associated with BMI, early-onset obesity, and waist circumference (Chambers et al. 2008; Wen et al. 2014; Wheeler et al. 2013). However, based on a GWAS catalogue search no associations between *MC1R-MC5R* variants and coronary artery disease have been reported.

2.6 INFLAMMATORY RISK FACTORS FOR CARDIOVASCULAR DISEASE

The INTERHEART study demonstrated that smoking, increased apolipoprotein B to apolipoprotein A1 ratio, hypertension, diabetes, abdominal obesity, psychosocial factors, daily consumption of fruit and vegetables, regular alcohol consumption, and regular physical activity accounted for 90 % of population attributable risk of acute myocardial infarction in men and 94 % in women (Yusuf et al. 2004). In addition to these nine risk factors, inflammation has also been implicated in the pathogenesis of cardiovascular disease (Kirii et al. 2003). In epidemiological studies, high-sensitivity C-reactive protein (hsCRP) has been associated with cardiovascular disease (Elkind et al. 2009; Makrygiannis et al. 2013). However, Mendelian randomization studies have questioned the causality of this association (Prins et al. 2016). Indeed, hsCRP might be a downstream biomarker of activated inflammatory response and the causal risk factors might reside in the upstream of the inflammatory cascade: Cholesterol crystals can be phagocytosed by human macrophages and these cholesterol crystals can induce interleukin-1beta (IL-1β) secretion in a dose-dependent manner (Rajamäki et al. 2010). IL-1β is a potent inducer of interleukin-6 (IL-6) in fibroblasts, endothelial cells, keratinocytes, and peripheral blood monocytes (Kirnbauer et al. 1989; Sironi et al. 1989; Tosato and Jones 1990). In addition to IL-1β, the IL-6 expression can also be increased by a tumor-necrosis factor alpha (TNF-α) (Sung et al. 2009). IL-6 (formerly known as interferon-β2) is able to initiate acute-phase response in liver cells. Stimulation of liver cells with IL-6 leads to the production of α1 acid lipoprotein, α1 proteinase inhibitor, α1 antichymotrypsin, haptoglobin, hemopexin, fibrinogen species, CRP, complement C3, complement factor B components, and serum amyloid A (Gauldie et al. 1987). Thus, it is possible that factors upstream from the circulating CRP are causally related to increased risk for cardiovascular disease. Supporting this notion, lack of IL-1β has been shown to decrease the atherosclerotic burden in *APOE*-deficient mice (Kirii et al. 2003). Furthermore, Mendelian randomization analyses support the causal role of IL-6 in the development of coronary artery disease (Niu et al. 2012). Currently, many clinical trials are exploring whether cardiovascular disease could be prevented by suppressing the inflammatory response. The IL-6 inhibitors tocilizumab and the TNF-α inhibitors adalimumab and infliximab are widely used in treatment of rheumatoid arthritis. Their role in treatment of cardiovascular disease may, however, be limited due to their adverse effects on the lipid profile (Di Minno et al. 2014). IL-1β signaling suppressing monoclonal antibody Anakinra (IL-1 receptor antagonist blocking both IL-1β and IL-1α) has been shown to improve left ventricular function and endothelium-dependent dilatation

of the brachial artery in patients with rheumatoid arthritis (Ikonomidis et al. 2008). A trial investigating the role of IL-1β blockage on markers of inflammation in non-ST elevation acute coronary syndrome is currently ongoing (Crossman et al. 2008). However, in Mendelian randomization studies concerns have also been raised regarding the cardiovascular safety of IL-1β blockade (Interleukin 1 Genetics Consortium 2015). Trials assessing the efficacy of low-dose methotrexate in the prevention of cardiovascular disease are also ongoing (Ridker 2009).

2.7 GROWTH FACTORS AND CARDIOVASCULAR DISEASE

Growth factors are protein molecules secreted by various types of cells. They instruct cell behavior, guide tissue repair and renewal (Mitchell et al. 2016). Due to their ability to guide tissue repair and renewal, growth factors hold great therapeutic potential. Randomized clinical trials investigating the role of VEGF in the treatment of coronary artery disease have already demonstrated that gene therapy targeting growth factors is effective in reducing angina pectoris symptoms in stable coronary heart disease. A transient improvement in oxygen supply has also been demonstrated in scintigraphy (Giusti et al. 2013; Kukuła et al. 2011).

Currently, percutaneous coronary intervention using balloon angioplasty and coronary stenting is a treatment of choice in acute coronary syndrome (Roffi et al. 2016; Windecker et al. 2014). Stenting is performed either with bare metal stents (BMS) or drug-eluting stents (DES) depending on the patient specific risk factors and site of coronary lesion. Bare metal stent requires shorter duration of dual antiplatelet therapy which is preferable in patients with a high bleeding risk (Levine et al. 2016). On the other hand, in-stent restenosis is more frequent after using BMS than DES (Ariotti et al. 2016; Ielasi et al. 2015). The in-stent restenosis may develop through neointimal hypertrophy or neoatherosclerosis (Looser, Kim, and Feldman 2016). Genetic variants within genes coding VEGF, basic fibroblast growth factor (FGF-basic), and basic platelet-derived growth factor (PDGFbb) might be associated with a risk of in-stent restenosis or narrowing of the stent lumen between follow-up visits (Osadnik et al. 2016). VEGF gene therapy administered during coronary angiography using adenovirus as the vector has been suggested as a therapy to prevent in-stent restenosis. The safety of the therapy has been demonstrated during a 10-year follow-up (Muona et al. 2012). However, data with hard clinical end-points are still lacking. Despite this, the growth factors have therapeutic potential for treatment of coronary heart disease and in-stent restenosis.

2.8 THE PATHOGENESIS OF ATHEROSCLEROSIS

Atherosclerosis is a lifelong process that causes narrowing of the arterial lumens which may eventually lead to clinically apparent coronary heart disease (McGill et al. 1997). Atherosclerosis results from lipid accumulation onto arterial wall and involves many stages. First, diffuse intimal thickening develops into atherosclerosis prone arteries including for example carotid artery, aorta, and coronary arteries. Diffuse intimal thickening is already present in 33 % of the 8-month old fetuses and it develops everyone before the age of 3 months. Diffuse intimal thickening is currently considered as a normal physiological process but it may play an important role in development of atherosclerosis since it is absent in atherosclerosis-resistant arteries (Nakashima et al. 2007). The American Heart Association has developed staging system for atherosclerotic lesions (Stary et al. 1994): Apolipoprotein B containing lipid particles and subsequently macrophages slowly build up in the pre-existing diffuse intimal thickenings (type I) and they eventually form the first grossly visible atherosclerotic lesions called fatty streaks (type II). The fatty streaks contain layers of macrophages filled with lipid droplets (foam cells). As the lesion progresses, lipids begin to accumulate to extracellular space (type III) (Nakashima et al. 2007; Stary et al. 1994). A Type IV lesion is a fully developed atheroma with a lipid core which is characteristic for this lesion type. The lipids in the core are thought to originate from dead foam cells (Nakashima et al. 2007). Type IV lesions are separated from arterial lumen by nearly normal intimal tissue. This tissue is later replaced by a fibrous cap (type V lesion). When the lesion surface gets ruptured and hematoma, hemorrhage, or thrombotic deposits are formed the lesion is classified as type VI lesion. The thrombus formed after lesion rupture may completely obstruct the blood flow symptoms of which depend upon the lesion cite within the vasculature and may include heart attack, stroke, or peripheral ischemia. Most of atherosclerosis mortality is due to ruptured type IV or type V lesions (Stary et al. 1995).

2.9 ASSESMENT OF PRECLINICAL ATHEROSCLEROSIS

Early atherosclerotic changes can be reliably measured by non-invasive vascular ultrasound methods (Raitakari et al. 2003). Measurements of arterial stiffness, intima-media thickness and endothelium dependent vasodilatation are methods currently used to evaluate the degree of preclinical atherosclerosis and vascular health.

2.9.1 CAROTID INTIMA-MEDIA THICKNESS

In 1986, Pignoli et al. demonstrated that histologically measured intima-media thickness (IMT) correlates strongly (correlation coefficient ranged from 0.76 to 0.82 depending on the severity of atherosclerotic changes in the specimen) with the thickness of the double line pattern seen in B-mode ultrasound. By dissecting and imaging structures of the arterial wall Pignoli et al. were able to show that the inner component of the double line pattern is caused by a lumen-intima interface and the outer component by a media-adventitia interface (Pignoli et al. 1986). The IMT measurement is thus performed simply by measuring the distance between the two echogenic lines seen in the arterial wall. These measurements can be considered as surrogate markers of atherosclerosis (de Groot et al. 2004) and each 0.10mm increment in IMT increases the risk of myocardial infarction by 12 to 15 % (Lorenz et al. 2007). Cardiovascular risk factors correlate with IMT and lipid lowering therapy can cause regression of IMT (Fleg et al. 2008; Raitakari et al. 2003; Wiegman et al. 2004). However, progression of IMT between two measurements has not currently been linked to cardiovascular events (Lorenz et al. 2012). Currently, IMT measurements are widely used to estimate the risk of cardiovascular disease.

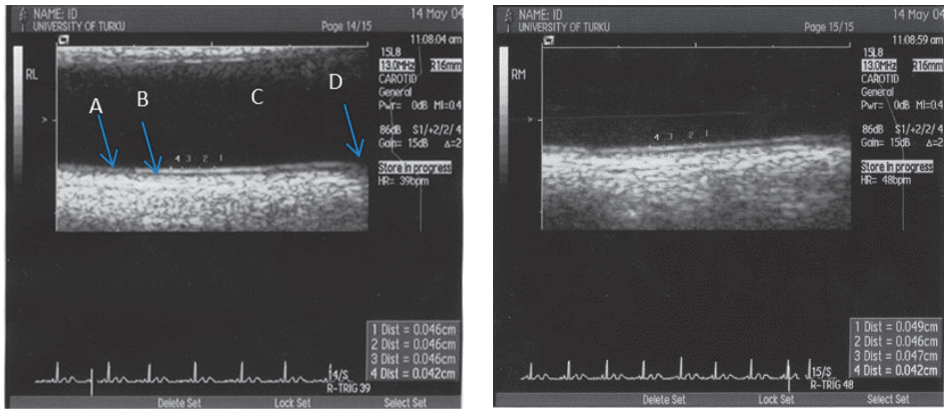


Figure 5 Ultrasound image obtained during measurement of carotid intima-media thickness. The image is taken just proximal to carotid bifurcation and the very beginning of the bifurcation can be seen on the right side of the image. A) Intima-lumen interface. B) Media-adventitia interface. C) Arterial lumen. D) Beginning of the bifurcation.

2.9.2 ARTERIAL ELASTICITY

In large arteries, such as the carotid artery, the arterial elasticity is mainly due to the high elastin to collagen ratio within the arterial wall. The arterial elasticity decreases during aging mainly due to progressive calcification, elastic fiber degeneration and an increase in collagen content (Avolio, Jones, and Tafazzoli-Shadpour 1998). Decrease in arterial elasticity between two measurements is greater in subjects with greater cardiovascular risk factor burden (Koskinen et al. 2012). In addition, decreased arterial elasticity is associated with cardiovascular events and all-cause mortality independently from other cardiovascular risk factors (van Sloten et al. 2014). The arterial stiffness can be measured ultrasonically by measuring carotid distensibility or Young's elastic modulus of the carotid wall. Carotid distensibility can be defined by measuring the difference in arterial diameter in end-diastole and end-systole and dividing the proportional diameter change by pulse pressure. The measurement of Young's elastic modulus is also based on change in the arterial diameter during cardiac cycle, and it also accounts for IMT.

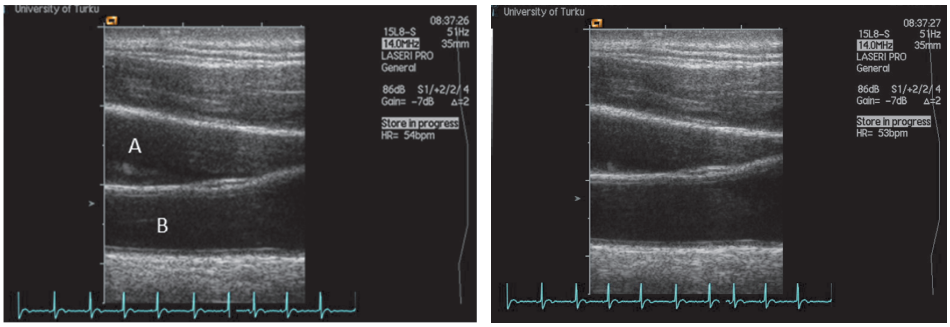


Figure 6 Measurement of the carotid elasticity in the Cardiovascular risk in the Young Finns Study. The picture on the left is taken at R-peak when the vessel diameter is at its maximum. The picture on the right is taken during diastole when the vessel diameter reaches its minimum value. The diameter is the distance between the intima-lumen interfaces. A) Internal jugular vein. B) Common carotid artery.

2.9.3 ENDOTHELIUM DEPENDENT VASODILATATION

In 1980, Furchgott and Zawadzki discovered that the endothelium is required for acetylcholine mediated vasodilatation (Furchgott and Zawadzki 1980). In 1985, Rubanyi et al. demonstrated that upon stimulation with acetylcholine the endothelial cells release vasoactive substance that was later named as the endothelium derived relaxing factor (EDRF) (Rubanyi, Lorenz, and Vanhoutte 1985). Further studies showed that the vasodilatation caused by EDRF was preceded by activation

of guanylyl cyclase and accumulation of cyclic guanosine monophosphate similar to the vasodilatation caused by nitrovasodilatators (Holzmann 1982). The exact chemical structure of EDRF remained unknown until 1987 when EDRF was recognized as nitric oxide (Ignarro et al. 1987). Nitric oxide production and vasodilatation of peripheral arteries is caused by hypoxemia (Doshi et al. 2001). The hypoxemia can be induced for example by compressing an upper limb by inflating a pneumatic cuff around the limb. The degree of vasodilatation caused by hypoxemia, or flow-mediated dilatation (FMD), can be considered as a measure of endothelial function and it can be measured with ultrasound after the release of pneumatic cuff. Decreased flow-mediated dilatation is associated with increased risk of cardiovascular disease events (Ras et al. 2013).

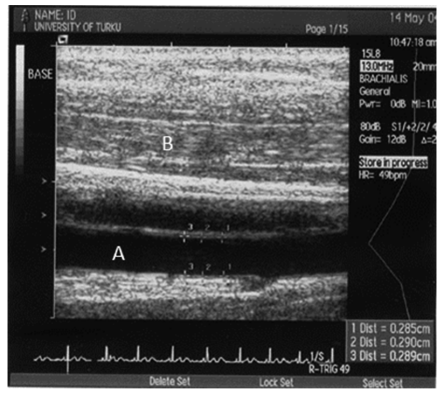


Figure 7 Measurement of flow-mediated dilatation from the brachial artery. A pneumatic cuff is inflated around the arm and subsequently released. The arterial diameter is measured before applying the cuff and after releasing at pre-specified time intervals. A) Brachial artery. B) Muscle tissue.

2.10 MULTIPLEX ANTIBODY ASSAYS

The advances made in assay technology are not limited to DNA assays as they also extend to the field of protein assays. Traditional 2-D gel electrophoresis is able to detect thousands of proteins, but this method is difficult to apply to large samples. Multiplexing allows simultaneous detection and quantification of single proteins from a solution of multiple proteins at a time. In multiplexing, capture antibodies are used to capture the analyte from the solution and a stained detection antibody is used to quantify the amount of bound analyte (**Figure 8**). Two kinds of antibody assays have been developed. Flat antibody assays are analogous to DNA assays where the capture antibody is attached to micro spots on a flat surface (Nielsen and

Geierstanger 2004). In bead assays, the capture antibodies are fixed on beads that are encoded, thus allowing the recognition of which capture antibody is bound to which bead (Evans, Sewter, and Hill 2003). Cross-reactivity has also been raised as a major concern in multiplex antibody assays. However, the assay manufacturers have taken much effort to reduce cross-reactivity when producing the antibodies. The use of separate detection and capture antibodies further reduces the risk of cross-reactivity(Juncker et al. 2014). The detailed molecular phenotypes obtained by multiplexing can help genetic studies to elucidate biochemical mechanisms that are responsible for the pathogenesis of complex diseases.

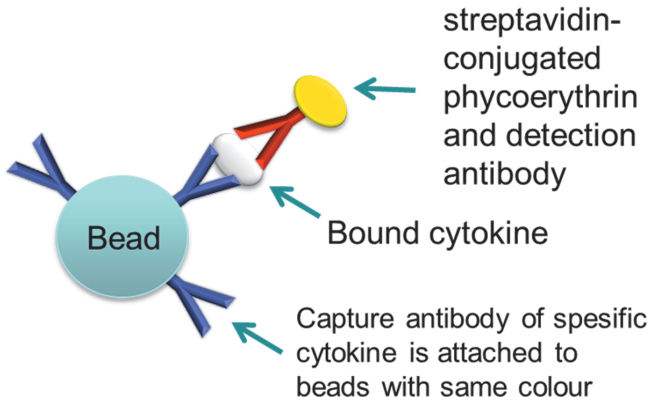


Figure 8 Schematic overview of a multiplex antibody assay. The bound analyte is recognized based on the bead color (or another physical property). Beads with the same color bind only one kind of analyte. The amount of bound analyte is measured via labeled detection antibody.

3 AIMS OF THE STUDY

This thesis is mainly based on findings from the Cardiovascular Risk in Young Finns study. In addition, subjects from The Special Turku coronary Risk Factor Intervention Project, FINRISK97, FINRISK02 and FINRISK07 have been included. The purpose was to employ genetic studies in humans to discover new drug targets for prevention and treatment of cardiovascular and inflammatory diseases.

The specific aims were:

1. To assess whether gene-environment interaction studies could be used to find drug targets for treatment of lipid disorders
2. To assess whether vascular changes observed in $MC1R^{e/e}$ mice could be replicated in humans carrying MC1R-signalling disrupting alleles.
3. To assess whether genetic variants are associated with concentration of circulating cytokines and whether the cytokine-associated variants predispose to complex diseases

4 MATERIALS AND METHODS

4.1 STUDY SAMPLES

4.1.1 THE CARDIOVASCULAR RISK IN YOUNG FINNS STUDY

4.1.1.1 Study participants

The Cardiovascular Risk in **Y**oung **F**inns **S**tudy (YFS) is an on-going multicenter follow-up study of Finnish subjects healthy at baseline. The subjects were randomly chosen from the Finnish cities of Helsinki, Tampere, Turku, Oulu, and Kuopio or their rural surroundings. The first cross-sectional survey was held in 1980 with 3596 subjects aged 3-18 years attending clinics. Follow-up studies that included physical examination, questionnaires, and measurements of known or suggested cardiovascular risk factors were conducted in 1983, 1986, 1989, 1992, 2001, 2007 and 2011.

