Tiia Kangas-Kontio

GENETIC BACKGROUND OF HDL-CHOLESTEROL AND ATHEROSCLEROSIS

LINKAGE AND CASE-CONTROL STUDIES IN THE NORTHERN FINNISH POPULATION
TIIA KANGAS-KONTIO

GENETIC BACKGROUND OF HDL-CHOLESTEROL AND ATHEROSCLEROSIS
Linkage and case-control studies in the Northern Finnish population

Academic dissertation to be presented with the assent of the Faculty of Medicine of the University of Oulu for public defence in Auditorium 101 A of the Faculty of Medicine (Aapistie 5 A), on 11 November 2011, at 12 noon

UNIVERSITY OF OULU, OULU 2011
Kangas-Kontio, Tiia, Genetic background of HDL-cholesterol and atherosclerosis. Linkage and case-control studies in the Northern Finnish population
University of Oulu, Faculty of Medicine, Institute of Clinical Medicine, Department of Internal Medicine, Biocenter Oulu, P.O. Box 5000, FI-90014 University of Oulu, Finland; Clinical Research Center, Oulu University Hospital, P.O. Box 5000, FI-90014 University of Oulu
Oulu, Finland

Abstract
Coronary heart disease (CHD), a manifestation of atherosclerosis, is the leading single cause of death in Finland. CHD is affected by numerous genetic and environmental factors, their combined effects and interactions between them. Low HDL-cholesterol (HDL-C) is an independent risk factor for atherosclerosis and the most common dyslipidemia associated with early onset CHD, but the mechanisms regulating HDL-C levels and protecting from atherosclerosis are still not completely understood. Adiponectin is a hormone that is secreted by adipose tissue and has several anti-atherosclerotic effects. There is multiple evidence suggesting that adiponectin could protect against CHD via positive effects on HDL metabolism. Vascular endothelial growth factor (VEGF) is a potent angiogenic growth factor that has a potentially conflicting role in atherosclerosis; it may have protecting or predisposing effects.

The objective of this thesis was to study the genetic background of HDL-C regulation and atherosclerosis. Three studies were executed using extended families with CHD or case-control setting, with samples collected from Northern Finland. In the first study, seven chromosomal regions showing suggestive evidence of linkage were identified for HDL-C regulation, using genome-wide linkage approach. In the second study, we found a strong correlation between HDL-C and adiponectin, but failed to show evidence of a shared genetic background. However, a genetic correlation between adiponectin and low-density lipoprotein-cholesterol was revealed. We also studied the genetic regulation of adiponectin, and for the first time its most active form, high-molecular weight adiponectin, and found suggestive evidence of linkage to three chromosomal regions. In the third study, it was discovered that the studied VEGF gene polymorphisms did not have a major effect on atherosclerosis quantified as carotid intima-media thickness or the risk of acute myocardial infarction (AMI).

This thesis presents potential regions for the genetic regulation of HDL-C and adiponectin and gives new information about their relationship and the effect of VEGF polymorphisms in atherosclerosis. The strong correlation between adiponectin and HDL-C was further strengthened, but we failed to show a shared genetic background between them.

Keywords: adiponectin, coronary heart disease, genetic linkage, genetic loci, genetic pleiotropy, genome-wide, haplotype, HDL-cholesterol, myocardial infarction, single nucleotide polymorphism, vascular endothelial growth factor A
Kangas-Kontio, Tiia, HDL-kolesterolin ja valtimonkovettumataudin geneettinen tausta. Kytkentä- ja tapaus-verrokkitutkimuksia pohjoissuomalaisessa väestössä
Oulun yliopisto, Lääketieteellinen tiedekunta, Kliinisen lääketieteen laitos, Sisäaudit, Biocenter Oulu, PL 5000, 90014 Oulun yliopisto; Kliinisen tutkimuksen keskus, Oulun yliopistollinen sairaala, PL 5000, 90014 Oulun yliopisto
Oulu

Tiivistelmä
Sepelvaltimotauti, eräs valtimonkovettumataudin ilmentymä, on yleisin yksittäinen kuolinsyy maassamme. Taudin syntyn vaikuttavat lukuisat genettiset ja ympäristötekijät sekä niiden väliset yhteis- ja vuorovaikutukset. Pieni HDL-kolesterolipitoisuus on valtimonkovettumataudin itsenäinen riskitekijä ja yleisin kolesterolipoikkeavuus, joka liittyy varhain ilmenevässä sepelvaltimotaudit. HDL-kolesterolin vaihtelun syitä tänän on tämän "lyvän kolesterolin" sepelvaltimotaudillä suojaa vaimutusmekanismeja ei kuitenkaan pystytä täysin selittämään. Adiponektiini on rasvakudoksen tuottama hormoni, jonka sepelvaltimotaudilta suojaa ominaisuuden on ehdottettu johtuvan siitä, että se vaikuttaisi HDL-kolesterolin aineenvaihduntaan. VEGF (vascular endothelial growth factor) on verisuonin sisäseinässä vaikuttava kasvutekijä, jolla saattaa olla joko sepelvaltimotaudilta suojaa tai sille altistavia vaikutuksia.


Tämä tutkimus tuo uutta tietoa HDL-kolesterolin ja adiponektiinin geneettisestä säätelyestä ja niiden suhteesta sekä VEGF-geenin nukleotidimuutosten osuudesta valtimonkovettumataudissa. Tutkimus vahvistaa edelleen HDL-kolesterolin ja adiponektiinin yhteyden, muttei pysty osoittamaan niille yhteistä geneettistä tekijää.

Asiasanat: adiponektiini, geenipaikat, geneettinen pleiotropia, haplotypeppi, HDL-kolesteroli, koko genomin kytentätutkimukset, sepelvaltimotauti, sydäninfarkti, verisuontien endoteelikasvutekijä A, yhden nukleotidin muutos
To my family
Acknowledgements

This study was carried out in the Department of Internal Medicine, Institute of Clinical Medicine and Biocenter Oulu, University of Oulu, and Clinical Research Center, Oulu University Hospital. I wish to acknowledge the institutes for providing excellent research facilities.

I would like to express my sincere gratitude to my supervisors Professor Markku Savolainen and Docent Sakari Kakko for their friendly guidance, invaluable expertise and never-ending optimism. From the beginning, you have encouraged me to find my own way of doing things. Markku has provided brilliant resources for conducting this work and Sakari has always been available for me to consult with research problems and to share thoughts and ideas.

I wish to thank all the collaborators and co-authors, especially the group of Professor Karl-Heinz Herzig for the adiponectin measurements, Professor Juha Kere for the kind and encouraging guidance at the Wellcome Trust Course for Human Genome Analysis and Professor Mika Ala-Korpela for breaking down my article and helping to reconstruct it again. I am grateful for Professor Aarno Palotie and Docent Katriina Aalto-Setälä for reviewing my theses; your inspiring comments are greatly appreciated. Anna Vuolteenaho is thanked for the language revision.

I would like to thank all the workmates in our research group and in CRC and our skilful and hard-working laboratory staff, Malla, Saara and Sari, without whom these projects would not have been possible. Secretary Terttu Niemelä is acknowledged for helping with practicalities. Special thanks go to my roommate Anna-Maria, who created a welcoming and delightful working environment when I came back to work from parental leave and with whom I have had the opportunity to share the excitement and frustration in the last stages of the writing of the thesis. I would also like to acknowledge the unselfish help of Tuire Salonurmi in dilemmas in science and life in general. Additionally, Antti, Tuija, Sanna, Merja, Mirella, Anne, Johanna, Elina, Meiju, Eija and Outi are thanked for the cheerful company and enjoyable conversations during lunch breaks, coffee and elsewhere, and for creating a relaxed atmosphere at work during all these years.

I want to express my heartfelt thanks to all my friends for balancing my life with other interesting matters besides science. Riikka, Piia, Kati, Janette, Paula, Maarit, Mari, Kaisa: thank you for all the memorable moments shared together and for the love and friendship that does not fade with time or geographic
distance. Minna, your sincere and unconditional support and friendship at work and leisure have been exceptional and changed my life when moving to Oulu.

I am indebted to my parents, Tarja and Jussi for their love and invincible faith in me. You have always trusted in my ability to make right decisions despite the occasional “bypaths” and encouraged me to challenge myself and to pursue a fulfilling and interesting career and life. It is too seldom that I express my gratitude for all you have done for me. My brother Tomi is thanked for reminding me to live life to the full! I am grateful for mummu for always being so proud of us and for loudly sticking up for us grandchildren.

My deepest gratitude goes to my dear husband, Ari, with whom I have been privileged to accomplish the most exciting “genetic project” of my life, Rosa. You have showed me the most important things in life so many times in so many ways and your help and support in good times and in bad have meant the world to me. Thank you for standing by me in the adventure of life!