4.1.1.2 Dietary records

Data from dietary intake at the age of 6 years was obtained by using 48-hour recall method. Food composition tables were used to extract information from the dietary intake of specific nutrients (Räsänen et al. 1985). At the 2007 follow-up, a food frequency questionnaire was used to gather information of dietary intake.

4.1.1.3 Biochemical analyses

The lipid measurements of Study I were performed in the Social Insurance Institution, Research and Development Center, Turku, Finland. Due to changes in determination methods and laboratory kits, the lipid concentrations were corrected to match lipid concentrations measured in 2001. The lipid measure corrections have been performed in two steps using regression equations. These equations were derived after re-analyzing a total of 275 frozen samples from earlier study visits. (Juonala et al. 2004; Porkka et al. 1997). In the first step, the year 1986 was chosen as the reference year because it was the most recent follow-up of total sample at

the time when the correction equations were derived. The laboratory reagents were changed in 1983 and thus measurements from 1980 were corrected to the new level with the equation:

Equation 1

$$Total\ cholesterol\ (1983)\left[\frac{mmol}{l}\right] = 0.997 \times Total\ cholesterol\ (1980)\left[\frac{mmol}{l}\right] + 0.287$$

The changes made in the total cholesterol measurement methods in 1983 affected the HDL-cholesterol measurements since the same method was used to measure HDL-cholesterol concentration after precipitation of ApoB containing particles.

Equation 2

$$HDL\ cholesterol(1983)\left[\frac{mmol}{l}\right] = 1.076 \times HDL\ cholesterol\ (1980) - 0.004\left[\frac{mmol}{l}\right]$$

No correction equations were needed for triglycerides in 1983 or 1986. In the second correction step, the obtained values were further corrected to the level of 2001with the following equations:

Equation 3

$$Total\ cholesterol\ (2001)\left[\frac{mmol}{l}\right] = 1.091 \times Total\ cholesterol\ (1983) - 0.271\left[\frac{mmol}{l}\right]$$

Equation 4

$$HDL\ cholesterol\ (2001)\left[\frac{mmol}{l}\right] = 1.068 \times HDL\ cholesterol\ (1983) - 0.0277\left[\frac{mmol}{l}\right]$$

Equation 5

$$Triglycerides\ (2001)\left[\frac{mmol}{l}\right] = 1.00756 \times triglycerides\ (1983) + 0.0582\left[\frac{mmol}{l}\right]$$

Values corrected to the level of 2001 were used in the primary analyses in study I.

4.1.1.4 NMR spectrometry

In study I, we assessed the interactions between genetic variants and serum LDL particle subclass concentrations measured in adulthood instead of childhood. This was because blood samples drawn from earlier follow-ups had deteriorated to a degree where reliable NMR measurements were not possible. The samples included in the study I, were drawn in 2007.

From each serum sample, two NMR spectra were recorded at a temperature of 37°C in a magnetic field of 500 MHz. Regression modeling was used to extract information from the NMR spectra to quantify three LDL subclasses, large, medium, and small (average diameters 25.5 nm, 23.0 nm and 18.7 nm, respectively).

4.1.1.5 Genotyping and imputation

The genotyping in the YFS study were performed with an Illumina 670k BeadChip containing a total of 670,000 SNPs and copy-number probes. The genotypes were called with Illumina's clustering algorithm. In study I, the imputation was performed with MACH software using HapMap release 22 as reference panel. The imputation resulted in a total of 2.5 million SNPs. In studies II and III, the imputation was performed with Impute2 software after pre-phasing with SHAPEIT software (Delaneau, Marchini, and Zagury 2012; B. N. Howie, Donnelly, and Marchini 2009). The 1000 Genomes September 2013 release was used as a reference panel resulting in ~10.6 million SNPs passing the quality control.

4.1.1.6 Measurements of gene expression

The gene expression was measured from peripheral blood cells drawn during the 2011 visit. After sample preparation, 1664 samples were analyzed with Illumina HT-12 v4 BeadChip containing 47231 expression probes. The expression data was quantile normalized and log2-transformed. Probes that were not expressed by over 5 % of the study subjects were excluded from further analyses resulting in a total of 19 637 included probes in study III.

4.1.1.7 Ultrasound studies

The non-invasive vascular ultrasound measurements were performed during the 2001 and 2007 follow-up visits with Sequoia 512 ultrasound mainframes (Acuson, CA, USA) with 13.0 MHz linear array transducers.

To measure the carotid intima-media thickness, the left carotid artery was scanned using a standardized protocol. The image was focused on the posterior wall of the carotid artery and an image sized 25 mm × 15 mm was recorded with a resolution box function. A video clip with a duration of 5 seconds including the beginning of carotid bifurcation and the common carotid artery, was stored in digital format for

subsequent offline analyses. These stored scans were analyzed by a single reader who was blinded to the subjects' details. From the 5 second video clip, at least four measurements were taken approximately 10 mm from the carotid bifurcation. Mean and maximal carotid IMTs were determined from these measures. To assess intra-individual reproducibility of the IMT measures, 57 subjects were re-examined 3 months after the study visit. To assess intra-observer reproducibility, the same reader interpreted these repeated ultrasound scans. The coefficient of variation was 6.4 % between visit, and the intra-observer coefficient of variation was 3.4 %.

4.1.1.8 Determining the cytokine concentrations

Concentrations of 48 cytokines and growth factors were measured from blood samples drawn during the 2011 follow-up visit. Bio-Rad's premixed Bio-Plex Pro Human Cytokine 27-plex and 21-plex assays were used in quantification. The assays were performed according to the manufacturer's instructions, except for the number of beads, where detection antibodies, and streptavidin-phycoerythrin conjugate were used at 50 % lower concentrations than recommended (Ritchie et al. 2015).

4.1.2 SPECIAL TURKU CORONARY RISK FACTOR INTERVENTION PROJECT

4.1.2.1 Study participants

The **S**pecial **T**urku coronary **R**isk factor **I**ntervention **P**roject (STRIP) started with a small pilot project (n=45) in October 1989. The main trial began half a year later and recruiting was discontinued in May 1992. Recruitment was performed during the 5-month visit at national well-baby clinics by nurses in the city of Turku, Finland. The study sample consisted of 1062 infants from 1054 voluntary families. After the pediatrician had explained the purpose and the design of the trial at the University of Turku, the children were randomized to an intervention group (n=540) or to a control group (n=522). The intervention group received individualized nutritional counseling targeting an optimal diet. The optimal diet was defined as containing 30-35 energy percent (E%) of fat with polyunsaturated:mono-

unsaturated:saturated fat ratio of 1:1:1, 10-15 E% of protein and 50-60 E% of carbohydrates. Using oils and soft margarines in food preparation was encouraged. Detailed instructions were given to replace saturated fat in the diet, for example, with leaner meat products or low-fat cheese. Serving fish at one of the main meals once or twice a week was recommended.

In study I, the randomization into intervention and control groups was omitted to better account for the actual dietary intake of the children. Instead of comparing intervention and control groups, we divided the study sample into two groups based on dietary fat quality. Subjects with an unsaturated-to-saturated fat ratio below the age group median were assigned to a bad fat quality group and subjects with a ratio above the median were assigned to a good fat quality group. Since puberty is known to affect circulating lipid concentration, only 5- and 7-years old STRIP participants were included in the study.

4.1.2.2 Dietary records

The nutritionists instructed the parents on how to correctly record food intake. Written instructions and drawings of ordinary foods were given to parents to help estimate the amount of food intake. The parents biannually kept a 4-day food record of their child's food intake. The nutritionist reviewed the records during the study visits and nutrient intakes were analyzed with Micro-Nutrica software. The software uses Food and Nutrient Database of the Social Insurance Institution and calculates 66 nutrients in over 650 foods and dishes that are most commonly used in Finland. The software is constantly updated and it provides an accurate estimation of the intake of many nutrients, including fats (Hakala et al. 1996). Supplemental vitamins and minerals were not included in the analyses.

4.1.2.3 Biochemical analyses

The STRIP blood samples were drawn in the fasting state at the age of 5 years and thereafter. All laboratory analyses were performed at the National Public Health Institute (currently called the National Institute for Health and Welfare) in Turku, Finland. After clotting and centrifugation, serum was separated and stored at -25 degrees of Celsius. Serum cholesterol concentration was determined with the enzymatic cholesterol oxidase p-aminophenazone method (Merck, Darmstadt, Germany) with an AU 510 automatic analyzer (Olympus, Hamburg, Germany). ApoB-containing lipoprotein particles were precipitated with dextran sulfate to

measure serum HDL-cholesterol concentration. The serum triglyceride concentration was measured by the colorimetric glycerol-3-phospate oxidase p-aminophenazone method (Merck) with an automatic Olympus AU 510 analyzer. The inter-assay coefficients of variation for HDL and total cholesterol were 1.9 % and 2.0 %, whereas the intra-assay coefficients of variation were 1.2 % and 1.5 %, respectively. The LDL-cholesterol concentration was estimated with the Friedewald Formula (Friedewald, Levy, and Fredrickson 1972).

4.1.2.4 Genotyping

A total of 666 STRIP participants were genotyped with Illumina Metabochip including 196 725 single nucleotide polymorphisms. The Metabochip is a targeted genotyping platform designed to capture the variation in genes with suggested role in development of cardiovascular disease –related traits. The targeted loci were selected based on GWAS meta-analyses of 23 cardiovascular disease –related traits (Voight et al. 2012). The DNA was extracted from peripheral blood leukocytes. To account for errors in the sample processing, individuals who had more than 5% genotype data missing were excluded. SNPs with minor allele frequency < 5 % and with more than 5 % of genotype data missing were excluded. To account for genotyping errors and population stratification SNPs deviating from the Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$) were excluded. Applying these filters 125 848 markers passed the quality control.

4.1.3 THE NATIONAL FINRISK STUDY

4.1.3.1 Study participants

FINRISK surveys are national cross-sectional surveys performed every five years to monitor risk factors for chronic diseases. Each survey is representative of the working age population of Finland and includes subjects aged 25-74 years. The subjects are drawn from population registries from five geographical areas of Finland: North Karelia, North Savo, Oulu and Lapland, the cities of Turku and Loimaa and their surroundings, and the cities of Helsinki and Vantaa. FINRISK surveys have been conducted in 1972, 1977, 1982, 1987, 1992, 1997, 2002, 2007, and 2012. At each study visit a blood sample is drawn and a clinical examination is

performed: in addition, a questionnaire is included. From a subset of FINRISK2007 participants living in the Helsinki area, a blood sample was drawn to measure mRNA expression profiles.

4.1.3.2 Biochemical analyses

The cytokine measurements from FINRISK participants were conducted as described for the YFS cohort.

4.1.3.3 Genotyping and imputation

Different genotyping chips have been used to acquire genome-wide data from subsets of FINRISK1997, FINRISK2002 and FINRISK2007 participants. The FINRISK1997 participants were genotyped with Illumina Core Exome Chip and the genotype calling was performed with Illuminus algorithm (Teo 2012). Impute2 was used as the imputation software (B. Howie, Marchini, and Stephens 2011). In FINRISK2007, two different genotyping chips were used: Illumina Omni Express and Illumina 610k. From both chips, the genotypes were called with Illuminus algorithm and the imputation was performed with Impute2 software using 1000 Genomes' September 2013 release as reference.

4.1.3.4 Measurements of gene expression

A total of 518 samples from FINRISK2007 were analyzed with Illumina HumanHT-12 version 3 Expression BeadChips. Probes not autosomal, were complementary to cDNA from erythrocyte globin components and those mapped to more than one genomic position were excluded. After filtering, 35 419 probes were included in the analyses (Inouye et al. 2010).

4.1.4 EXPERIMENTAL STUDIES: THE RODENTS

$MC1R^{e/e}$ mice, which lack functional MC1R signaling due to single base pair deletion and their wild-type control littermates, were obtained from Jackson Laboratory (strain# 000060, Bar Harbor, Maine, USA). All experiments were conducted on male adult mice aged 4 to 6 months. The animals were housed in a 12-hour

light/dark cycle and fed on a regular chow diet. All animal experiments were approved by the National Animal Experiment Board. In all the experiments, the mice were euthanized with carbon dioxide asphyxiation, the blood was withdrawn and thoracic aortae were harvested for further experimentation.

4.2 STATISTICAL ANALYSES

The statistical analyses were performed with SAS versions 9.2 or 9.3 (SAS institute, Inc., Cary, NC) unless otherwise stated. The distributions of continuous variables and assumptions of normality were assessed by examining histograms and estimates of skewness and kurtosis. In study III, large deviations from normality were observed in cytokine distributions and thus inverse normal transformation was used.

4.2.1 STUDY I

To describe the characteristics of study participants', mean and standard deviations were calculated for the main variables used in the study. To compare differences between boys and girls, a 2-tailed t-test was used. To assess whether genetic variants modulated the association between the quality of fats in diet and circulating LDL-cholesterol concentration, linear interaction models were used. LDL-cholesterol concentrations at the age of 5 years and at the age of 7 years were used as dependent variable. The dietary fat quality was described as the ratio of unsaturated and saturated fats. In order to include a locus in the replication stage, a variant within the locus needed to have an interaction term p-value (dietary fat quality × SNP) of less than 0.01 with consistent effect directions in both age groups.

4.2.2 STUDY II

To assess the effects of MC1-R signaling on vascular health in humans, ultrasound data from 2001 and 2007 were used. An additive genetic score was constructed based on five variants previously linked to weak MC1R signaling. To account for intra-individual between measurements correlation, mixed models were used in the modeling of longitudinal data. The models were adjusted for age, sex, and BMI. Based on Akaike's information criterion, compound symmetry was selected as the covariance structure. To assess whether the changes observed in the hemodynamic

profile affect arterial ultrasound measurements, the model was further adjusted for mean arterial pressure. Statistical significance was inferred at a two-tailed P-value of less than 0.05.

4.2.3 STUDY III

The cytokine concentrations were adjusted for age, sex, body-mass index and the first 10 genetic principal components prior to GWAS by fitting a linear model to the data and using inverse transformed model residuals as a phenotype in GWAS. The GWASs were performed with SNPTEST version 2.4beta software (Marchini and Howie 2010). SNPs with minor allele count < 10, info < 0.7, and model fit statistics < 0.7 were excluded from further analyses. Meta-analyses were performed with METAL software (Willer, Li, and Abecasis 2010). SAS software and linear models with age and sex as covariates were used to assess whether cytokine-associated SNPs have eQTLs in whole blood. In the meta-analyses, statistical significance was inferred at a two-tailed p-value of less than 1.2×10^{-9} (5×10^{-8} divided by 41 which is the number of cytokines included to the study). Despite correlations between the concentrations of various cytokines, this conservative threshold was selected to avoid false-positive associations.

4.3 ETHICAL STATEMENT

The study participants (or their parents) gave written informed consent and the studies have been approved by local ethics committees.

5 RESULTS

5.1 STUDY I

The mean circulating LDL-cholesterol concentrations in 5-year-old boys and girls were 2.87 ± 0.63 mmol/l and 3.08 ± 0.66 mmol/l, respectively. There was statistically significant difference between boys and girls ($P < 0.0001$). The dietary fat quality, as described by the ratio of monounsaturated and polyunsaturated fat to saturated fats ((PUFA + MUFA):SFA), was 1.26 and 1.25 in 5-year-old boys and girls, respectively. By the age of 7 years, the dietary fat quality had improved to 1.35 in boys and to 1.30 in girls (**Table 1)**. These age groups were selected to discount the effect of puberty on circulating lipid concentration.

Table 1 STRIP participant characteristics

Variable	Girls Mean ± SD	n	Boys Mean ± SD	n
Age 5				
Height (cm)	111 ± 4	303	112 ± 4	339
Weight (kg)	19.3 ± 2.9	304	19.4 ± 2.3	339
LDL-cholesterol (mmol/l)	3.08 ± 0.66	284	2.87 ± 0.63	319
Total cholesterol (mmol/l)	4.62 ± 0.75	284	4.41 ± 0.73	319
HDL-cholesterol (mmol/l)	1.22 ± 0.24	284	1.24 ± 0.26	319
Triglycerides (mmol/l)	0.67 ± 0.22	284	0.63 ± 0.24	319
Dietary fat quality ((PUFA+MUFA)/SFA)	1.25 ± 0.32	279	1.26 ± 0.31	302
Dietary SAFA (E %)	13.0 ± 3.0	279	13.0 ± 2.7	302
Dietary Palmitic acid (E %)	6.29 ± 1.31	279	6.30 ± 1.18	302
Dietary strearic acid (E %)	2.67 ± 0.59	279	2.71 ± 0.60	302
Dietary linolic acid (E %)	3.93 ± 0.93	279	4.07 ± 1.07	302
Dietary α-linolenic acid (E %)	0.82 ± 0.25	279	0.81 ± 0.24	302
Age 7				
Height (cm)	124 ± 5	306	125 ± 5	336
Weight (kg)	25.0 ± 4.6	306	24.6 ± 3.5	336
LDL-cholesterol (mmol/l)	3.07 ± 0.68	298	2.84 ± 0.59	331
Total cholesterol (mmol/l)	4.72 ± 0.77	298	4.46 ± 0.68	331
HDL-cholesterol (mmol/l)	1.32 ± 0.23	298	1.32 ± 0.26	331
Triglycerides (mmol/l)	0.69 ± 0.27	298	0.64 ± 0.28	331
Dietary fat quality ((PUFA+MUFA)/SFA)	1.30 ± 0.30	274	1.35 ± 0.31	295
Dietary SAFA (E %)	12.8 ± 2.6	274	12.5 ± 2.5	295
Dietary Palmitic acid (E %)	6.11 ± 1.19	274	6.08 ± 1.12	295
Dietary strearic acid (E %)	2.62 ± 0.58	274	2.65 ± 0.56	295
Dietary linolic acid (E %)	4.18 ± 1.10	274	4.30 ± 1.18	295
Dietary α-linolenic acid (E %)	0.88 ± 0.28	274	0.87 ± 0.28	295

Table 1. Study cohort characteristics from STRIP in two age points.
*Dietary fat quality is expressed as unsaturated to saturated fats ratio. Modified from original publication I.

To assess whether genetic polymorphisms modulate the relationship between dietary fat quality and circulating LDL-cholesterol concentration, genetic interaction analyses were performed. A total of 21 SNPs had statistically significant interaction and consistent direction of effect in both age groups (**Table 2)** (Ahola-Olli et al. 2014). The main effects for these SNPs are depicted in **Table 3**.