The studies would not have been possible without the volunteer study subjects and I want to express my gratitude especially towards the multiple Northern Finnish family members who enabled us to assemble exceptional pedigrees. This work was financially supported by the Academy of Finland, the Finnish Academy of Science, the Finnish Cultural Foundation, the Finnish Foundation for Alcohol studies, the Finnish Foundation for Cardiovascular Research, the Finnish Medical Foundation, the Finnish Medical Society Duodecim, the Ida Montin Foundation, the Onni and Hilja Tuovinen Foundation, the Oulu University Scholarship Foundation, the Orion-Farmos Research Foundation, the Paavo Nurmi Foundation, the Sigrid Jusélius Foundation and the US National Institutes of Health. They are all gratefully acknowledged.

Oulu, September 28th 2011

Tiia Kangas-Kontio
Abbreviations

ABCA1 ATP binding cassette sub-family A, member 1
ADIPOQ adiponectin gene
AMI acute myocardial infarction
apoA-I apolipoprotein A-I
APOA1 apoA-I gene
BIF bifurcation enlargement
BMI body mass index
CCA common carotid artery
CE cholesteryl esters
CETP cholesterol ester transfer protein
CHD coronary heart disease
CombIMT combined IMT–plaque thickness
CRP C-reactive protein
DNA deoxyribonucleic acid
GWAS genome-wide association studies
HDL high-density lipoprotein
HDL-C high-density lipoprotein cholesterol
HL hepatic lipase
HMW high-molecular weight
IBD identical by descent
IBS identical by state
ICA internal carotid artery
IDL intermediate-density lipoprotein
IMT intima-media thickness
LCAT lecithin-cholesterol acyltransferase
LD linkage disequilibrium
LDL low-density lipoprotein
LDL-C low-density lipoprotein cholesterol
LDL-PLA2 LDL-associated phospholipase A2
LOD score logarithm-of-odds score
LPL lipoprotein lipase
mRNA messenger ribonucleic acid
NPL non-parametric linkage
OPERA Oulu Project Elucidating Risk of Atherosclerosis
PLTP phospholipid transfer protein
PPARD: peroxisome proliferator-activated receptor delta
PPARG: peroxisome proliferator-activated receptor gamma
PPARGC1A: peroxisome proliferator-activated receptor gamma, coactivator 1 alpha
QTL: quantitative trait locus
RCT: reverse cholesterol transport
RXRB: retinoid X receptor, beta
SNP: single nucleotide polymorphism
SR-B1: scavenger receptor class B, type 1
SSRs: simple sequence repeats
TG: triglycerides
VCLA: variance component linkage analysis
VEGF: vascular endothelial growth factor
VLDL: very-low-density lipoprotein
List of original articles

The thesis is based on the following original articles, which are referred to in the text by their Roman numerals:


*with equal contribution
Contents

Abstract .............................. 9
Tiivistelmä .................................. 10
Acknowledgements .................... 9
Abbreviations .......................... 11
List of original articles ............... 13
Contents ................................. 15

1 Introduction ......................... 19
2 Review of the literature .......... 23
   2.1 Atherosclerosis – a complex disease .............................................. 23
   2.2 Pathogenesis and risk factors of atherosclerosis ............................. 24
   2.3 HDL, adiponectin and vascular endothelial growth factor .......... 25
       (VEGF): factors protecting from atherosclerosis? ......................... 25
       2.3.1 HDL .................................................................................. 25
       2.3.2 Adiponectin ..................................................................... 29
       2.3.3 VEGF ............................................................................. 32
   2.4 Structure and analysis of the human genome ............................... 33
       2.4.1 Organization of the human genome ................................... 34
       2.4.2 Linkage and linkage disequilibrium ................................. 36
   2.5 Genetic analysis of complex diseases ........................................... 36
       2.5.1 Selection of the study population ................................... 36
       2.5.2 Traditional linkage and modern association? ...................... 37
       2.5.3 Non-parametric linkage analysis ................................... 38
       2.5.4 Heritability ................................................................. 40
       2.5.5 Variance component linkage analysis ............................ 40
       2.5.6 Bivariate linkage analysis ............................................. 41
       2.5.7 Association and haplotype analysis ............................... 42
   2.6 Genetics of HDL-cholesterol (HDL-C), adiponectin and VEGF ........ 42
       2.6.1 HDL-C ............................................................................ 42
       2.6.2 Adiponectin .................................................................... 44
       2.6.3 VEGF ............................................................................ 45

3 Aims of the study ................. 47

4 Subjects and methods ............ 49
   4.1 Study subjects ................................................................. 49
       4.1.1 Northern Finnish families with early onset CHD and low HDL-C (I, II) 49
4.1.2 OPERA controls (III) ................................................................. 49
4.1.3 Acute myocardial infarction (AMI) survivors (III) .......... 50
4.2 Methods ..................................................................................... 50
4.2.1 Selection and genotyping of the polymorphisms (I, II, III) ...... 50
4.2.2 Laboratory methods ............................................................... 52
4.2.3 Measurement of intima-media thickness (IMT) (III) ........... 53
4.2.4 Statistical methods ................................................................. 53

5 Results .......................................................................................... 57
5.1 Genetic loci regulating HDL-C levels in Northern Finnish families with early onset CHD and low HDL-C levels (I) .......... 57
5.1.1 Basic characteristics, heritability and significance
thresholds .................................................................................. 57
5.1.2 Quantitative and qualitative linkage analysis ................. 57
5.1.3 Association analysis ............................................................. 63
5.2 Genetic and environmental determinants of total and high-
molecular weight adiponectin in Northern Finnish families with early onset CHD and low HDL-C levels (II) .................. 63
5.2.1 Sex difference and association with CHD ...................... 63
5.2.2 Correlations of the adiponectins and lipids ...................... 64
5.2.3 Heritability and linkage analyses ....................................... 64
5.3 VEGF gene in atherosclerosis, quantified as carotid IMT and the risk of AMI (III) ................................................................. 66
5.3.1 Power analysis ................................................................. 66
5.3.2 The effect of VEGF variation on IMT and on the risk of AMI ................................................................. 66
5.3.3 The effect of VEGF haplotypes on IMT and on the risk of AMI ................................................................. 68

6 Discussion ................................................................. 71
6.1 Genetic loci regulating HDL-C levels (I) ..................... 71
6.2 The adiponectins and lipids in atherosclerosis – genetic and environmental determinants (I, II) ...................... 73
6.2.1 The relationship between the adiponectins ......................... 73
6.2.2 Genetic and environmental correlations ............................. 73
6.2.3 A link between adiponectin, HDL-C and LDL-C? .............. 74
6.2.4 Adjusting for BMI reveals a possible link between total adiponectin and apolipoproteins ............................. 75
6.2.5 Overlapping linkage regions on chromosome 6 ............. 76
6.2.6 The adiponectins and lipids in atherosclerosis - conclusions................................................................................... 77
6.3 VEGF and atherosclerosis (III)............................................................. 79
6.4 Combined results................................................................................... 80
6.5 Methodological considerations ........................................................... 81
   6.5.1 Genotypings and study populations.............................................. 82
   6.5.2 IMT............................................................................................... 83
6.6 Future studies...................................................................................... 83

7 Conclusions .......................................................................................... 85
References .................................................................................................. 87
Original articles ......................................................................................... 111
1 Introduction

Atherosclerosis is the leading cause of death in Western societies and it manifests mainly as peripheral vascular disease, cerebrovascular disease and coronary heart disease (CHD). In Finland CHD is the cause of every fourth death in the working-age population aged 15–64 and the most common single cause of death in the whole population (Official Statistics of Finland (OSF) 2009). In general, the factors protecting from atherosclerosis have beneficial effects on the lipid profile, body mass index (BMI), glucose metabolism and vascular biology. Among others, high-density lipoprotein cholesterol (HDL-C), adiponectin and vascular endothelial growth factor (VEGF) are associated with these protective effects.

Low plasma HDL-C level is an independent risk factor for atherosclerosis (Gordon & Rifkind 1989, NCEP expert panel 2002) and it is the most common dyslipidemia associated with premature and familial CHD (Genest et al. 1992). The earliest findings of the beneficial effects of HDL were made already in the 1950s, and the Finnish internist Esko Nikkilä was among the first researchers to discover this effect. The mechanisms by which HDL protects from atherosclerosis are still partly unknown, but reverse cholesterol transport (RCT) is thought to represent the basis for its anti-atherogenic properties (von Eckardstein et al. 2001, Kontush & Chapman 2006b). In this process HDL removes excess cholesterol from the cells in the arterial wall to the liver to be excreted (Glomset 1968, Bruce et al. 1998). Deeper understanding of HDL-C metabolism would give us new tools for atherosclerosis prevention and treatment, which would be greatly improved, if HDL function could be made more efficient with drugs. The current medication is based on statins and the lowering of low-density lipoprotein cholesterol (LDL-C), which decrease the risk of major coronary events by 15% to 40% even though statins consistently and efficiently lower LDL-C (Baigent et al. 2005). It has also been argued that a reduction in coronary endpoints beyond 20–25% is not achievable by LDL-C lowering alone (Linsel-Nitschke & Tall 2005). Therefore, there is considerable interest in the therapeutic potential of increasing HDL-C levels. The identification of new susceptibility genes and disease risk markers could also enable us to detect some high-risk individuals earlier and focus on effective primary prevention.