Table 2 Results from the discovery analyses

						Age 5		Age 7	
Locus	**SNP**	**Position**	**Alleles**	**MAF**	**Gene**	**Beta**	**P**	**Beta**	**P**
2q11.1	rs2579547	96759437	A/G	0.46	*ADRA2B*	-0.405	5.1×10^{-4}	-0.326	4.9×10^{-3}
2q11.1	rs11687113	96658588	A/G	0.44	*ADRA2B*	-0.351	2.5×10^{-3}	-0.327	4.8×10^{-3}
2q33.3	rs17537366	208768540	A/G	0.13	*PLEKHM3*	-0.465	7.8×10^{-3}	-0.554	1.7×10^{-3}
3p14.2	rs9833522	63131003	C/T	0.09	*SYNPR*	1.192	1.8×10^{-3}	1.013	6.9×10^{-3}
3p25.1	rs13083662	16282442	C/G	0.09	*GALNTL2*	0.597	4.0×10^{-3}	0.528	9.7×10^{-3}
3q21.3	rs7428297	127078340	C/T	0.03	*LOC101927123*	-0.541	5.7×10^{-3}	-0.618	1.2×10^{-3}
3p24.2	rs9838720	24590744	A/G	0.10	*OXSM*	0.55	7.1×10^{-3}	0.722	3.2×10^{-4}
4q26	rs9995961	117785341	C/T	0.22	*TRAM1L1*	-0.365	8.7×10^{-3}	-0.415	2.5×10^{-3}
5q23.2	rs9327448	126975062	G/T	0.03	*PRRC1*	0.512	2.5×10^{-3}	0.441	9.1×10^{-3}
5p14.3	rs10065818	18877165	C/T	0.13	*CDH10*	-1.049	5.1×10^{-3}	-0.937	8.3×10^{-3}
6q25.3	rs3777392	160556642	A/G	0.20	*SLC22A1*	0.456	1.1×10^{-3}	0.464	8.8×10^{-4}
6q26	rs9364628	162339286	C/T	0.46	*PARK2*	0.302	8.4×10^{-3}	0.319	5.7×10^{-3}
6q26	rs9458419	162344712	C/T	0.42	*PARK2*	-0.302	9.6×10^{-3}	-0.301	9.4×10^{-3}
7q21.13	rs10487028	88263202	G/T	0.43	*STEAP1*	0.31	1.0×10^{-2}	0.461	1.2×10^{-4}
8q12.3	rs17179841	63704368	C/T	0.13	*NKAIN3*	0.733	5.9×10^{-3}	0.946	5.7×10^{-4}
8p21.3	rs4557718	19890654	C/T	0.05	*LPL*	-0.472	8.5×10^{-3}	-0.474	9.0×10^{-3}
11q13.4	rs7116713	74276271	C/T	0.04	*LIPT2*	-0.939	2.7×10^{-3}	-0.892	9.0×10^{-3}
11q23.1	rs522188	111040593	A/G	0.13	*COLCA1*	-0.461	8.4×10^{-3}	-0.484	4.6×10^{-3}
12p12.2	rs17376366	20492724	C/T	0.20	*SLCO1B3*	0.378	7.3×10^{-3}	0.365	9.3×10^{-3}
16q23.1	rs11150015	77827259	A/G	0.14	*KIAA1576*	-0.454	6.0×10^{-3}	-0.441	7.7×10^{-3}
22q13.1	rs2267366	38480659	C/T	0.17	*BAIAP2L2*	0.442	4.2×10^{-3}	0.425	6.2×10^{-3}

Two interaction analyses were performed in the Special Turku Coronary Risk Factor Intervention Project (STRIP) at two age points with a low density lipoprotein (LDL) -cholesterol as the phenotype. P-values and betas are for SNP×fat quality group interaction. The position refers to GRCh37/hg19. From original publication I.

Table 3 Main effects for SNPs mentioned in Table 2

			Age 5				Age 7			
Locus	**SNP**	**Alleles**	**Beta (FQ)**	**P (FQ)**	**Beta (SNP)**	**P (SNP)**	**Beta (FQ)**	**P (FQ)**	**Beta (SNP)**	**P (SNP)**
2q11.1	rs2579547	A/G	0.212	0.10	0.642	5.7×10^{-4}	0.135	0.30	0.481	1.0×10^{-2}
2q11.1	rs11687113	A/G	0.180	0.18	0.536	4.0×10^{-3}	0.149	0.27	0.490	8.5×10^{-3}
2q33.3	rs17537366	A/G	-0.025	0.79	0.681	1.4×10^{-2}	-0.018	0.84	0.752	6.5×10^{-3}
3p14.2	rs9833522	C/T	-0.207	1.3×10^{-2}	-1.445	1.4×10^{-2}	-0.205	1.0×10^{-2}	-1.313	2.8×10^{-2}
3p25.1	rs13083662	C/G	-0.244	5.9×10^{-3}	-0.995	3.0×10^{-3}	-0.240	6.4×10^{-3}	-0.788	1.2×10^{-2}
3q21.3	rs7428297	C/T	-0.046	0.61	1.056	7.6×10^{-4}	-0.034	0.70	1.068	5.5×10^{-4}
3p24.2	rs9838720	A/G	-0.243	6.5×10^{-3}	-0.878	6.7×10^{-3}	-0.278	1.6×10^{-3}	-1.023	1.0×10^{-3}
4q26	rs9995961	C/T	0.011	0.91	0.614	4.6×10^{-3}	0.028	0.78	0.674	2.4×10^{-3}
5q23.2	rs9327448	G/T	-0.284	2.3×10^{-3}	-0.896	9.4×10^{-4}	-0.268	3.7×10^{-3}	-0.739	5.9×10^{-3}
5p14.3	rs10065818	C/T	-0.104	0.21	2.099	8.4×10^{-4}	-0.094	0.26	1.645	2.1×10^{-2}
6q25.3	rs3777392	A/G	-0.342	7.0×10^{-4}	-0.727	1.0×10^{-3}	-0.342	5.8×10^{-4}	-0.747	8.3×10^{-4}
6q26	rs9364628	C/T	-0.421	1.4×10^{-3}	-0.435	1.6×10^{-2}	-0.441	9.1×10^{-4}	-0.477	7.5×10^{-3}
6q26	rs9458419	C/T	0.110	0.39	0.443	1.7×10^{-2}	0.108	0.40	0.435	1.5×10^{-2}
7q21.13	rs10487028	G/T	-0.415	1.7×10^{-3}	-0.527	5.7×10^{-3}	-0.540	3.1×10^{-5}	-0.689	2.2×10^{-4}
8q12.3	rs17179841	C/T	-0.213	1.3×10^{-2}	-1.394	1.8×10^{-3}	-0.237	5.3×10^{-3}	-1.587	6.1×10^{-4}
8p21.3	rs4557718	C/T	-0.028	0.77	0.710	1.2×10^{-2}	-0.037	0.68	0.621	3.0×10^{-2}
11q13.4	rs7116713	C/T	-0.087	0.30	1.850	3.4×10^{-4}	-0.089	0.29	1.351	5.0×10^{-3}
11q23.1	rs522188	A/G	-0.037	0.69	0.795	6.7×10^{-3}	-0.028	0.76	0.751	6.3×10^{-3}
12p12.2	rs17376366	C/T	-0.308	2.2×10^{-3}	-0.452	4.5×10^{-2}	-0.298	2.5×10^{-3}	-0.443	4.2×10^{-2}
16q23.1	rs11150015	A/G	-0.013	0.89	0.791	1.7×10^{-3}	-0.034	0.71	0.753	5.2×10^{-3}
22q13.1	rs2267366	C/T	-0.289	2.9×10^{-3}	-0.826	9.8×10^{-4}	-0.291	2.6×10^{-3}	-0.736	4.0×10^{-3}

Two interaction analyses were performed in the Special Turku Coronary Risk Factor Intervention Project (STRIP) at two age points with a low density lipoprotein (LDL) -cholesterol as the phenotype. The table represent betas and p-values for the main effects of the following model: LDL = SNP + fat quality + SNP × fat quality.

Two of the 21 SNPs were from the same locus (*PARK2)*. The PARK2 locus was also the only locus replicated in the subsequent replication study in YFS subjects (**Table 4)**. The lead SNP from *PARK2* locus was rs12207186. However, the interaction between rs9364628 and dietary fat intake did not replicate in YFS, but was consistent with the discovery study that the dietary fat quality had a more profound effect on circulating LDL-cholesterol concentration in rs9364628-TT carriers (**Figure 9**).

Table 4 Replication phase results from genetic interaction analysis between SNPs in the discovered genomic loci and fat quality ((MUFA+PUFA)/SFA) on serum LDL cholesterol concentration at two age points in the Young Finns Study subjects.

		Age 6						Age 9					
Gene	**Chr**	**Top hit**	**Beta**	**Alleles**	**MAF**	**P**	**FDR**	**Top hit**	**Beta**	**Alleles**	**MAF**	**P**	**FDR**
ADRA2B	2	rs12717775	-0.37	A/C	0.28	7.67×10^{-2}	0.68	rs4062969	2.82	C/T	0.01	5.19×10^{-2}	0.88
ADRA2B	2	rs1724120	-0.42	C/T	0.33	3.49×10^{-2}	0.71	rs11687923	0.33	C/G	0.09	0.23	0.96
PLEKHM3	2	rs2021858	-0.41	T/C	0.30	3.53×10^{-2}	0.87	rs4234084	-496.89	C/T	2.0×10^{-4}	1.57×10-2	0.93
LOC101927123	3	rs17282119	0.61	A/C	0.20	6.07×10^{-3}	0.18	rs2594205	0.54	G/A	0.14	2.43×10-3	0.23
OXSM	3	rs13086718	-0.62	A/C	0.23	$2.86E\times10^{-3}$	0.29	rs7644196	0.43	C/G	0.44	8.63×10-4	0.06
GALNTL2	3	rs690522	-0.46	T/C	0.46	8.87×10^{-3}	0.54	rs6787251	0.51	C/T	0.29	2.51×10-4	0.14
SYNPR	3	rs6445305	-1635.67	T/C	2.0×10^{-4}	4.73×10^{-3}	0.80	rs12631618	0.44	T/C	0.25	2.95×10-3	0.11
TRAM1L1	4	rs13106909	-0.93	G/T	0.11	5.80×10^{-3}	0.54	rs1440313	2.18	C/T	0.02	6.50×10-2	0.96
PRRC1	5	rs6874681	-2.99	A/C	0.02	3.58×10^{-3}	0.17	rs791326	0.38	A/G	0.40	3.63×10-3	0.78
CDH10	5	rs17663246	0.93	G/C	0.12	3.03×10^{-3}	0.30	rs11747402	12.94	C/T	2.8×10^{-4}	2.61×10-2	0.48
PARK2	6	rs12207186	2.23	G/A	0.04	$\mathbf{9.44\times10^{-5}}$	**0.02**	rs1623209	-0.38	C/T	0.22	1.05×10-2	0.43
SLC22A1	6	rs316170	0.54	T/C	0.38	1.08×10^{-2}	0.25	rs1564348	0.41	T/C	0.14	3.46×10-2	1.00
NKAIN3	8	rs12056842	0.93	G/A	0.08	8.65×10^{-3}	0.85	rs16929308	-46.58	C/T	2.0×10^{-4}	3.57×10-2	0.91
LPL	8	rs12680065	-6.35	C/T	0.01	1.21×10^{-2}	0.99	rs1441772	0.40	C/T	0.28	4.80×10-3	0.55
COLCA1	11	rs11605670	0.54	C/A	0.28	1.28×10^{-2}	0.64	rs5020096	-0.38	G/C	0.30	6.81×10-3	0.81
LIPT2	11	rs17132991	2.77	G/A	0.02	1.75×10^{-2}	0.99	rs12805294	-0.59	C/T	0.16	9.35×10-3	0.51
SLCO1B3	12	rs4762759	0.57	A/G	0.25	4.73×10^{-3}	0.45	rs274977	0.62	G/T	0.11	6.70×10-3	0.68
KIAA1576	16	rs9935467	-0.71	T/C	0.33	3.07×10^{-3}	0.97	rs10781989	-0.47	A/C	0.38	3.70×10-4	0.24
BAIAP2L2	22	rs132987	-0.97	G/A	0.07	1.71×10^{-2}	0.75	rs11089864	-1.69	G/C	0.03	5.72×10-3	1.00

MUFA = Monounsaturated fatty acid intake. PUFA = Polyunsaturated fatty acid intake. SFA = Saturated fatty acid intake. LDL = Low-denisty lipoprotein. MAF = Minor allele frequency. FDR = False-discovery rate. From original publication I.

Further analyses in STRIP subjects demonstrated that the interaction is mainly mediated through linoleic acid intake (**Table 5)**. Detailed metabolomic analyses in adult YFS subjects showed that the interaction is primarily on small LDL particles (**Table 6)**. The frequency of T allele of rs9364628 polymorphism was 55.0 % in the STRIP cohort.

Table 5 Individual dietary fatty acids

Dietary lipid	Beta(rs9364628)	Beta(FA)	Beta(interaction)	P(rs9364628)	P(FA)	P(interaction)
Age 5						
Palmitic	0.096	-0.050	-0.056	0.610	0.709	0.627
Stearic	0.143	0.074	-0.084	0.446	0.578	0.470
α-linolenic	0.025	-0.113	-0.003	0.894	0.394	0.982
Linoleic	-0.341	-0.351	0.237	0.060	0.008	**0.040**
SFA	0.183	0.137	-0.109	0.327	0.306	0.348
MUFA	-0.120	-0.196	0.088	0.520	0.140	0.447
PUFA	-0.240	-0.271	0.167	0.195	0.041	0.148
Age 7						
Palmitic	-0.118	-0.026	0.072	0.518	0.848	0.538
Stearic	-0.160	-0.149	0.098	0.375	0.263	0.396
α-linolenic	-0.345	-0.322	0.223	0.058	0.015	0.053
Linoleic	-0.468	-0.458	0.300	0.009	0.001	**0.009**
SFA	-0.061	0.034	0.034	0.738	0.796	0.771
MUFA	-0.417	-0.369	0.270	0.021	0.005	**0.019**
PUFA	-0.436	-0.410	0.280	0.016	0.002	**0.015**

The main effects and interactions between individual dietary lipid intakes and the *PARK2* variant rs9364628 on serum low-density lipoprotein (LDL) -cholesterol concentration in **S**pecial **T**urku Coronary **R**isk Factor **I**ntervention **P**roject (STRIP) cohort. The dietary intakes are measured in grams. The variables were z-transformed (mean=0 and SD=1) prior to analysis. The model for interaction is LDL = fatty acid + SNP + fatty acid × SNP. SFA = Saturated fatty acid intake. MUFA = Monounsaturated fatty acid intake. PUFA = Polyunsaturated fatty acid intake. From original publication I.

Table 6 NMR metabolomics of adult participants in The Cardiovascular Risk in Young Finns Study

Lipid	n	Beta(SNP)	P(SNP)	Beta(Interaction)	P(Interaction)
Small LDL					
Concentration of small LDL particles	1742	-0.230	0.037	0.140	**0.044**
Total lipids in small LDL	1742	-0.220	0.046	0.140	**0.044**
Total cholesterol in small LDL	1742	-0.189	0.087	0.118	0.090
Medium LDL					
Concentration of medium LDL particles	1742	-0.157	0.154	0.094	0.174
Total lipids in medium LDL	1742	-0.150	0.173	0.091	0.192
Total cholesterol in medium LDL	1742	-0.137	0.214	0.083	0.235
Cholesterol esters in medium LDL	1742	-0.135	0.221	0.080	0.251
Phospholipids in medium LDL	1742	-0.151	0.172	0.087	0.210
Large LDL					
Concentration of large LDL particles	1742	-0.102	0.355	0.056	0.416
Total lipids in large LDL	1742	-0.094	0.391	0.052	0.455
Total cholesterol in large LDL	1742	-0.137	0.214	0.083	0.235
Cholesterol esters in large LDL	1742	-0.092	0.403	0.049	0.480
Phospholipids in large LDL	1742	-0.085	0.441	0.048	0.489
Free cholesterol in large LDL	1742	-0.034	0.755	0.016	0.814

Interactions between the *PARK2* variant rs9364628 and dietary fat quality on serum low-density lipoprotein (LDL) subclasses (small, medium, and large) measured using nuclear magnetic resonance spectroscopy in adult participants in The Cardiovascular Risk in Young Finns Study. From original publication I.

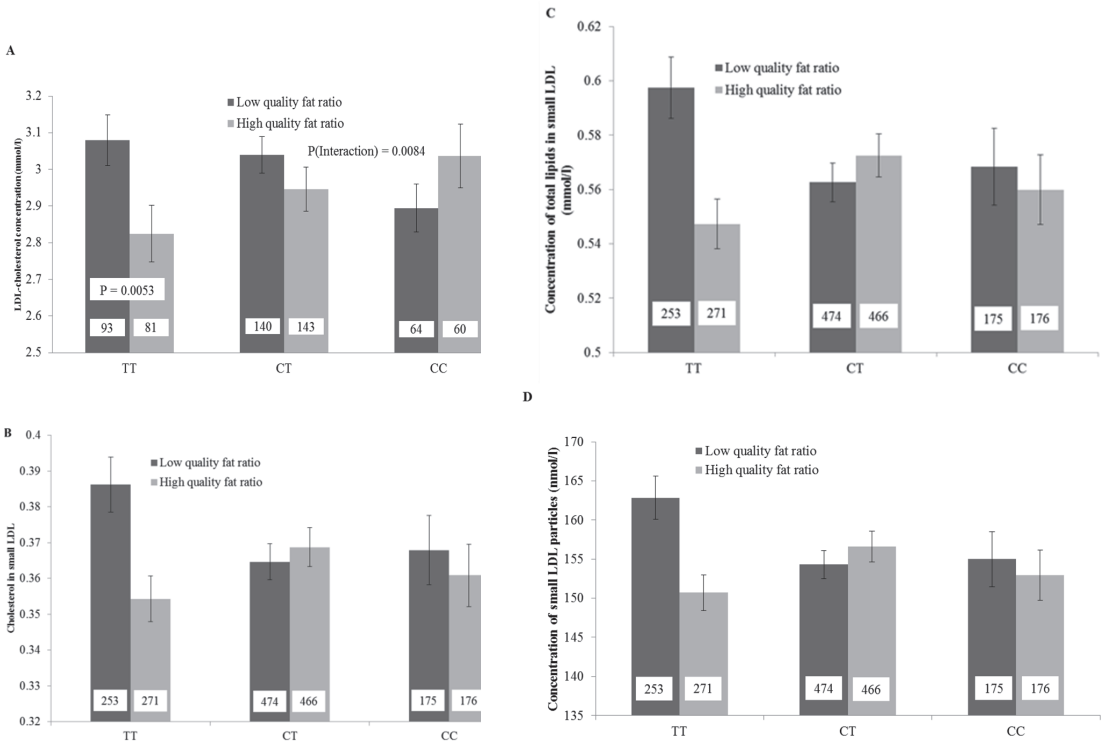


Figure 9 The relationship between dietary fat intake and circulating lipids. The rs9364628 modifies the association between dietary fat intake and circulating LDL-cholesterol. The interaction effect is statistically significant in the STRIP (P = 0.0084). The effect of dietary fat on circulating LDL-cholesterol concentration is greatest among the carriers of rs9364628-TT genotype (**A**). In the YFS cohort, the variant rs9364628 modifies the association between dietary fat intake and circulating cholesterol concentration in small LDL particles, which are the most atherogenic form of LDL. Similarly, the greatest effect is noted in carriers of the rs9364628-TT genotype (**B**). The same is noted for concentration of total lipids in LDL and concentration of small LDL particles (**C** and **D**). Variant rs9364628 is located within the *PARK2* gene. The fat ratio describing dietary intake of fatty acids is calculated as (polyunsaturated fats + monounsaturated fats) ÷ saturated fats. The bars represent standard error and the columns represent mean. The number of subjects in each category is indicated in white boxes within the columns. LDL = Low-density lipoprotein. YFS = The Cardiovascular Risk in Young Finns.