In Finland CHD mortality, especially in men, has decreased dramatically (80%) in 35 years (Vartiainen et al. 2010). It has been estimated that 75% of the observed decline in coronary mortality in middle-aged men could be explained by the decline in blood pressure, cholesterol and smoking, but since the mid-1980s
advances in medicine might explain most of the remaining decline (Laatikainen et al. 2005). However, a new concern is a worldwide epidemic of obesity, which is setting the scene for a new wave of premature cardiovascular disease. Obesity also increases the risk of insulin resistance and diabetes, and it is very alarming that obesity and type 2 diabetes have increased even among children especially in Western societies during the last few years. Adiponectin, an adipocytokine secreted by adipose tissue, has several anti-atherosclerotic effects, and it has repeatedly been associated with HDL-C levels and suggested to be involved in HDL-C regulation by affecting RCT. Low levels of adiponectin are strongly correlated with low HDL-C levels and high BMI, both associated with CHD, obesity and diabetes. Hypoadiponectinemia could be a mechanistic link between visceral obesity, insulin resistance and dyslipidemia and adiponectin could be a molecular link between metabolic syndrome and atherosclerosis. Therefore, better understanding of the adiponectin metabolism and regulation could give us new targets for prevention of obesity-related diseases.

VEGF is a potent angiogenic growth factor that may have a dual role in atherosclerosis. It is possible that VEGF can serve as a protecting factor by providing relief for tissue ischemia caused by atherosclerosis by enhancing neovascularization, but on the other hand, plaque angiogenesis induced by VEGF could lead to the growth and rupture of these atherosclerotic plaques causing acute myocardial infarction (AMI). Studies to date have had promising results with treatments where VEGF DNA is injected into the heart muscle, stimulating the growth of new blood vessels to feed the heart tissue no longer efficiently served by the old blood vessels. This kind of experimental treatment has caused improvement in the symptoms of angina pectoris and in the contractile function of the heart muscle (Reilly et al. 2005) (Alber et al. 2005). On the other hand, a substantial concern with this treatment is that it may cause accelerated atherosclerosis, since the neovascularization of atherosclerotic lesions has been correlated with their risk of rupture in human aorta (Moreno et al. 2004). Therefore, more studies about the effect of the VEGF gene in atherosclerosis are needed before the experimental treatments can be offered to other than no-option patients.

The intimate relationship between genetics and epidemiology was discussed by Neel and Schull (Neel & Schull 1954), only a year after Watson and Crick reported the discovery of the deoxyribonucleic acid (DNA) double helix. Genetic epidemiology studies the role of genetic factors in health and disease. Its primary goal is to resolve whether the studied trait has a genetic component and what the
relative contributions of genes and environment are. This thesis was planned to disentangle the complex genetic background of HDL-C regulation and CHD and to clarify the role of adiponectin in the process. In this study, we used a genome-wide linkage approach to find genetic loci that regulate HDL-C and predispose to CHD. We also studied the relationship between the adiponectins (total adiponectin and its most active form, high-molecular weight (HMW) adiponectin) and conventional CHD risk factors, and investigated genetic and environmental determinants and genetic regulation of the adiponectins. In addition, we aimed at revealing a possible common genetic background for the adiponectins and HDL-C. Finally, we investigated whether the polymorphisms of the VEGF gene, previously associated with VEGF levels, were associated with intima-media thickness (IMT), a surrogate marker of atherosclerosis, and the risk of AMI.
2 Review of the literature

2.1 Atherosclerosis – a complex disease

Fig. 1. Coronary heart disease (CHD) is a complex disease. CHD is affected by numerous environmental and genetic factors. Some of the factors predisposing to CHD are related to our lifestyle choices, others are unchangeable, such as genes and gender. Obesity, lipid abnormalities, blood pressure and thrombotic propensity are called intermediate traits that can be quantitatively measured. They are, as well, affected by environmental and genetic factors and involved in the development of CHD. All these different factors have combined effects and interactions between them and a lot of the mechanisms are unknown. CHD shares many predisposing factors and traits with other, associated diseases, such as metabolic syndrome and diabetes, which possibly reflects shared factors and mechanisms in the development of the diseases.

Atherosclerosis is a complex disease affected by multiple environmental and genetic factors, their combined effects and interactions between them (Figure 1). It is not simply an inevitable degenerative consequence of ageing, but rather a chronic inflammatory condition that can be converted into an acute clinical event by plaque rupture and thrombosis (Lusis 2000). Atherosclerosis is also not
inevitably progressive, since it can be arrested and even relieved by lifestyle changes and drug treatment.

2.2 Pathogenesis and risk factors of atherosclerosis

Atherosclerosis was previously identified as a lipid-storage disease, but it is now recognized as a subacute inflammatory condition of the vessel wall, characterized by infiltration of macrophages and T cells, which interact with one another and with cells of the arterial wall (Rocha & Libby 2009). Atherosclerosis is a progressive disease where the accumulation of lipids and fibrous tissue forms lesions on the arterial wall, which can lead to the thickening of intima, the innermost layer of the artery, and narrowing of the artery. The formation of a lesion is a slow process and can begin already in the childhood. Fatty streaks are prevalent even among young people and may progress to atherosclerotic lesion or eventually disappear. The core of the lesion is formed by extracellular lipid droplets and foam cells, which are surrounded by a cap of smooth muscle cells and collagen-rich matrix. T-cells, macrophages and mast cells infiltrate the lesion and are abundant in the region where the lesion grows. A stable lesion can transform into a vulnerable, unstable structure that can rupture, induce a thrombus, and elicit an acute coronary syndrome (Hansson 2005).

Dyslipidemia (meaning low HDL-C, high LDL-C or high triglycerides (TG) or combinations of these) smoking, physical inactivity, obesity, hypertension, diabetes and genetic predisposition are well-established risk factors for atherosclerosis. In addition, certain environment or lifestyle can be particularly harmful for an individual having a certain genetic background. For example, Western-style diet high in saturated fat and cholesterol could be especially hazardous for a person having multiple genetic factors predisposing to problems in RCT.

Different risk factors are emphasized in different environments and are therefore associated with time, culture, nationality and population. In 1962, James V. Neel introduced the thrifty gene hypothesis (Neel 1962), which is based on the fact that for thousands of years human populations experienced alternating periods of feast and famine and in order to adapt to these extreme changes in caloric needs, people developed a thrifty gene that allowed them to store fat during times of plenty so that they would not starve during times of famine. This explains the current increase in obesity and related diseases by having “stone-age genes in a space-age environment” (Eaton et al. 1998). In this bountiful
environment, the thrifty genes that provided a reproductive advantage to those individuals who could efficiently store energy (in the form of fat) are now the same genes that confer greater susceptibility to the so-called life-style diseases. Contemporary humans are, in the genetic sense, still Stone Agers and remain adapted to a preagricultural nutritional pattern (Eaton & Eaton, III 2000).

There is also a sex difference in the disease risk, since men have a greater risk of developing early CHD than women. An individual’s age at disease onset, the rate of progression and the severity of the disease are affected by the sum of the predisposing and protecting factors. Quantitatively measurable risk factors, such as obesity, lipid abnormalities, blood pressure and thrombotic propensity are called intermediate traits. The division of the overall CHD risk into quantitative intermediate traits can be used to make the hunt for atherosclerosis genes somewhat easier.

2.3 HDL, adiponectin and vascular endothelial growth factor (VEGF): factors protecting from atherosclerosis?

Classical factors protecting from atherosclerosis are mostly the same as the risk factors - however, with the opposite values or activities. For example, a sufficient amount of exercise, a healthy diet and moderate alcohol drinking usually have beneficial effects on the lipid profile, BMI, glucose metabolism and vascular biology. HDL-C, adiponectin and VEGF are associated with these beneficial effects. In addition to having been associated with environmental factors, their genetic backgrounds have been actively studied. Multiple genome scans have been published for loci affecting HDL-C and adiponectin levels and variations in the VEGF-gene have been shown to affect their plasma concentrations. The roles of HDL-C, adiponectin and VEGF in atherosclerosis and their complex relationships with other protecting or risk factors are still somewhat unclear.

2.3.1 HDL

Lipids are necessary for us. They are needed as components of cell structure, in the synthesis of bile acids and steroid hormones and as a source of energy. The main lipids in human plasma are cholesterol, cholesteryl esters (CE), TG and phospholipids. Because lipids are hydrophobic, they have to be transported in the circulation as components of water-soluble lipoproteins, which are commonly divided into five density classes by ultracentrifugation: chylomicrons, very-low-
density lipoproteins (VLDL), intermediate-density lipoproteins (IDL), LDL and HDL (Havel et al. 1955). LDL is the major carrier of cholesterol in humans and responsible for delivering cholesterol to the peripheral cells, whereas HDL transfers cholesterol from the cells to the liver (Figure 2). HDL is a class of heterogeneous lipoproteins with various subclasses, which vary in quantitative and qualitative content of lipids, apolipoproteins, enzymes and lipid transfer proteins. A high concentration of plasma HDL-C has consistently been found to be a powerful negative predictor of premature CHD and there is strong evidence that HDLs protect against atherosclerosis (Gordon et al. 1977, Barter & Rye 1994). A meta-analysis (Lewington et al. 2007) with HDL-C measures available in 150,000 individuals demonstrated a strong negative association with CHD mortality, with no evidence of a threshold beyond which higher HDL-C was no longer associated with lower mortality. On average 0.34mmol/l higher HDL-C was associated with about a third lower CHD mortality.