To assess the mechanisms by which the identified variants within the *PARK2* locus could contribute to lipid traits, we queried the polymorphisms from Ensembl's variant effect predictor. This tool predicted that the C allele of rs9364628 leads to non-sense mediated decay of the *PARK2* transcript. The *PARK2* probes in the YFS transcriptomic data did not pass quality control and thus it was not possible to

assess the effect of *PARK2* variants on *PARK2* transcription. However, it was possible to search for trans-eQTLs for *PARK2* variants (rs9364628, rs12207186 and rs9458419) within the genome, however, significant associations between these SNPs and gene expression was not detected.

5.2 STUDY II

The immunostaining procedure demonstrated the existence of MC1R on the luminal side of both human and wild-type mice arteries. The $MC1R^{e/e}$ mice had increased collagen deposition in the aorta and lower diastolic blood pressure compared to wild-type mice. However, no difference was observed in systolic blood pressure or heart rate. These hemodynamic changes led to increased pulse pressure in $MC1R^{e/e}$ mice. In wire-myograph measurements, the contractile responses of thoracic aortae to potassium and phenylephrine was reduced in $MC1R^{e/e}$ compared to the wild-type. In addition, the acetylcholine response was diminished. The eNOS inhibitor L-NNA had a greater effect on acetylcholine induced vasodilatation in wild-type mice indicating that the nitric oxide dependent component of vasodilatation was reduced in $MC1R^{e/e}$ mice. Consistent with these findings, the NOS activity and concentration of NOS metabolites were reduced in $MC1R^{e/e}$ mice (Rinne et al. 2015).

To assess whether the results obtained in a genetically modified mouse strain applied to humans, we determined whether human subjects carrying *MC1R* signaling disrupting mutations have impaired vascular function. The subjects carrying alleles previously associated with weak MC1R function had lower brachial FMD ($P = 0.037$), carotid distensibility ($P = 0.029$) and Young's elastic modulus of the carotid artery ($P = 0.013$). Since MC1R signaling has been previously associated with skin color, and vitamin D status has been suggested to have an impact on vascular function, we assessed whether *MC1R* variants are associated with circulating vitamin D concentration. Subjects with weak MC1R signaling had higher circulating vitamin D concentration (58.0 vs. 54.6 nom/l, $P = 0.014$) but further adjustment of models concerning vascular function parameters with vitamin D concentration did not substantially change the results. As opposed to the results in mice, no association between MC1R signaling and pulse pressure or diastolic blood pressure was observed in humans (**Table 7** and **Figure 10**).

Table 7 Longitudinal analysis of genetic risk score and vessel wall elasticity related variables (n = 2001)

		95 % Confidence limits		
	Estimate	**Lower**	**Upper**	**P**
Carotid artery distensibility	-0.0427	-0.080	-0.005	0.026
Flow-mediated dilation	-0.0106	-0.019	-0.003	0.010
Young's elastic modulus	11.1813	2.603	19.760	0.011
Intima-media thickness	0.0008	-0.005	0.006	0.761
Carotid artery systolic diameter	0.0048	-0.028	0.037	0.771
Carotid artery diastolic diameter	0.0129	-0.017	0.043	0.393

The model was adjusted for age, sex, BMI, and mean arterial pressure. Compound symmetry was used as covariance structure. The results are based on longitudinal model and data from 2001 and 2007 follow-up visits. From original publication II.

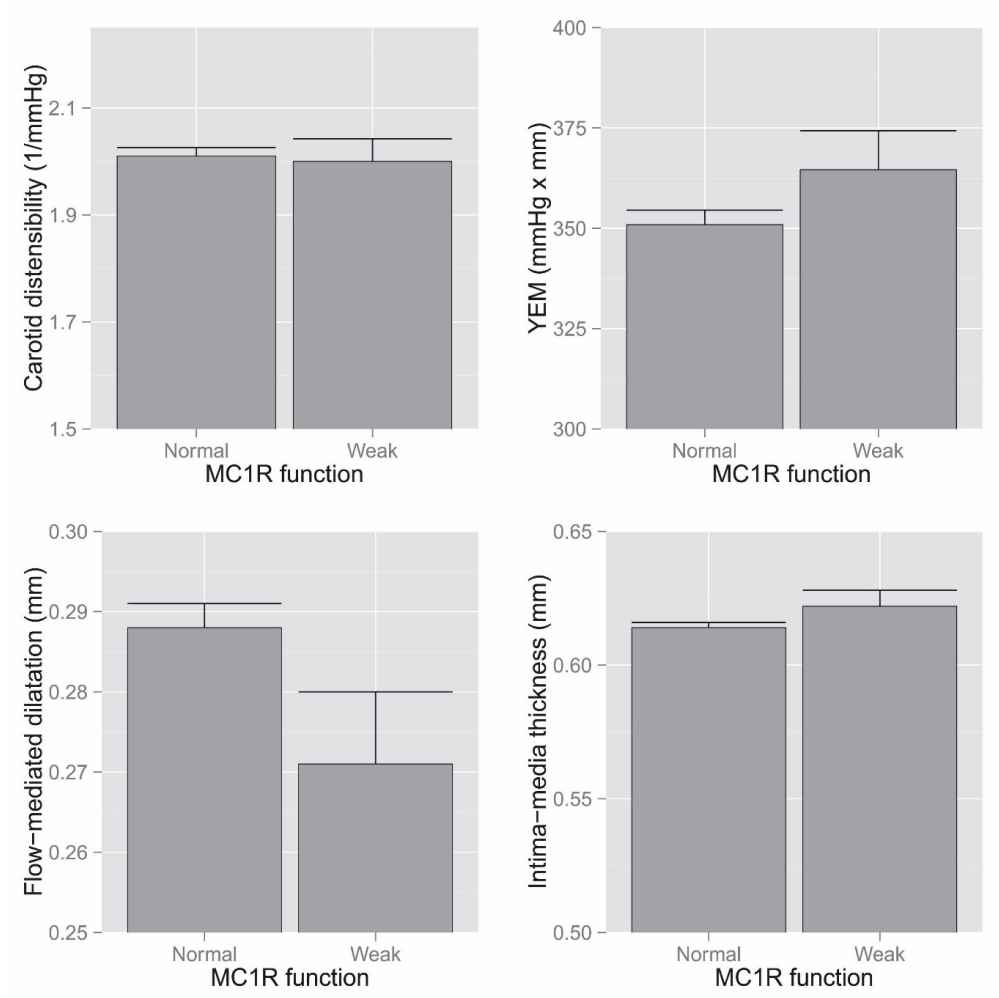


Figure 10 Mean carotid distensibility, Young's elastic modulus (YEM), flow-mediated dilatation and carotid intima-media thickness by genetic MC1R signaling capacity. MC1R = Melanocortin 1 receptor.

5.3 STUDY III

The correlations between the 41 phenotypes included in the study are depicted in **Figure 11**. A module of five cytokines (VEGF, IL7, IL10, IL12p70, and IL13) with a stronger correlation to each other can be identified from Figure 5. In addition, another module consisting of seven different cytokines (MCP1, IL1b, TNFa, IFNg, IL2, IL6, and IL1ra) can be seen in the upper left-hand corner of Figure 5. After quality control, approximately 10.7 million SNPs were included in the meta-

analyses of 41 circulating cytokines. A total of 27 loci reached genome-wide significance ($P < 1.2 \times 10^{-9}$), 17 of which had not previously been linked to cytokine or growth factor concentration, CRP concentration or white blood cell count (**Table 8** and **Figure 12**). The strongest association was detected between rs192989810 and TRAIL concentration ($\beta = 2.2$ SD; $p = 8.4 \times 10^{-175}$; MAF = 1 %). The minor allele frequency of the identified variants ranged between 1 % and 47 %. After identifying the cytokine associated loci, we conditioned the model with the lead variant until any significant associations could no longer be detected. A total of five variants within 17q12 were associated with MIP1b concentration. All MIP1b concentration associated variants (including the five variants (rs113877493, rs76842834, rs142499028, rs184920307, and rs11263779) in 17q12 and two variants (rs138591554 and rs113341849) in 3p21.31) explained altogether 34 % of the total variance (both environmental and genetic) in circulating MIP1b concentration. Two growth factors with a large fraction of variance explained by the identified variants were VEGF (15 %) and TRAIL (14 %). The G allele of rs12075 within *ACKR1* was associated with decreased concentration of eotaxin, MCP1, and GROa. The smallest association test p-values was detected for MCP1 ($\beta = 0.219$; $P = 1.44 \times 10^{-44}$; MAF = 0.47). No association or a considerably weaker association (depending on the cytokine) was detected between rs12075 and eotaxin, MCP1, or GROa in FINRISK2002 where the cytokine quantification was conducted with heparin plasma. Variants in 6p21.1 were significantly associated with VEGF, IL-10, IL-7, and IL-13. The concentrations of these cytokines were also correlated. To assess whether the association between these four cytokines is causal, we performed a Mendelian randomization analysis (**Figure 13**). We constructed a genetic score for VEGF and compared the causal effect estimates with epidemiological estimates which corresponded well with each other. Since 6p21.1 contains the *VEGFA* gene encoding VEGF, the results indicated that IL-10, IL-7, and IL-13 are causally regulated by VEGF.

In the whole blood transcriptomic analyses, we identified 18 eQTLs for 10 of the 27 lead variants at significance level of $p < 6.2 \times 10^{-8}$. As an example, we found interesting biology for the A allele of rs4778636 that was associated with increased concentration of the *IL16* gene transcript but decreased concentration of IL16 cytokines (**Table 9**). Ensembl's variant effect predictor predicted that the A allele of rs4778636 leads to non-sense mediated decay of *IL16* mRNA. Supporting this prediction, no IL16 cytokine was detected in rs4778636 AA homozygotes (**Figure 14**). To match cytokine associations with disease risk variants, we searched the GWAS catalogue in order to identify loci harboring variants that had previously been associated with complex diseases and that were also associated with cytokine

concentration. The cytokine associated SNPs are also associated to Behcet's disease, celiac disease, inflammatory bowel diseases, and multiple sclerosis (**Table 10**) (Ahola-Olli et al. 2017).

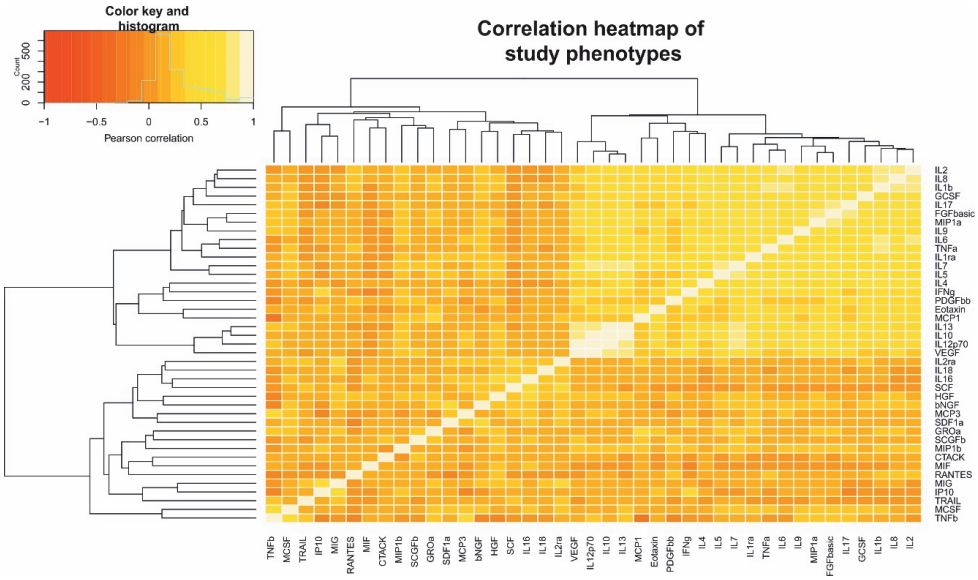


Figure 11 Correlation coefficients and clustering of the studied phenotypes. From original publication III.

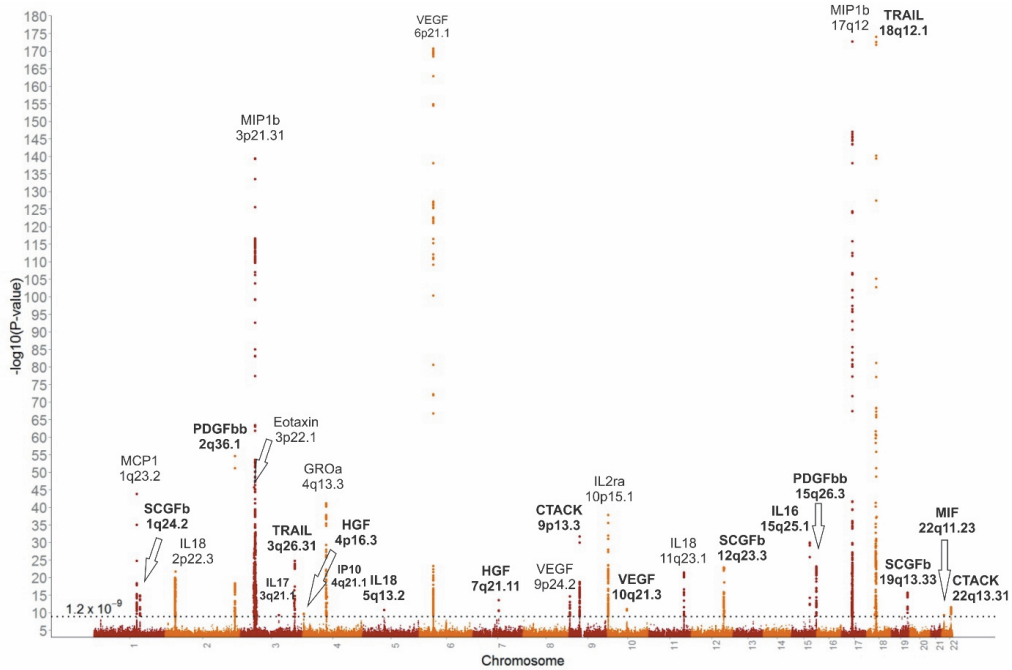


Figure 12 Combined Manhattan plot from the genome-wide association study of 41 circulating cytokines. The novel loci are shown in bold in the Figure. The dashed line indicates the limit for genome-wide significance accounting for 41 cytokine measures. From original publication III.

Table 8 All loci associated with the concentration of at least one of the studied cytokines

Cytokine	Rs-id	Gene	Locus	Position	A1	A2	MAF	Info	Beta	SE	P-value	HetPVal
MCP1	rs12075	*DARC*	1q23.2	159175354	G	A	0.47	1.00	0.219	0.015	1.44×10^{-44}	8.76×10^{-78}
SCGFb	rs4656185	*F5*	1q24.2	169476326	G	A	0.29	1.00	0.205	0.026	1.16×10^{-15}	7.37×10^{-6}
IL18	rs385076	*NLRC4*	2p22.3	32489851	T	C	0.34	0.97	0.243	0.025	1.66×10^{-22}	0.07
PDGFbb	rs13412535	*SERPINE2*	2q36.1	224874874	G	A	0.19	0.87	0.335	0.021	2.46×10^{-55}	1.86×10^{-10}
Eotaxin	rs 2228467	*CCBP2*	3p22.1	42906116	T	C	0.07	1.00	0.416	0.029	2.27×10^{-46}	0.02
MIP1b	rs113010081	*CCRL2*	3p21.31	46457412	T	C	0.13	0.89	0.595	0.024	3.85×10^{-140}	0.33
IL17	rs1530455	*PDIA5*	3q21.1	122854899	T	C	0.36	0.95	-0.108	0.017	4.71×10^{-10}	0.81
TRAIL	esv2656942	*TNFSF10*	3q26.31	172274209	A	-	0.04	0.98	-0.431	0.041	1.74×10^{-25}	0.74
HGF	rs3748034	*HGFAC*	4p16.3	3446091	G	T	0.12	0.97	0.150	0.023	1.79×10^{-10}	0.50
GROa	rs508977	*CXCL1*	4q13.3	74762383	T	G	0.24	1.00	0.380	0.028	7.56×10^{-42}	0.03
IP10	rs141053179	*CXCL10*	4q21.1	77589911	C	G	0.01	0.99	1.103	0.140	2.81×10^{-15}	0.97
IL18	rs17229943	*OCLN*	5q13.2	68682536	A	C	0.08	0.84	0.312	0.046	1.62×10^{-11}	0.98
VEGF	rs6921438	*VEGFA*	6p21.1	43925607	G	A	0.47	0.98	-0.490	0.018	2.09×10^{-171}	2.19×10^{-47}
HGF	rs5745687	*HGF*	7q21.11	81359051	C	T	0.04	1.00	-0.307	0.041	2.71×10^{-14}	0.29
VEGF	rs7030781	*VLDLR*	9p24.2	2686273	A	T	0.42	0.98	-0.137	0.017	2.57×10^{-15}	0.01
CTACK	rs2070074	*IL11RA*	9p13.3	34649442	A	G	0.11	1.00	-0.447	0.037	1.79×10^{-32}	5.55×10^{-5}
IL2ra	rs12722497	*IL2RA*	10p15.1	6095928	C	A	0.06	0.99	0.628	0.049	1.57×10^{-38}	0.18
VEGF	rs10761731	*JMJD1C*	10q21.3	65027610	A	T	0.36	0.99	0.119	0.017	1.01×10^{-11}	0.54
IL18	rs71478720	*IL18*	11q23.1	112009605	C	T	0.24	0.99	-0.267	0.028	3.71×10^{-22}	0.59
SCGFb	rs187503377	*STAB2*	12q23.3	104261835	C	T	0.02	0.93	0.965	0.097	1.34×10^{-23}	0.07
IL16	rs4778636	*IL16*	15q25.1	81591639	G	A	0.04	1.00	-0.727	0.063	1.11×10^{-30}	0.22
PDGFbb	rs4965869	*PCSK6*	15q26.3	101990320	C	T	0.26	1.00	0.184	0.018	5.66×10^{-24}	7.2×10^{-3}
MIP1b	rs113877493	*CCL4L1*	17q12	34812273	C	T	0.17	0.92	-0.612	0.022	1.62×10^{-173}	3.58×10^{-7}
TRAIL	rs192989810	*MEP1B*	18q12.1	29783353	C	T	0.01	0.96	2.221	0.078	8.40×10^{-175}	5.92×10^{-3}
SCGFb	rs116924815	*CLEC11A*	19q13.33	51230733	C	T	0.03	0.97	0.608	0.074	1.74×10^{-16}	0.03
MIF	rs2330634	*MIF*	22q11.23	24250795	C	G	0.37	0.99	-0.156	0.025	4.53×10^{-10}	0.11
CTACK	rs201003839	*PPARA*	22q13.31	46534944	A	ACGGGC	0.27	0.99	0.190	0.027	2.42×10^{-12}	1.36×10^{-3}

The same locus is occasionally associated with more than one cytokine but only the strongest trait-association is reported in the table. Allele 2 is the effect allele. Beta and SE are in SD-units. All positions refer to the human genome build 37. The Info column reports the imputation quality metric calculated. HetPVal reports the effect size heterogeneity statistics between the cohorts. From original publication III.