HDL-C and reverse cholesterol transport

RCT is thought to be one of the most important mechanisms by which HDL can protect from atherosclerosis and counteract the effects of LDL by removing excess cholesterol from the cells in the arterial intima. In the process cholesterol is effluxed by peripheral cells onto acceptor particles for transport to the liver for excretion as bile (Figure 2). In RCT apolipoprotein A-I (ApoA-I) and lipid-poor pre-β HDL, rich in Apo-AI is synthesized in the liver or intestinal mucosa and released into circulation. There it promotes the transfer of excess cellular-free cholesterol (FC) from macrophages to Apo-AI by interacting with ATP-binding cassette transporter A1 (ABCA1) in arterial-wall macrophages. Plasma lecithin-cholesterol acyltransferase (LCAT) converts FC in pre-β HDL to CE allowing more efficient packaging of the cholesterol for transport and resulting in maturation of pre-β HDL to mature HDL. Phospholipid transfer protein (PLTP) transfers phospholipids from TG-rich lipoproteins to HDL and regulates the size of HDL particles. ABCG1 and ABCG4 stimulate cholesterol efflux from macrophages to both smaller (HDL-3) and larger (HDL-2) subclasses but not to lipid-poor apoA-I (Wang et al. 2004). HDL-C is transported to the liver by direct or indirect pathway. In the direct pathway selective uptake of CE by hepatocytes occurs with the scavenger receptor, class B, type 1 (SR-B1). In the indirect pathway, CE in HDL are exchanged for TG in apolipoprotein-B-rich particles, LDL, VLDL or IDL through cholesterol-ester transfer protein (CETP), with
uptake of CE by the liver through the LDL receptor (LDL-R). Cholesterol that is returned to the liver is secreted as bile acids and cholesterol. Acquired TG in the modified HDL particle are subjected to hydrolysis by hepatic lipase (HL) expressed in the liver, thereby regenerating small HDL particles and pre-β HDL for participation in RCT (Ashen & Blumenthal 2005). In the hypertriglyceridemic state, CETP is highly expressed and RCT leads to the formation of small dense low-density lipoprotein (LDL) and small dense HDL, both of which are involved in the progression of atherosclerosis (Kolovou et al. 2009). RCT is a multistep process affected by several different enzymes, transfer proteins and receptors, and failure in one of these steps may cause problems in removing excess cholesterol and therefore predispose to dyslipidemia and atherosclerosis. Genetic deficiencies in these key molecules, such as mutations in the genes of ABCAI, HDL apolipoproteins, lipoprotein lipase (LPL), HL and LCAT may alter HDL function.

**Other anti-atherosclerotic actions of HDL**

In addition to RCT, normal HDL has multiple anti-atherosclerotic properties (Figure 3). It is anti-inflammatory (Cockerill et al. 1995), anti-oxidative (Parthasarathy et al. 1990), anti-apoptotic (Sugano et al. 2000) and anti-thrombotic (Viswambharan et al. 2004) and can contribute to endothelial function and vasodilation (Spieker et al. 2002). However, the importance of the single mechanisms in atherosclerosis prevention remains uncertain. Atherosclerosis is a multistep process, and the various properties of HDL can intervene at different stages. For example, there is evidence that HDL can inhibit oxidative modification of LDL within the arterial intima, which is an early event in atherosclerosis. Anti-inflammatory HDLs can inhibit the cell-surface expression of adhesion molecules, such as VCAM-1 and E-selectin, and therefore the adhesion of monocytes to endothelial cells, which is one of the earliest events in atherogenesis (Barter et al. 2003, Kontush & Chapman 2006a).
Fig. 2. Lipoprotein metabolism: emphasizing reverse cholesterol transport (RCT). In RCT HDL transports cholesterol from the peripheral cells to the liver to be excreted. ABCA1/G1/G4, ATP binding cassette sub-family A/G/G, member 1/1/4; apoA-I, apolipoprotein A-I; CETP, cholesterol ester transfer protein; HDL, high-density lipoprotein; HL, hepatic lipase; IDL, intermediate-density lipoprotein; LCAT, lecithin-cholesterol acyltransferase; LDL, low-density lipoprotein; LDL-R, low-density lipoprotein receptor; LPL, lipoprotein lipase; PLTP, phospholipid transfer protein; SR-B1, scavenger receptor class B, type 1; VLDL, very-low-density lipoprotein.
2.3.2 Adiponectin