Table 9 Results from eQTL analyses for locus specific lead SNPs

SNP locus	SNP	Probe	Probe locus	Gene	Other allele	Effect allele	Beta	SE	P-value	HetPval
1q23.2	rs12075	ILMN_1723684	1q23.2	*DARC*	G	A	0.111	0.017	6.95×10^{-11}	1.00
2q36.1	rs13412535	ILMN_1655595	2q36.1	*SERPINE2*	G	A	-0.309	0.018	9.65×10^{-63}	1.00
3p21.31	rs113010081	ILMN_1763322	3p21.31	*CCR3*	T	C	0.372	0.031	4.80×10^{-33}	0.87
4q13.3	rs508977	ILMN_1745522	4q13.3	*PF4V1*	T	G	0.257	0.031	7.54×10^{-17}	0.66
5q13.2	rs17229943	ILMN_1679620	5q13.2	*LOC728519*	A	C	0.342	0.036	6.72×10^{-22}	0.35
5q13.2	rs17229943	ILMN_1760189	5q13.2	*NAIP*	A	C	0.256	0.031	3.05×10^{-16}	0.36
5q13.2	rs17229943	ILMN_2260082	5q13.2	*NAIP*	A	C	0.228	0.032	8.09×10^{-13}	0.79
5q13.2	rs17229943	ILMN_1767377	5q13.2	*LOC153561*	A	C	0.226	0.037	6.88×10^{-10}	0.25
9p13.3	rs2070074	ILMN_1664912	9p13.3	*IL11RA**	A	G	0.192	0.018	2.07×10^{-26}	0.36
9p13.3	rs2070074	ILMN_1720024	9p13.3	*IL11RA**	A	G	0.081	0.013	8.30×10^{-10}	1.00
9p13.3	rs2070074	ILMN_1657475	9p13.3	*GALT*	A	G	-0.155	0.016	2.70×10^{-21}	0.90
15q25.1	rs4778636	ILMN_2290628	15q25.1	*IL16**	G	A	0.636	0.023	2.70×10^{-161}	1.00
17q12	rs113877493	ILMN_2100209	17q12	*CCL4L1*	C	T	-0.389	0.020	2.14×10^{-87}	1.00
17q12	rs113877493	ILMN_2105573	17q12	*CCL3L3*	C	T	-0.391	0.039	4.51×10^{-23}	0.47
17q12	rs113877493	ILMN_2218856	17q12	*CCL3L1*	C	T	-0.180	0.021	1.72×10^{-18}	0.70
17q12	rs113877493	ILMN_1716276	17q12	*CCL4L2*	C	T	-0.164	0.021	1.27×10^{-14}	0.15
19q13.33	rs116924815	ILMN_1807359	19q13.33	*CLEC11A**	C	T	0.450	0.023	1.99×10^{-85}	1.00
22q11.23	rs2330634	ILMN_1690982	22q11.23	*DDT*	C	G	0.196	0.006	4.75×10^{-211}	0.77

Probe IDs are from Illumina HumanHT-12 v.4 Expression BeadChip. * = The SNP is associated with cytokine concentration and with mRNA coding the cytokine. CTACK is also known as IL11RA and stem cell growth factor beta (SCGFb) as CLEC11A. From original publication III.

Table 10 SNPs previously associated with disease and in this study found to be associated with cytokine concentrations.

Locus	Rs-id	Disease	Cytokine	Other allele	Effect allele / Disease risk increasing allele	EAF	Beta	StdErr	P-value
3p21.31	rs7616215	Behcet's disease	MIP1b	T	C	0.40	-0.142	0.016	4.8×10^{-19}
3p21.31	rs13096142	Celiac disease	MIP1b	C	T	0.29	-0.163	0.017	4.8×10^{-21}
6p21.1	rs943072	Ulcerative colitis	IL12p70	T	G	0.11	0.155	0.025	3.3×10^{-10}
6p21.1	rs943072	Ulcerative colitis	VEGF	T	G	0.11	0.176	0.027	2.0×10^{-11}
10p15.1	rs12722489	Crohn's disease + MS	IL2ra	T	C	0.87	0.325	0.034	4.1×10^{-21}
10p15.1	rs12722515	Inflammatory bowel disease	IL2ra	A	C	0.87	0.328	0.034	1.6×10^{-21}

Each SNP reported in the table have been previously associated with the corresponding disease or trait. P-values denote the SNP-cytokine association. EAF = Effect allele frequency. MS=Multiple sclerosis. From original publication III.

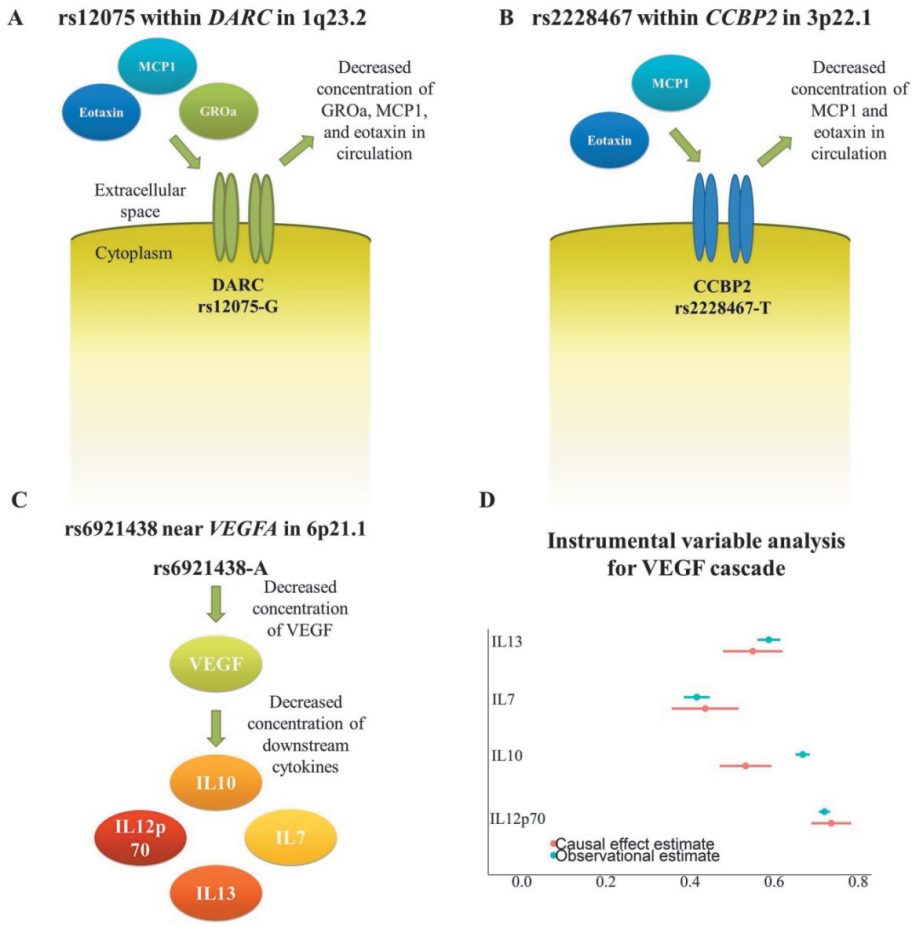


Figure 13 Variant rs12075 is associated at genome-wide significance with concentration of three cytokines. The effect of rs12075 on these cytokines is mediated via a DARC receptor (**Panel A**). The variant rs2228467 is located within the chemokine-binding protein 2 (*CCBP2*) and is associated at genome-wide significance with a concentration of eotaxin and the monocyte chemotactic protein 1 (MCP1) (**Panel B**). Variant rs6921438 near *VEGFA* is associated with concentrations of IL10, IL12p70, IL7 and IL13. The Mendelian randomization analysis suggests that the effect is mediated via circulating VEGF (**Panels C-D**). From original publication III.

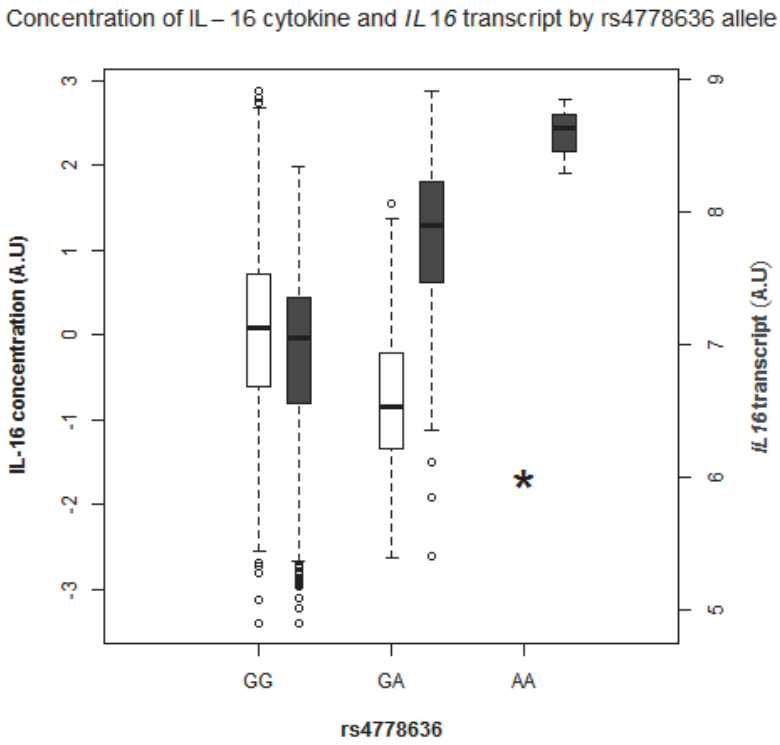


Figure 14 The inverse normal transformed IL-16 cytokine concentration (white boxes) and *IL16* expression (dark boxes) by the rs4778636 allele. The rs4778636-A in 15q25.1 is associated with decreased IL16 concentration. The rs4778636-AA homozygotes had an undetectable IL-16 concentration although the *IL16* expression is high. According to Ensembl, rs4778636 creates a splice cite modification that leads to non-sense mediated decay of *IL16* mRNA. The vertical bars within the boxes represent the median, and the top and the bottom of the boxes indicate 75th and 25th percentile. The whiskers indicate the 75th or 25th percentile ± 1.5 × interquartile range. From original publication III.

6 DISCUSSION

6.1 STUDY I

Study I demonstrated that the *PARK2* gene might be involved in lipid metabolism in healthy children. The TT allele of rs9364628 predisposes to serum LDL-cholesterol elevation on a diet which is rich in saturated fat and poor in unsaturated fat. The effect of rs9364628 is mainly seen on small LDL particles which are the most atherogenic form of LDL. However, despite the consistent pattern between dietary fat quality and circulating LDL-cholesterol in subjects with rs9364628-TT genotype, the interaction effect seen in STRIP could not be replicated with the same SNP in YFS cohort. Since the replication of the exact SNP could not be performed, further research is needed before the interaction between rs9364628 and dietary fat quality can be confirmed. However, consistent patterns in different lipid measures in different cohorts lend support for the possibility that the results in Study I have true biological meaning instead of presenting only statistical interaction (Thomas 2010).

PARK2 encodes a protein called Parkin which is a ubiquitin ligase tagging specific proteins for degradation into proteasomes and lysosomes (Shimura et al. 2000; Zhang et al. 2000). The loss of ligase activity in Parkin leads to dopaminergic neuron loss in autosomal juvenile Parkinsonism and therefore *PARK2* polymorphisms have been extensively studied in the development of central nervous system disorders (Kitada et al. 1998; Shaltouki et al. 2015). However, *PARK2* also has an emerging role in the regulation of energy metabolism and homeostasis (K.-Y. Kim et al. 2011; K.-Y. Kim and Sack 2012; Locke et al. 2015). Since *PARK2* deletion in mice has been shown to protect against metabolic changes induced by a high fat diet (such as hepatic steatosis), our results suggest that extraneuronal metabolic side-effects from interventions aimed to increase expression of Parkin can be assumed in humans. Based on current knowledge, it is impossible to evaluate whether the metabolic and the neuronal effects of Parkin are mediated via the same pathways: Kim et al. noted that the Parkin increases the stability of CD36 (a fatty acid transporter) and promotes fatty acid uptake whereas increased intake of PUFA and MUFA has been linked to decreased risk of Parkinson's disease (K.-Y. Kim et al. 2011; de Lau et al. 2005; D. D. Wang et al. 2016). However, since Parkin has several ubiquitination targets it is also possible that the metabolic and neuronal effects are mediated via two distinct pathways. Thus, further research is needed to unravel this important area (Gao et al. 2015).

Nutrigenomics and nutrigenetics refer to the scientific disciplines studying inter-individual differences in responses to dietary exposures caused by genetic variants, gene expression or epigenetic changes. The American Society for Nutrition have identified these as one of the five concepts that are critical to nutrition research in the 21st century (Ohlhorst et al. 2013). Replication of the results shown in Study I is vital prior to any genotype based individualized nutritional counseling. In addition, the effect of rs9364628 on cardiovascular risk needs to be addressed with respect to the dietary fat intake before giving any recommendations to the general population. The results of Study I should be considered rather as a discovery of new lipid biology and drug targets in humans as opposed to a pursuit of tools for individualized nutritional counseling.

An important limitation of study I is the small sample size. However, it is strengthened by inclusion of detailed metabolomic profiling and assessment of individual dietary fatty acids. Despite the fact that rs9364628 did not have a statistically significant interaction in the replication sample, the larger effect of dietary fat quality on circulating LDL-cholesterol in subjects with rs9364628-TT is constantly replicated across study populations, different age groups, and different lipid markers. Further research is needed to assess the effects of *PARK2* variants on cardio-metabolic outcomes such as fatty liver in humans.

6.2 STUDY II

Study II demonstrated that MC1R signaling contributes to the regulation of vascular tone both in mice and humans. Since MC1R signaling has previously been linked to pigmentation, a difference in the vitamin D status was not unexpected (Han et al. 2008; Liu et al. 2015). However, the results from Study II further demonstrated that the effects of MC1R signaling are not mediated through vitamin D.

In their previous study, Rinne et al. demonstrated that pharmacological activation of the melanocortin system with melanotan II has positive effects on vascular health in mice (Rinne et al. 2014). However, since melanotan II is not specific for any of the melanocortin receptors, the receptor type through which the effects are mediated remained unknown. The current study demonstrated that *MC1R* is expressed on endothelial cells in mice and humans. Furthermore, MC1R co-localizes with its natural ligand α-MSH in vasculature indicating that the melanocortin signaling might occur in an autocrine or paracrine manner. In addition, the study

demonstrates that deficiency in MC1R signaling leads to increased arterial stiffness and impaired endothelial function in mice and humans. These vascular changes are surrogate markers of atherosclerosis and indicate higher cardiovascular risk (Ras et al. 2013; van Sloten et al. 2014). Thus, drugs targeting MC1R might have potential in the treatment of atherosclerosis. An important limitation of Study II is that we were not able to link MC1R signaling to clinical cardiovascular end-points, such as myocardial infarction due to the small number of events in YFS, which is a result of the current age distribution of the cohort. However, results from recently published GWAS studying coronary artery disease could be exploited to assess whether the *MC1R* variants included in our genetic risk score contribute to the risk of coronary artery disease or myocardial infarction (Nikpay et al. 2015).

Recently, an MC1R-selective derivative was developed from melanotan II (Doedens et al. 2010). Therefore, future studies should assess whether this derivative contributes to vascular health in mice. Since melanotan II has already been tested on humans, further studies could evaluate the effect of melanotan II on vascular parameters, such as FMD (Fitzgerald et al. 2006).

6.3 STUDY III

Study III identified 17 new loci contributing to the concentrations of circulating cytokines and growth factors. Understanding the molecular mechanisms behind immune system regulation will facilitate setting-up therapies for diseases with an inflammatory component, such as coronary artery disease. In addition to identifying new loci associated with inflammatory traits, the study suggested that a circulating vascular endothelial growth factor (VEGF) regulates the concentrations of four other cytokines: IL7, IL10, IL12p70, and IL13. Therefore, drugs inhibiting VEGF will also affect the concentration of the four cytokines. This might lead to unwanted side-effects of VEGF-targeting drugs. VEGFA signals through cell surface receptor tyrosine kinases (VEGFR-1, VEGFR-2, and VEGFR-3). For example, angiogenesis is mainly mediated through VEGF-2 (Shibuya 2015). Therefore, in some instances targeting a specific VEGF-receptor instead of causing alterations to circulating concentration of VEGF (and therefore to concentrations of IL7, IL10, IL12p70, and IL13) might be an alternative to gain a suitable treatment effect that avoids side effects (**Figure 15)** (Holmes et al. 2007).

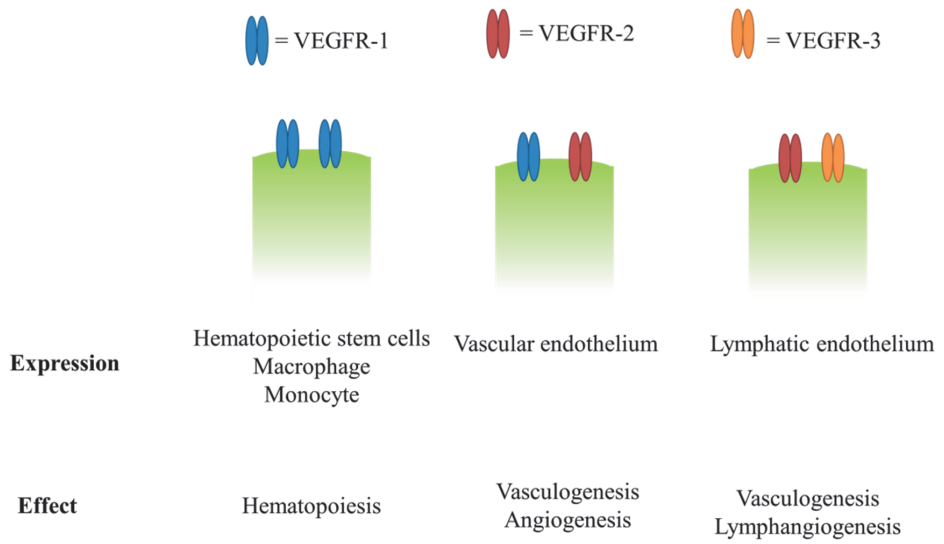


Figure 15 Signaling through VEGF-receptors. VEGF-signaling is mediated through three different receptors, the expression of which differs between tissues. Targeting specific receptors instead of circulating VEGF might help to reduce side-effects.

Since the effect of 6p21.1 variants (rs6921438 and rs12214617) on IL-7, IL-10, IL-12p70 are mediated via VEGF, the core assumptions of Mendelian randomization are not violated and thus 6p21.1 variants can be used as instruments in Mendelian randomization. In contrast, the effect of rs12075-G on circulating eotaxin, MCP1 and GROa is mediated through another protein (ACKR1) which violates the assumptions of Mendelian randomization and thus rs12075 cannot be used as an instrument for GROa, MCP1, or eotaxin (Davey Smith and Hemani 2014). By explaining the large proportion of variance (34 %) in the circulating MIP1b concentration, we have provided a strong instrument for future Mendelian randomization studies which will now be able to assess the causality of circulating MIP1b in the pathogenesis of inflammatory diseases.

Another interesting finding from Study III was the association between rs4778636 and circulating IL16 concentration. Homozygous rs4778636-AA study participants had undetectable amounts of circulating IL16 but increased concentration of *IL16* mRNA. This is due to splice cite modification caused by rs4778636-A. IL16 has been suggested to contribute to the pathogenesis of HIV (Baier et al. 1995). The frequency of rs4778636-A in the YRI population (Yoruba, Ibadan, Nigeria) is 25 % although in European populations it is considerably smaller. The small allele frequency in the European population combined with the small sample size or possible recessive effects might explain why the association between *IL16* locus and

HIV have not been detected in GWASs (Johnson et al. 2015).Therefore, our results will assist studies of HIV progression in humans lacking *IL16* signaling.

Finally, we were able to link variants previously associated with inflammatory disease to cytokine signaling. Variants associated with decreased concentration of interleukin-2 receptor α subunit (IL2ra) were associated with increased risk of multiple sclerosis and inflammatory bowel diseases (Jostins et al. 2012; Patsopoulos et al. 2011). Circulating IL2ra is cleaved from a membrane bound IL2-receptor (H. P. Kim, Imbert, and Leonard 2006). IL2ra targeting the antibody daclizumab has been shown to be effective in treatment of multiple sclerosis (Gold et al. 2013). These results suggest that daclizumab might be effective also in the treatment of inflammatory bowel diseases. In addition, variants previously linked to Behcet's disease (a rare small-artery vasculitis) and celiac disease were also associated with concentration of MIP1b (Garner et al. 2014; Kirino et al. 2013). However, rs13096142-T was associated with increased concentration of MIP1b and increased risk for celiac diseases whereas rs7616215-T was associated with increased risk for Behcet's disease but with a decreased concentration of MIP1b. Therefore, MIP1b-signaling inhibiting agents designed to treat Behcet's disease might potentially increase the risk for celiac disease.

Surprisingly, no association was detected between cardiovascular diseases and cytokine-associated loci although cardiovascular diseases have an important inflammatory component in their pathogenesis (Golia et al. 2014; Jialal and Devaraj 2015). However, a coronary artery disease GWAS based on a 1000 Genomes imputation was published after we downloaded the GWAS catalogue database (Nikpay et al. 2015). Therefore, by regularly re-downloading the database and conducting fresh analyses the loci might continuously reveal new biology behind inflammatory diseases since new studies are constantly being added to the database. The chromosomal positions are in the Human Genome Build 37 whereas the current version of the GWAS catalogue uses Human Genome Build 38 and thus the genome positions must be changed to matching reference genome. This can be done, for example, with the LiftOver tool available via USCS Genome Browser(Rosenbloom et al. 2015). In addition, the constant growth of The Genotype-Tissue Expression (GTEx) portal enables anyone to access tissue specific transcriptomic data (Carithers et al. 2015). These data contain eQTL information from tissues relevant to cardiovascular disease. The database has expression data, for example, from arteries and adipose tissue. Therefore, GTEx database could be used to search whether cytokine-associated loci harbor variants that are associated with cytokine gene expression in arteries since circulating cytokines might not be

as relevant for cardiovascular disease development than locally produced cytokines (Helseth et al. 2015). The cytokine gene expression within the arterial wall could be used as a proxy for local cytokine production.

In conclusion, Study III identified 17 new loci associated with the concentration of circulating cytokines. These results will help to clarify the causal role of cytokine signaling in the development of inflammatory diseases and set the molecular basis for future studies assessing the molecular regulation of the immune system.

6.4 FUTURE PROSPECTS

Genetic studies with continuously increasing sample sizes continue to identify loci which predispose to complex diseases. A recent study assessing human height reached a sample size of 700,000 subjects and increased the variance explained by all known variants to 27.4 % from the heritability of height (heritability estimate 0.80) (Marouli et al. 2017; Visscher et al. 2006). In addition to increasing sample sizes, it is time to concentrate on detailed molecular phenotyping which might be a successful approach with only few thousands of subjects, as demonstrated by Study III. Detailed molecular phenotypes could be measured, for example, by using proteomics or untargeted metabolomics. To derive causal effect estimates based on the results from these molecular studies it is of importance that complete summary statistics from genetic studies of complex diseases continue to be placed on publically accessible servers (Stephen Burgess et al. 2015). In addition, the continuously increasing amount of biobank data makes it possible to exploit reverse genetics (i.e. the correlating variants associated with molecular phenotype with range of diagnosis retrieved from hospital registries) to evaluate the side effects of drugs targeting the molecule associated with the pertinent variants.

The most important finding of this thesis is that the variants increasing IL2ra concentration also increase the risk for Crohn's disease and multiple sclerosis. Since IL2ra antibody daclizumab reduced disease relapses in multiple sclerosis considerably, it is of importance to evaluate the efficacy of IL2ra antibodies in treatment of Crohn's disease as well. Since daclizumab has gained marketing approval for other indications earlier it could be possible to directly launch phase II trials without the need for i.e. toxicity studies which are necessary before doing clinical testing with a completely novel pharmaceutical compound.

7 CONCLUSIONS

In this thesis, the potential of three different study designs to identify new drug targets were evaluated. Study I supports the role of *PARK2* in regulation of lipid metabolism. Since the mutations, which cause the loss of ubiquitin ligase activity of Parkin, are associated with Parkinson's disease, targeting Parkin with drugs might cause unwanted side-effects. Study II demonstrated that the beneficial vascular effects of melanotan-II are mediated via the MC1R receptor in mice. This receptor also contributes to vascular health in humans: Human subjects with variants previously linked to weak MC1R signaling have impaired vascular function. These functional changes have been linked to increased cardiovascular risk. Therefore, drugs targeting MC1R signaling might have a role in treatment and prevention of cardiovascular diseases. Study III linked novel loci to circulating concentration of cytokines and growth factors which point to novel therapeutic targets which might have role in treatment and prevention of inflammatory disorders. The results also set the framework for further studies to evaluate the role of the immune system in health and disease. Overall, this thesis demonstrates how various human genetic approaches can be used to point novel drug targets with potential to treat complex diseases which are a major cause of mortality and morbidity world-wide. The most robust results were obtained via detailed molecular phenotypes. However, the results also indicate how human genetics and animal studies can be used to complement each other.

8 ACKNOWLEDGEMENTS

This study was implemented at the Research Centre of Applied and Preventive cardiovascular Medicine (CAPC) in collaboration with the Department of Internal Medicine, the Department of Clinical Physiology and Nuclear Medicine and the Department of Pharmacology, Drug Development and Pharmaceutics, University of Turku, Turku, Finland, during 2010-2015. This thesis work was financially supported by the Juho Vainio Foundation, Julin Foundation, the Finnish Medical Foundation, and The Lions Club.

I owe my deepest gratitude to my three excellent supervisors who have guided and taught me during my doctoral studies. I am deeply indebted to my supervisor, Professor Olli Raitakari. His excellent knowledge in the fields of atherosclerosis, statistics and medical writing has been invaluable for this study. Despite his busy schedules, he has been able to allocate time to share his vast expertise. Without his intelligent ideas, this thesis would not have been possible. I consider it an honor to work with my supervisor, Docent Johannes Kettunen. His proficiency in the fields of genetics and statistics has been invaluable for this thesis. Thank you for your patience, kind support and encouraging attitude that you have devoted to my work. I continue to be inspired by all our discussions. It is with immense gratitude that I acknowledge my supervisor, PhD Niina Pitkänen for her profound scientific guidance and constructive comments during writing this thesis.

I express my sincere gratitude to emeritus professors Jorma Viikari and Tapani Rönnemaa for their thoughtful comments, constructive criticism and productive ideas while preparing this thesis. Their tremendous knowledge in the field of medicine has greatly inspired me during these years.

I express my sincere thanks to all other co-authors: Vera Mikkilä, Terho Lehtimäki, Mika Kähönen, Katja Pahkala, Harri Niinikoski, Antti J. Kangas, Pasi Soininen, Olli Simell, Petteri Rinne, Salla Nuutinen, Emilia Koskinen, Katja Kaipio, Kim Eerola, Markus Juonala, Aki Havulinna, Peter Würtz, Kristiina Aalto, Emma Raitoharju, Ilkka Seppälä, Antti-Pekka Sarin, Samuli Ripatti, Aarne Palotie, Markus Perola, Sirpa Jalkanen, Mikael Maksimow, Veikko Salomaa, Marko Salmi, and Eriika Savontaus.

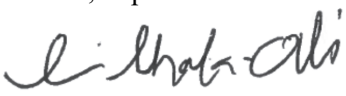
I wish to express my gratitude to the official reviewers Docent William Hennah and Dr Jussi Naukkarinen whose comments have greatly helped to improve this thesis.

I also express my gratitude to volunteer subjects who has participated to the study. Without them this study would have not been possible.

I warmly thank Ville Aalto, Irina Lisinen, Noora Kartiosuo, and Jarno Kankaanranta for their statistical help during preparation of this thesis. Their excellent data management skills have been also a great help. Several discussions with all of them have helped me to improve this thesis. My sincere thanks to Nina Ruotsalainen, Nina Aalto, Tiina Peromaa, Minna Hyppönen, and Marja Piippo for helping me any time I needed assistance. I thank the talented investigators of CAPC, Petri Kresanov, Atte Voipio, Sonja Firtser, Lauri Vähämurto, Joel Nuotio, Olli Hartiala, Jonna Juhola, Juha Koskinen, Juho Raiko, Suvi Rovio, Saku Ruohonen, Noora Mattson, Emmi Suomela, Ina Anttila, Mari Nupponen, Hanna Mikola, Jarkko Heiskanen, Emmi Suominen, Behnoush Foroutan, Petri Niinisalo, Martha Snäll, Tomi Laitinen, Harri Helajärvi, Jari Kaikkonen, Olli Heinonen, and Johanna Jaakkola for countless happy moments during these years. Especially, I want to thank Juha Mykkänen from inspiring conversations concerning human genetics. I am profoundly indebted to Leo-Pekka Lyytikäinen from the University of Tampere for the help he has kindly provided.

I am thankful to all my family members and friends from their support throughout my life and their interest toward my scientific work. Finally, I want to thank my love Jaana for her never-ending support and understanding during the thesis preparation. I have always been able to count on her knowledge whenever I have asked. Especially, I want to thank her for every great moment we have experienced together during these years.

Turku, September 2017



Ari Ahola-Olli

REFERENCES

Abifadel, Marianne et al. 2003. "Mutations in PCSK9 Cause Autosomal Dominant Hypercholesterolemia." *Nature genetics* 34(2): 154–56.

Ahola-Olli, Ari V et al. 2014. "Interactions between Genetic Variants and Dietary Lipid Composition: Effects on Circulating LDL Cholesterol in Children." *The American journal of clinical nutrition* 100(6): 1569–77.

Ahola-Olli, Ari V. et al. 2017. "Genome-Wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors." *The American Journal of Human Genetics* 100(1): 40–50.

Anttonen, Anna-Kaisa et al. 2017. "ZNHIT3 Is Defective in PEHO Syndrome, a Severe Encephalopathy with Cerebellar Granule Neuron Loss." *Brain* 140(5): 1267–79.

Ariotti, Sara et al. 2016. "Is Bare-Metal Stent Implantation Still Justifiable in High Bleeding Risk Patients Undergoing Percutaneous Coronary Intervention?" *JACC: Cardiovascular Interventions* 9(5): 426–36.

Arrowsmith, John, and Philip Miller. 2013. "Trial Watch: Phase II and Phase III Attrition Rates 2011-2012." *Nature reviews. Drug discovery* 12(8): 569.

Avolio, A, D Jones, and M Tafazzoli-Shadpour. 1998. "Quantification of Alterations in Structure and Function of Elastin in the Arterial Media." *Hypertension* 32(1): 170–75.

Baier, M et al. 1995. "HIV Suppression by Interleukin-16." *Nature* 378(6557): 563.

Barrett, Jennifer H et al. 2011. "Genome-Wide Association Study Identifies Three New Melanoma Susceptibility Loci." *Nature genetics* 43(11): 1108–13.

Barter, Philip J et al. 2007. "Effects of Torcetrapib in Patients at High Risk for Coronary Events." *The New England journal of medicine* 357(21): 2109–22.

Bernstein, Bradley E et al. 2012. "An Integrated Encyclopedia of DNA Elements in the Human Genome." *Nature* 489(7414): 57–74.

Bochud, Murielle. 2012. "Genetics for Clinicians: From Candidate Genes to Whole Genome Scans (Technological Advances)." *Best practice & research. Clinical endocrinology & metabolism* 26(2): 119–32.

Bodmer, Walter, and Carolina Bonilla. 2008. "Common and Rare Variants in Multifactorial Susceptibility to Common Diseases." *Nature genetics* 40(6): 695–701.

Boring, L, J Gosling, M Cleary, and I F Charo. 1998. "Decreased Lesion Formation in CCR2-/- Mice Reveals a Role for Chemokines in the Initiation of Atherosclerosis." *Nature* 394(6696): 894–97.

Bumgarner, Roger. 2001. Chapter 22 Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.] *DNA Microarrays: Types, Applications and Their Future Roger*. eds. Frederick M. Ausubel et al. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Burgess, S. 2014. "Sample Size and Power Calculations in Mendelian Randomization with a Single Instrumental Variable and a Binary Outcome." *International Journal of Epidemiology* 43(3): 922–29.

Burgess, Stephen et al. 2015. "Using Published Data in Mendelian Randomization: A Blueprint for Efficient Identification of Causal Risk Factors." *European Journal of Epidemiology* 30(7): 543–52.

Cao, Richard Yang et al. 2017. "Effects of p53-Knockout in Vascular Smooth Muscle Cells on Atherosclerosis in Mice" ed. Andreas Zirlik. *PLOS ONE* 12(3): e0175061.

Carithers, Latarsha J. et al. 2015. "A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project." *Biopreservation and Biobanking* 13(5): 311–19.

Chambers, John C et al. 2008. "Common Genetic Variation near MC4R Is Associated with Waist Circumference and Insulin Resistance." *Nature genetics* 40(6): 716–18.

Chheda, Himanshu et al. 2017. “Whole-Genome View of the Consequences of a Population Bottleneck Using 2926 Genome Sequences from Finland and United Kingdom.” *European journal of human genetics : EJHG* 25(4): 477–84.

Cho, Yoon Shin et al. 2012. “Meta-Analysis of Genome-Wide Association Studies Identifies Eight New Loci for Type 2 Diabetes in East Asians.” *Nature genetics* 44(1): 67–72.

Cole, Christopher B., Majid Nikpay, and Ruth McPherson. 2015. “Gene–environment Interaction in Dyslipidemia.” *Current Opinion in Lipidology* 26(2): 133–38.

Collins, F S, and V A McKusick. 2001. “Implications of the Human Genome Project for Medical Science.” *JAMA* 285(5): 540–44.

Crossman, David C et al. 2008. “Investigation of the Effect of Interleukin-1 Receptor Antagonist (IL-1ra) on Markers of Inflammation in Non-ST Elevation Acute Coronary Syndromes (The MRC-ILA-HEART Study).” *Trials* 9(1): 8.

Davey Smith, George, and Gibran Hemani. 2014. “Mendelian Randomization: Genetic Anchors for Causal Inference in Epidemiological Studies.” *Human molecular genetics* 23(R1): R89-98.

Dawson, Elisabeth et al. 2002. “A First-Generation Linkage Disequilibrium Map of Human Chromosome 22.” *Nature* 418(6897): 544–48.

Dehghan, Abbas et al. 2011. “Meta-Analysis of Genome-Wide Association Studies in >80 000 Subjects Identifies Multiple Loci for C-Reactive Protein Levels.” *Circulation* 123(7): 731–38.

Delaneau, Olivier, Jonathan Marchini, and Jean-François Zagury. 2012. “A Linear Complexity Phasing Method for Thousands of Genomes.” *Nature methods* 9(2): 179–81.

Deloukas, Panos et al. 2013. “Large-Scale Association Analysis Identifies New Risk Loci for Coronary Artery Disease.” *Nature genetics* 45(1): 25–33.

DiMasi, Joseph A, Ronald W Hansen, and Henry G Grabowski. 2003. “The Price of Innovation: New Estimates of Drug Development Costs.” *Journal of health economics* 22(2): 151–85.

Doedens, Lucas et al. 2010. “Multiple N-Methylation of MT-II Backbone Amide Bonds Leads to Melanocortin Receptor Subtype hMC1R Selectivity: Pharmacological and Conformational Studies.” *Journal of the American Chemical Society* 132(23): 8115–28.

Doshi, S N et al. 2001. “Flow-Mediated Dilatation Following Wrist and Upper Arm Occlusion in Humans: The Contribution of Nitric Oxide.” *Clinical science (London, England : 1979)* 101(6): 629–35.

Ehret, Georg B et al. 2011. “Genetic Variants in Novel Pathways Influence Blood Pressure and Cardiovascular Disease Risk.” *Nature* 478(7367): 103–9.

Elkind, M.S.V. et al. 2009. “High-Sensitivity C-Reactive Protein Predicts Mortality but Not Stroke: The Northern Manhattan Study.” *Neurology* 73(16): 1300–1307.

Evans, Mark, Ciaran Sewter, and Elizabeth Hill. 2003. “An Encoded Particle Array Tool for Multiplex Bioassays.” *Assay and drug development technologies* 1(1 Pt 2): 199–207.

Ferguson, J A, F J Steemers, and D R Walt. 2000. “High-Density Fiber-Optic DNA Random Microsphere Array.” *Analytical chemistry* 72(22): 5618–24.

Fitzgerald, Liesel M, Jayne L Fryer, Terence Dwyer, and Stuart M Humphrey. 2006. “Effect of MELANOTAN, [Nle(4), D-Phe(7)]-Alpha-MSH, on Melanin Synthesis in Humans with MC1R Variant Alleles.” *Peptides* 27(2): 388–94.

Fleg, Jerome L et al. 2008. “Effect of Statins Alone versus Statins plus Ezetimibe on Carotid Atherosclerosis in Type 2 Diabetes: The SANDS (Stop Atherosclerosis in Native Diabetics Study) Trial.” *Journal of the American College of Cardiology* 52(25): 2198–2205.

Friedewald, W T, R I Levy, and D S Fredrickson. 1972. “Estimation of the Concentration of Low-Density Lipoprotein Cholesterol in Plasma, without Use of the Preparative Ultracentrifuge.” *Clinical chemistry* 18(6): 499–502.

Furchgott, R F, and J V Zawadzki. 1980. “The Obligatory Role of Endothelial Cells in the Relaxation of Arterial Smooth Muscle by Acetylcholine.” *Nature* 288(5789): 373–76.

Gabriel, S. B. 2002. “The Structure of Haplotype Blocks in the Human Genome.” *Science* 296(5576): 2225–29.

Gao, Feng et al. 2015. “The Mitochondrial Protein BNIP3L Is the Substrate of PARK2 and Mediates Mitophagy in PINK1/PARK2 Pathway.” *Human molecular genetics* 24(9): 2528–38.

García-Closas, Montserrat et al. 2005. “NAT2 Slow Acetylation, GSTM1 Null Genotype, and Risk of Bladder Cancer: Results from the Spanish Bladder Cancer Study and Meta-Analyses.” *Lancet (London, England)* 366(9486): 649–59.

Garner, Chad et al. 2014. “Genome-Wide Association Study of Celiac Disease in North America Confirms FRMD4B as New Celiac Locus.” *PloS one* 9(7): e101428.

Gauldie, J et al. 1987. “Interferon Beta 2/B-Cell Stimulatory Factor Type 2 Shares Identity with Monocyte-Derived Hepatocyte-Stimulating Factor and Regulates the Major Acute Phase Protein Response in Liver Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 84(20): 7251–55.

Genovese, Giulio et al. 2010. “Association of Trypanolytic ApoL1 Variants with Kidney Disease in African Americans.” *Science (New York, N.Y.)* 329(5993): 841–45.

Giusti, Imarilde I. et al. 2013. “High Doses of Vascular Endothelial Growth Factor 165 Safely, but Transiently, Improve Myocardial Perfusion in No-Option Ischemic Disease.” *Human Gene Therapy Methods* 24(5): 298–306.

Gold, Ralf et al. 2013. “Daclizumab High-Yield Process in Relapsing-Remitting Multiple Sclerosis (SELECT): A Randomised, Double-Blind, Placebo-Controlled Trial.” *Lancet* 381(9884): 2167–75.

Golia, Enrica et al. 2014. “Inflammation and Cardiovascular Disease: From Pathogenesis to Therapeutic Target.” *Current Atherosclerosis Reports* 16(9): 435.

Gordon, Derek, and Stephen J Finch. 2005. “Factors Affecting Statistical Power in the Detection of Genetic Association.” *The Journal of clinical investigation* 115(6): 1408–18.

Gordon, T et al. 1977. “High Density Lipoprotein as a Protective Factor against Coronary Heart Disease. The Framingham Study.” *The American journal of medicine* 62(5): 707–14.

Grieco, P et al. 2003. “Extensive Structure-Activity Studies of Lactam Derivatives of MT-II and SHU-9119: Their Activity and Selectivity at Human Melanocortin Receptors 3, 4, and 5.” *The journal of peptide research : official journal of the American Peptide Society* 62(5): 199–206.

Groenwold, Rolf H.H. et al. 2016. “Sensitivity Analysis for the Effects of Multiple Unmeasured Confounders.” *Annals of Epidemiology* 26(9): 605–11.

de Groot, Eric et al. 2004. “Measurement of Arterial Wall Thickness as a Surrogate Marker for Atherosclerosis.” *Circulation* 109(23 Suppl 1): III33-8.

Guldberg, P et al. 1995. “Phenylketonuria in a Low Incidence Population: Molecular Characterisation of Mutations in Finland.” *Journal of medical genetics* 32(12): 976–78.

Gunderson, Kevin L et al. 2004. “Decoding Randomly Ordered DNA Arrays.” *Genome research* 14(5): 870–77.

Hakala, Paula et al. 1996. “Calculated vs Analysed Nutrient Composition of Weight Reduction Diets.” *Food Chemistry* 57(1): 71–75.

Han, Jiali et al. 2008. “A Genome-Wide Association Study Identifies Novel Alleles Associated with Hair Color and Skin Pigmentation.” *PLoS genetics* 4(5): e1000074.

Helseth, Ragnhild et al. 2015. “Associations between Circulating Proteins and Corresponding Genes Expressed in Coronary Thrombi in Patients with Acute Myocardial Infarction.” *Thrombosis Research* 136(6): 1240–44.

Hennekens, Charles H. et al. 1996. “Lack of Effect of Long-Term Supplementation with Beta Carotene on the Incidence of Malignant Neoplasms and Cardiovascular Disease.” *New England Journal of Medicine* 334(18): 1145–49.

Hime, Neil J, Audrey S Black, Josh J Bulgrien, and Linda K Curtiss. 2008. “Leukocyte-Derived Hepatic Lipase Increases HDL and Decreases En Face Aortic Atherosclerosis in LDLr-/- Mice Expressing CETP.” *Journal of lipid research* 49(10): 2113–23.

Holmes, Katherine, Owain Ll Roberts, Angharad M. Thomas, and Michael J. Cross. 2007. “Vascular Endothelial Growth Factor Receptor-2: Structure, Function, Intracellular Signalling and Therapeutic Inhibition.” *Cellular Signalling* 19(10): 2003–12.

Holzmann, S. 1982. “Endothelium-Induced Relaxation by Acetylcholine Associated with Larger Rises in Cyclic GMP in Coronary Arterial Strips.” *Journal of cyclic nucleotide research* 8(6): 409–19.

Hooper, Lee, Nicole Martin, Asmaa Abdelhamid, and George Davey Smith. 2015. “Reduction in Saturated Fat Intake for Cardiovascular Disease.” *The Cochrane database of systematic reviews* 6: CD011737.

Howie, Bryan, Jonathan Marchini, and Matthew Stephens. 2011. “Genotype Imputation with Thousands of Genomes.” *G3 (Bethesda, Md.)* 1(6): 457–70.

Howie, Bryan N, Peter Donnelly, and Jonathan Marchini. 2009. “A Flexible and Accurate Genotype Imputation Method for the next Generation of Genome-Wide Association Studies.” *PLoS genetics* 5(6): e1000529.

Hruby, V J et al. 1995. “Cyclic Lactam Alpha-Melanotropin Analogues of Ac-Nle4-cyclo[Asp5, D-Phe7,Lys10] Alpha-Melanocyte-Stimulating Hormone-(4-10)-NH2 with Bulky Aromatic Amino Acids at Position 7 Show High Antagonist Potency and Selectivity at Specific Melanocortin Receptors.” *Journal of medicinal chemistry* 38(18): 3454–61.

Hunt, S C et al. 2000. “Genetic Localization to Chromosome 1p32 of the Third Locus for Familial Hypercholesterolemia in a Utah Kindred.” *Arteriosclerosis, thrombosis, and vascular biology* 20(4): 1089–93.

Ielasi, Alfonso et al. 2015. “Everolimus-Eluting Stent versus Bare-Metal Stent in Elderly (≥75years) versus Non-Elderly (<75years) Patients with ST-Elevation Myocardial Infarction Undergoing Primary Percutaneous Coronary Intervention: Insights from the Examination Trial.” *International Journal of Cardiology* 179: 73–78.

Ignarro, L J, R E Byrns, G M Buga, and K S Wood. 1987. “Endothelium-Derived Relaxing Factor from Pulmonary Artery and Vein Possesses Pharmacologic and Chemical Properties Identical to Those of Nitric Oxide Radical.” *Circulation research* 61(6): 866–79.

Ikonomidis, I. et al. 2008. “Inhibition of Interleukin-1 by Anakinra Improves Vascular and Left Ventricular Function in Patients With Rheumatoid Arthritis.” *Circulation* 117(20): 2662–69.

Inouye, Michael et al. 2010. “An Immune Response Network Associated with Blood Lipid Levels.” *PLoS genetics* 6(9): e1001113.

Interleukin 1 Genetics Consortium. 2015. “Cardiometabolic Effects of Genetic Upregulation of the Interleukin 1 Receptor Antagonist: A Mendelian Randomisation Analysis.” *The Lancet Diabetes & Endocrinology* 3(4): 243–53.

International HapMap Consortium. 2003. “The International HapMap Project.” *Nature* 426(6968): 789–96.

———. 2005. “A Haplotype Map of the Human Genome.” *Nature* 437(7063): 1299–1320.

Jialal, I, and S Devaraj. 2015. “Anti-Inflammatory Strategies to Prevent Diabetic Cardiovascular Disease.” *Clinical Pharmacology & Therapeutics* 98(2): 121–23.

Johnson, Eric O et al. 2015. “Novel Genetic Locus Implicated for HIV-1 Acquisition with Putative Regulatory Links to HIV Replication and Infectivity: A Genome-Wide Association Study.” *PloS one* 10(3): e0118149.

Jostins, Luke et al. 2012. “Host-Microbe Interactions Have Shaped the Genetic Architecture of Inflammatory Bowel Disease.” *Nature* 491(7422): 119–24.

Juncker, David, Sébastien Bergeron, Veronique Laforte, and Huiyan Li. 2014. “Cross-Reactivity in Antibody Microarrays and Multiplexed Sandwich Assays: Shedding Light on the Dark Side of Multiplexing.” *Current opinion in chemical biology* 18: 29–37.

Juonala, M et al. 2004. "The 21-Year Follow-up of the Cardiovascular Risk in Young Finns Study: Risk Factor Levels, Secular Trends and East-West Difference." *Journal of internal medicine* 255(4): 457–68.

Kannel, W B et al. 1961. "Factors of Risk in the Development of Coronary Heart Disease--Six Year Follow-up Experience. The Framingham Study." *Annals of internal medicine* 55: 33–50.

Keene, Daniel, Clare Price, Matthew J Shun-Shin, and Darrel P Francis. 2014. "Effect on Cardiovascular Risk of High Density Lipoprotein Targeted Drug Treatments Niacin, Fibrates, and CETP Inhibitors: Meta-Analysis of Randomised Controlled Trials Including 117,411 Patients." *BMJ (Clinical research ed.)* 349: g4379.

Kettunen, Johannes et al. 2016. "Genome-Wide Study for Circulating Metabolites Identifies 62 Loci and Reveals Novel Systemic Effects of LPA." *Nature communications* 7: 11122.

Kim, Hyoung Pyo, Jean Imbert, and Warren J Leonard. 2006. "Both Integrated and Differential Regulation of Components of the IL-2/IL-2 Receptor System." *Cytokine & growth factor reviews* 17(5): 349–66.

Kim, Kye-Young et al. 2011. "Parkin Is a Lipid-Responsive Regulator of Fat Uptake in Mice and Mutant Human Cells." *The Journal of clinical investigation* 121(9): 3701–12.

Kim, Kye-Young, and Michael N Sack. 2012. "Parkin in the Regulation of Fat Uptake and Mitochondrial Biology: Emerging Links in the Pathophysiology of Parkinson's Disease." *Current opinion in lipidology* 23(3): 201–5.

Kirii, Hirokazu et al. 2003. "Lack of Interleukin-1beta Decreases the Severity of Atherosclerosis in ApoE-Deficient Mice." *Arteriosclerosis, thrombosis, and vascular biology* 23(4): 656–60.

Kirino, Yohei et al. 2013. "Genome-Wide Association Analysis Identifies New Susceptibility Loci for Behçet's Disease and Epistasis between HLA-B*51 and ERAP1." *Nature genetics* 45(2): 202–7.

Kirnbauer, R et al. 1989. "IFN-Beta 2, B Cell Differentiation Factor 2, or Hybridoma Growth Factor (IL-6) Is Expressed and Released by Human Epidermal Cells and Epidermoid Carcinoma Cell Lines." *Journal of immunology* 142(6): 1922–28.

Kitada, T et al. 1998. "Mutations in the Parkin Gene Cause Autosomal Recessive Juvenile Parkinsonism." *Nature* 392(6676): 605–8.

Kittles, R A et al. 1998. "Dual Origins of Finns Revealed by Y Chromosome Haplotype Variation." *American journal of human genetics* 62(5): 1171–79.

Kooner, Jaspal S et al. 2011. "Genome-Wide Association Study in Individuals of South Asian Ancestry Identifies Six New Type 2 Diabetes Susceptibility Loci." *Nature genetics* 43(10): 984–89.

Koskinen, Juha et al. 2012. "Effect of Age, Gender and Cardiovascular Risk Factors on Carotid Distensibility during 6-Year Follow-Up. The Cardiovascular Risk in Young Finns Study." *Atherosclerosis* 224(2): 474–79.

Kukuła, Krzysztof et al. 2011. "Intramyocardial Plasmid-Encoding Human Vascular Endothelial Growth Factor A165/basic Fibroblast Growth Factor Therapy Using Percutaneous Transcatheter Approach in Patients with Refractory Coronary Artery Disease (VIF-CAD)." *American Heart Journal* 161(3): 581–89.

de la Chapelle, A. 1993. "Disease Gene Mapping in Isolated Human Populations: The Example of Finland." *Journal of medical genetics* 30(10): 857–65.

Lander, E S et al. 2001. "Initial Sequencing and Analysis of the Human Genome." *Nature* 409(6822): 860–921.

de Lau, L M L et al. 2005. "Dietary Fatty Acids and the Risk of Parkinson Disease: The Rotterdam Study." *Neurology* 64(12): 2040–45.

Lehtimäki, T et al. 1995. "Association between Serum Lipids and Apolipoprotein E Phenotype Is Influenced by Diet in a Population-Based Sample of Free-Living Children and Young Adults: The Cardiovascular Risk in Young Finns Study." *Journal of lipid research* 36(4): 653–61.

Levine, Glenn N. et al. 2016. "2016 ACC/AHA Guideline Focused Update on Duration of Dual Antiplatelet Therapy in Patients With Coronary Artery Disease." *Journal of the American College of Cardiology* 68(10): 1082–1115.

Levy, D et al. 2000. "Evidence for a Gene Influencing Blood Pressure on Chromosome 17. Genome Scan Linkage Results for Longitudinal Blood Pressure Phenotypes in Subjects from the Framingham Heart Study." *Hypertension* 36(4): 477–83.

Light, Donald W, and Rebecca N Warburton. 2005. "Extraordinary Claims Require Extraordinary Evidence." *Journal of health economics* 24(5): 1030-3-53.

Lim, Elaine T et al. 2014. "Distribution and Medical Impact of Loss-of-Function Variants in the Finnish Founder Population." ed. David Cutler. *PLoS genetics* 10(7): e1004494.

Liu, Fan et al. 2015. "Genetics of Skin Color Variation in Europeans: Genome-Wide Association Studies with Functional Follow-Up." *Human genetics* 134(8): 823–35.

Locke, Adam E. et al. 2015. "Genetic Studies of Body Mass Index Yield New Insights for Obesity Biology." *Nature* 518(7538): 197–206.

Looser, Patrick M., Luke K. Kim, and Dmitriy N. Feldman. 2016. "In-Stent Restenosis: Pathophysiology and Treatment." *Current Treatment Options in Cardiovascular Medicine* 18(2): 10.

Lorenz, Matthias W et al. 2007. "Prediction of Clinical Cardiovascular Events with Carotid Intima-Media Thickness: A Systematic Review and Meta-Analysis." *Circulation* 115(4): 459–67.

———. 2012. "Carotid Intima-Media Thickness Progression to Predict Cardiovascular Events in the General Population (the PROG-IMT Collaborative Project): A Meta-Analysis of Individual Participant Data." *Lancet* 379(9831): 2053–62.

Magnani, J. W. et al. 2014. "Sequencing of SCN5A Identifies Rare and Common Variants Associated With Cardiac Conduction: Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium." *Circulation: Cardiovascular Genetics* 7(3): 365–73.

Makrygiannis, Stamatis S. et al. 2013. "Prognostic Usefulness of Serial C-Reactive Protein Measurements in ST-Elevation Acute Myocardial Infarction." *The American Journal of Cardiology* 111(1): 26–30.

Marchini, Jonathan, and Bryan Howie. 2010. "Genotype Imputation for Genome-Wide Association Studies." *Nature reviews. Genetics* 11(7): 499–511.

Marouli, Eirini et al. 2017. "Rare and Low-Frequency Coding Variants Alter Human Adult Height." *Nature* 542(7640): 186–90.

McGill, H C et al. 1997. "Effects of Serum Lipoproteins and Smoking on Atherosclerosis in Young Men and Women. The PDAY Research Group. Pathobiological Determinants of Atherosclerosis in Youth." *Arteriosclerosis, thrombosis, and vascular biology* 17(1): 95–106.

Miller, J K, and W M Barnes. 1986. "Colony Probing as an Alternative to Standard Sequencing as a Means of Direct Analysis of Chromosomal DNA to Determine the Spectrum of Single-Base Changes in Regions of Known Sequence." *Proceedings of the National Academy of Sciences of the United States of America* 83(4): 1026–30.

Di Minno, Matteo Nicola Dario et al. 2014. "Lipid Profile Changes in Patients with Rheumatic Diseases Receiving a Treatment with TNF-α Blockers: A Meta-Analysis of Prospective Studies." *Annals of Medicine* 46(2): 73–83.

Mitchell, Aaron C., Priscilla S. Briquez, Jeffrey A. Hubbell, and Jennifer R. Cochran. 2016. "Engineering Growth Factors for Regenerative Medicine Applications." *Acta Biomaterialia* 30: 1–12.

Morehouse, Lee A et al. 2007. "Inhibition of CETP Activity by Torcetrapib Reduces Susceptibility to Diet-Induced Atherosclerosis in New Zealand White Rabbits." *Journal of lipid research* 48(6): 1263–72.

Morgan, Steve et al. 2011. "The Cost of Drug Development: A Systematic Review." *Health policy (Amsterdam, Netherlands)* 100(1): 4–17.

Mountjoy, K G, L S Robbins, M T Mortrud, and R D Cone. 1992. "The Cloning of a Family of Genes That Encode the Melanocortin Receptors." *Science (New York, N.Y.)* 257(5074): 1248–51.

Muona, K et al. 2012. "10-Year Safety Follow-up in Patients with Local VEGF Gene Transfer to Ischemic Lower Limb." *Gene Therapy* 19(4): 392–95.

Musunuru, Kiran et al. 2010. "From Noncoding Variant to Phenotype via SORT1 at the 1p13 Cholesterol Locus." *Nature* 466(7307): 714–19.

Müller, Franz Hermann. 1940. "Tabakmißbrauch Und Lungencarcinom." *Zeitschrift für Krebsforschung* 49(1): 57–85.

Nakashima, Y. et al. 2007. "Early Human Atherosclerosis: Accumulation of Lipid and Proteoglycans in Intimal Thickenings Followed by Macrophage Infiltration." *Arteriosclerosis, Thrombosis, and Vascular Biology* 27(5): 1159–65.

Namboodiri, K K et al. 1985. "The Collaborative Lipid Research Clinics Family Study: Biological and Cultural Determinants of Familial Resemblance for Plasma Lipids and Lipoproteins." *Genetic epidemiology* 2(3): 227–54.

Nelson, Matthew R et al. 2015. "The Support of Human Genetic Evidence for Approved Drug Indications." *Nature Genetics* 47(8): 856–60.

Nielsen, Ulrik B, and Bernhard H Geierstanger. 2004. "Multiplexed Sandwich Assays in Microarray Format." *Journal of immunological methods* 290(1–2): 107–20.

Niinikoski, Harri et al. 2012. "Effect of Repeated Dietary Counseling on Serum Lipoproteins from Infancy to Adulthood." *Pediatrics* 129(3): e704-13.

Nikpay, Majid et al. 2015. "A Comprehensive 1,000 Genomes-Based Genome-Wide Association Meta-Analysis of Coronary Artery Disease." *Nature genetics* 47(10): 1121–30.

Niu, Wenquan et al. 2012. "Association of Interleukin-6 Circulating Levels with Coronary Artery Disease: A Meta-Analysis Implementing Mendelian Randomization Approach." *International Journal of Cardiology* 157(2): 243–52.

Oh, Kyungwon et al. 2005. "Dietary Fat Intake and Risk of Coronary Heart Disease in Women: 20 Years of Follow-up of the Nurses' Health Study." *American journal of epidemiology* 161(7): 672–79.

Ohlhorst, Sarah D et al. 2013. "Nutrition Research to Affect Food and a Healthy Life Span." *The American journal of clinical nutrition* 98(2): 620–25.

Oleksyk, Taras K et al. 2010. "Worldwide Distribution of the MYH9 Kidney Disease Susceptibility Alleles and Haplotypes: Evidence of Historical Selection in Africa." *PloS one* 5(7): e11474.

Osadnik, Tadeusz et al. 2016. "The Relationships between Polymorphisms in Genes Encoding the Growth Factors TGF-β1, PDGFB, EGF, bFGF and VEGF-A and the Restenosis Process in Patients with Stable Coronary Artery Disease Treated with Bare Metal Stent" ed. Graham R. Wallace. *PLOS ONE* 11(3): e0150500.

Patsopoulos, Nikolaos A et al. 2011. "Genome-Wide Meta-Analysis Identifies Novel Multiple Sclerosis Susceptibility Loci." *Annals of neurology* 70(6): 897–912.

Perheentupa, J. 1972. "[Hereditary Diseases in Finland--from the Clinician's and Scientist's Point of View]." *Duodecim; laaketieteellinen aikakauskirja* 88(1): 1–3.

Pignoli, P et al. 1986. "Intimal plus Medial Thickness of the Arterial Wall: A Direct Measurement with Ultrasound Imaging." *Circulation* 74(6): 1399–1406.

Polvi, Anne et al. 2013. "The Finnish Disease Heritage Database (FinDis) Update-A Database for the Genes Mutated in the Finnish Disease Heritage Brought to the Next-Generation Sequencing Era." *Human Mutation* 34(11): 1458–66.

Porkka, K V et al. 1997. "Trends in Serum Lipid Levels during 1980-1992 in Children and Young Adults. The Cardiovascular Risk in Young Finns Study." *American journal of epidemiology* 146(1): 64–77.

Preston, Samuel H., and Andrew Stokes. 2014. "Obesity Paradox." *Epidemiology* 25(3): 454–61.

Prins, Bram. P. et al. 2016. “Investigating the Causal Relationship of C-Reactive Protein with 32 Complex Somatic and Psychiatric Outcomes: A Large-Scale Cross-Consortium Mendelian Randomization Study” ed. Phillipa J. Hay. *PLOS Medicine* 13(6): e1001976.

Qi, Qibin et al. 2012. “Sugar-Sweetened Beverages and Genetic Risk of Obesity.” *New England Journal of Medicine* 367(15): 1387–96.

———. 2014. “Fried Food Consumption, Genetic Risk, and Body Mass Index: Gene-Diet Interaction Analysis in Three US Cohort Studies.” *BMJ (Clinical research ed.)* 348: g1610.

Raitakari, Olli T et al. 2003. “Cardiovascular Risk Factors in Childhood and Carotid Artery Intima-Media Thickness in Adulthood: The Cardiovascular Risk in Young Finns Study.” *JAMA : the journal of the American Medical Association* 290(17): 2277–83.

Rajamäki, Kristiina et al. 2010. “Cholesterol Crystals Activate the NLRP3 Inflammasome in Human Macrophages: A Novel Link between Cholesterol Metabolism and Inflammation.” ed. Derya Unutmaz. *PloS one* 5(7): e11765.

Ramesh, V et al. 1988. “Molecular Basis of Ornithine Aminotransferase Deficiency in B-6-Responsive and -Nonresponsive Forms of Gyrate Atrophy.” *Proceedings of the National Academy of Sciences of the United States of America* 85(11): 3777–80.

Ras, Rouyanne T, Martinette T Streppel, Richard Draijer, and Peter L Zock. 2013. “Flow-Mediated Dilation and Cardiovascular Risk Prediction: A Systematic Review with Meta-Analysis.” *International journal of cardiology* 168(1): 344–51.

Ridker, Paul M. 2009. “Testing the Inflammatory Hypothesis of Atherothrombosis: Scientific Rationale for the Cardiovascular Inflammation Reduction Trial (CIRT).” *Journal of Thrombosis and Haemostasis* 7: 332–39.

Rinne, Petteri, Wendy Nordlund, et al. 2013. “α-Melanocyte-Stimulating Hormone Regulates Vascular NO Availability and Protects against Endothelial Dysfunction.” *Cardiovascular research* 97(2): 360–68.

Rinne, Petteri, Anna-Maija Penttinen, et al. 2013. “α-MSH Analogue Attenuates Blood Pressure Elevation in DOCA-Salt Hypertensive Mice.” *PloS one* 8(8): e72857.

Rinne, Petteri et al. 2014. “Pharmacological Activation of the Melanocortin System Limits Plaque Inflammation and Ameliorates Vascular Dysfunction in Atherosclerotic Mice.” *Arteriosclerosis, Thrombosis, and Vascular Biology* 34: 1346–54.

———. 2015. “Deficiency in Melanocortin 1 Receptor Signaling Predisposes to Vascular Endothelial Dysfunction and Increased Arterial Stiffness in Mice and Humans.” *Arteriosclerosis, Thrombosis, and Vascular Biology* 35(7): 1678–86.

———. 2017. “Melanocortin 1 Receptor Signaling Regulates Cholesterol Transport in Macrophages.” *Circulation*.

Ritchie, Scott C et al. 2015. “The Biomarker GlycA Is Associated with Chronic Inflammation and Predicts Long-Term Risk of Severe Infection.” *Cell systems* 1(4): 293–301.

Roffi, Marco et al. 2016. “2015 ESC Guidelines for the Management of Acute Coronary Syndromes in Patients Presenting without Persistent ST-Segment Elevation.” *European Heart Journal* 37(3): 267–315.

Rosenbloom, Kate R et al. 2015. “The UCSC Genome Browser Database: 2015 Update.” *Nucleic acids research* 43(Database issue): D670-81.

Rubanyi, G M, R R Lorenz, and P M Vanhoutte. 1985. “Bioassay of Endothelium-Derived Relaxing Factor(s): Inactivation by Catecholamines.” *The American journal of physiology* 249(1 Pt 2): H95-101.

Räsänen, L, M Ahola, R Kara, and M Uhari. 1985. “Atherosclerosis Precursors in Finnish Children and Adolescents. VIII. Food Consumption and Nutrient Intakes.” *Acta paediatrica Scandinavica. Supplement* 318: 135–53.

Sabatine, Marc S. et al. 2017. “Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease.” *New England Journal of Medicine*: NEJMoa1615664.

Sattar, Naveed et al. 2010. “Statins and Risk of Incident Diabetes: A Collaborative Meta-Analysis of Randomised Statin Trials.” *The Lancet* 375(9716): 735–42.

Schramm, Tina Ken et al. 2008. “Diabetes Patients Requiring Glucose-Lowering Therapy and Nondiabetics with a Prior Myocardial Infarction Carry the Same Cardiovascular Risk: A Population Study of 3.3 Million People.” *Circulation* 117(15): 1945–54.

Scott, L. J. et al. 2007. “A Genome-Wide Association Study of Type 2 Diabetes in Finns Detects Multiple Susceptibility Variants.” *Science* 316(5829): 1341–45.

Selvin, Elizabeth et al. 2014. “Diabetes Mellitus, Prediabetes, and Incidence of Subclinical Myocardial Damage.” *Circulation* 130(16): 1374–82.

Shaltouki, Atossa et al. 2015. “Mitochondrial Alterations by PARKIN in Dopaminergic Neurons Using PARK2 Patient-Specific and PARK2 Knockout Isogenic iPSC Lines.” *Stem cell reports* 4(5): 847–59.

Shibuya, Masabumi. 2015. “VEGF-VEGFR System as a Target for Suppressing Inflammation and Other Diseases.” *Endocrine, metabolic & immune disorders drug targets* 15(2): 135–44.

Shimura, H et al. 2000. “Familial Parkinson Disease Gene Product, Parkin, Is a Ubiquitin-Protein Ligase.” *Nature genetics* 25(3): 302–5.

Sironi, M et al. 1989. “IL-1 Stimulates IL-6 Production in Endothelial Cells.” *Journal of immunology* 142(2): 549–53.

van Sloten, Thomas T et al. 2014. “Local Stiffness of the Carotid and Femoral Artery Is Associated with Incident Cardiovascular Events and All-Cause Mortality: The Hoorn Study.” *Journal of the American College of Cardiology* 63(17): 1739–47.

Stary, H C et al. 1994. “A Definition of Initial, Fatty Streak, and Intermediate Lesions of Atherosclerosis. A Report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association.” *Circulation* 89(5): 2462–78.

———. 1995. “A Definition of Advanced Types of Atherosclerotic Lesions and a Histological Classification of Atherosclerosis. A Report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association.” *Circulation* 92(5): 1355–74.

Stein, Evan A et al. 2012. “Effect of a Monoclonal Antibody to PCSK9 on LDL Cholesterol.” *The New England journal of medicine* 366(12): 1108–18.

Sung, Si-Chan et al. 2009. “7-Ketocholesterol Upregulates Interleukin-6 via Mechanisms That Are Distinct from Those of Tumor Necrosis Factor-α, in Vascular Smooth Muscle Cells.” *Journal of Vascular Research* 46(1): 36–44.

Swerdlow, Daniel I et al. 2015. “HMG-Coenzyme A Reductase Inhibition, Type 2 Diabetes, and Bodyweight: Evidence from Genetic Analysis and Randomised Trials.” *Lancet (London, England)* 385(9965): 351–61.

Teo, Yik Ying. 2012. “Genotype Calling for the Illumina Platform.” In *Methods in Molecular Biology*, , 525–38.

Teslovich, Tanya M et al. 2010. “Biological, Clinical and Population Relevance of 95 Loci for Blood Lipids.” *Nature* 466(7307): 707–13.

The 1000 Genomes Project Consortium. 2010. “A Map of Human Genome Variation from Population-Scale Sequencing.” *Nature* 467(7319): 1061–73.

Thomas, Duncan. 2010. “Gene--Environment-Wide Association Studies: Emerging Approaches.” *Nature reviews. Genetics* 11(4): 259–72.

Tikkanen, M J, J K Huttunen, C Ehnholm, and P Pietinen. “Apolipoprotein E4 Homozygosity Predisposes to Serum Cholesterol Elevation during High Fat Diet.” *Arteriosclerosis (Dallas, Tex.)* 10(2): 285–88.

Tosato, G, and K D Jones. 1990. “Interleukin-1 Induces Interleukin-6 Production in Peripheral Blood Monocytes.” *Blood* 75(6): 1305–10.

Valverde, Paloma et al. 1995. “Variants of the Melanocyte–stimulating Hormone Receptor Gene Are Associated with Red Hair and Fair Skin in Humans.” *Nature Genetics* 11(3): 328–30.

Vineis, P, and R Pirastu. 1997. “Aromatic Amines and Cancer.” *Cancer causes & control : CCC* 8(3): 346–55.

Visscher, Peter M. et al. 2006. “Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings.” *PLoS Genetics* 2(3): e41.

Voight, Benjamin F et al. 2010. “Twelve Type 2 Diabetes Susceptibility Loci Identified through Large-Scale Association Analysis.” *Nature genetics* 42(7): 579–89.

———. 2012. “The Metabochip, a Custom Genotyping Array for Genetic Studies of Metabolic, Cardiovascular, and Anthropometric Traits.” *PLoS genetics* 8(8): e1002793.

Walt, D R. 2000. “Techview: Molecular Biology. Bead-Based Fiber-Optic Arrays.” *Science (New York, N.Y.)* 287(5452): 451–52.

Walter, Klaudia et al. 2015. “The UK10K Project Identifies Rare Variants in Health and Disease.” *Nature* 526(7571): 82–90.

Wang, Dong D et al. 2016. “Association of Specific Dietary Fats With Total and Cause-Specific Mortality.” *JAMA internal medicine* 176(8): 1134–45.

Wang, Xiaobin et al. 2002. “Maternal Cigarette Smoking, Metabolic Gene Polymorphism, and Infant Birth Weight.” *JAMA* 287(2): 195–202.

Wen, Wanqing et al. 2014. “Meta-Analysis of Genome-Wide Association Studies in East Asian-Ancestry Populations Identifies Four New Loci for Body Mass Index.” *Human molecular genetics* 23(20): 5492–5504.

Wheeler, Eleanor et al. 2013. “Genome-Wide SNP and CNV Analysis Identifies Common and Low-Frequency Variants Associated with Severe Early-Onset Obesity.” *Nature genetics* 45(5): 513–17.

Wiegman, Albert et al. 2004. “Efficacy and Safety of Statin Therapy in Children with Familial Hypercholesterolemia: A Randomized Controlled Trial.” *JAMA* 292(3): 331–37.

Willer, Cristen J et al. 2008. “Newly Identified Loci That Influence Lipid Concentrations and Risk of Coronary Artery Disease.” *Nature genetics* 40(2): 161–69.

———. 2013. “Discovery and Refinement of Loci Associated with Lipid Levels.” *Nature genetics* 45(11): 1274–83.

Willer, Cristen J, Yun Li, and Gonçalo R Abecasis. 2010. “METAL: Fast and Efficient Meta-Analysis of Genomewide Association Scans.” *Bioinformatics (Oxford, England)* 26(17): 2190–91.

Windecker, Stephan et al. 2014. “Management of ST-Elevation Myocardial Infarction according to European and American Guidelines.” *EuroIntervention* 10(T): T23–31.

Wood, Andrew R et al. 2013. “Imputation of Variants from the 1000 Genomes Project Modestly Improves Known Associations and Can Identify Low-Frequency Variant-Phenotype Associations Undetected by HapMap Based Imputation.” *PloS one* 8(5): e64343.

Yusuf, Salim et al. 2004. “Effect of Potentially Modifiable Risk Factors Associated with Myocardial Infarction in 52 Countries (the INTERHEART Study): Case-Control Study.” *The Lancet* 364(9438): 937–52.

Zeggini, Eleftheria et al. 2008. “Meta-Analysis of Genome-Wide Association Data and Large-Scale Replication Identifies Additional Susceptibility Loci for Type 2 Diabetes.” *Nature genetics* 40(5): 638–45.

Zhang, Y et al. 2000. “Parkin Functions as an E2-Dependent Ubiquitin- Protein Ligase and Promotes the Degradation of the Synaptic Vesicle-Associated Protein, CDCrel-1.” *Proceedings of the National Academy of Sciences of the United States of America* 97(24): 13354–59.

Annales Universitatis Turkuensis

Turun yliopisto
University of Turku

Painosalama Oy, Turku, Finland 2017

ISBN 978-951-29-6913-5 (PRINT)
ISBN 978-951-29-6914-2 (PDF)
ISSN 0355-9483 (PRINT) | ISSN 2343-3213 (PDF)