Adipose tissue is an active endocrine organ, secreting a number of bioactive substances known as adipocytokines or adipokines (Shimomura et al. 1996, Yamauchi et al. 2003). One of these adipokines is adiponectin, which is an important circulating plasma protein that exhibits beneficial effects on lipid metabolism, glucose homeostasis, vascular biology, inflammation and endothelial functions (Ouchi et al. 1999, Hotta et al. 2000, Ouchi et al. 2001, Okamoto et al. 2002, Ouchi et al. 2004, Ohashi et al. 2010) (Figure 4). Although adiponectin is secreted from adipose tissue, obese subjects have significantly lower plasma concentrations of adiponectin than non-obese subjects (Arita et al. 1999). In addition, it has been shown that adiponectin mRNA levels are lower in visceral versus subcutaneous adipose tissue (Lihn et al. 2004). Concentrations are also gender-dependent as women have higher adiponectin levels than men.

Adiponectin has several anti-atherosclerotic effects (Goldstein et al. 2009). Low plasma levels of adiponectin are associated with early CHD onset and multiple atherosclerotic lesions in coronary arteries (Hashimoto et al. 2006).
addition, low level of adiponectin is an independent risk factor for CHD (Frystyk et al. 2007, Persson et al. 2010) and has been suggested to predict the progression of coronary atherosclerosis, quantified as coronary artery calcification (Maahs et al. 2005). Furthermore, high plasma adiponectin concentrations are associated with lower risk of AMI in men (Pischon et al. 2004). In patients with CHD, HDL-C concentration is an independent predictor of the serum adiponectin value (Wang et al. 2010a). Accumulating evidence suggests that adiponectin could protect against CHD via positive effects on HDL metabolism (Verges et al. 2006, Matsuura et al. 2007, Lara-Castro et al. 2007, Tsubakio-Yamamoto et al. 2008) (Figure 5) or that HDL could affect adipocyte metabolism and adiponectin expression (Van Linthout et al. 2010). The positive correlation between plasma adiponectin and HDL-C is strong and independent of BMI or visceral adiposity (Cnop et al. 2003, Okada et al. 2006, von Eynatten et al. 2006). This could be due to a shared genetic or environmental background regulating both levels simultaneously. Low adiponectin levels are consistently associated with diabetes and metabolic syndrome, also sharing a low HDL-C phenotype, which emphasizes the strong and likely functional correlation between adiponectin- and HDL-C-levels. Hypoadiponectinemia could be a mechanistic link between visceral obesity, insulin resistance and dyslipidemia, and adiponectin could be a molecular link between metabolic syndrome and atherosclerosis.

Adiponectin is secreted by adipocytes in three different forms, namely as trimers, low-molecular weight hexamers (LMW) and high-molecular weight (HMW) isoforms (12-18 mers) (Wang et al. 2006). The HMW form is highly active in endothelial cells (Ouchi et al. 2004) and liver, but less relevant in skeletal muscle (Wang et al. 2006). Different multimeric forms can determine the activity of adiponectin, and multiple studies (Kobayashi et al. 2004, Pajvani et al. 2004, Aso et al. 2006, Inoue et al. 2007, von Eynatten et al. 2008) have suggested that higher proportion of HMW multimers, not the absolute amount of plasma total adiponectin, is clinically more important, at least with respect to CHD.
Fig. 4. Adiponectin and atherosclerosis. Inflammation plays an important part in the development of atherosclerosis and metabolic syndrome. Obese adipose tissue is characterized by adipocyte hypertrophy and infiltration of inflammatory macrophages and lymphocytes, leading to increased production of pro-inflammatory adipokines. By contrast, the secretion of adiponectin, an adipokine with anti-inflammatory activities, is decreased in obesity. Adiponectin has beneficial effects on lipid metabolism, glucose homeostasis and atherosclerosis, thus reducing the classical risk factors for CHD. Accordingly, low levels of adiponectin are associated with CHD, metabolic syndrome and diabetes. ABCAI, ATP-binding cassette, sub-family A (ABC1), member 1; apo-AI, apolipoprotein A-I; BMI, body mass index; CHD, coronary heart disease; CRP, C-reactive protein; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TG, triglycerides.
3.2.3 VEGF

VEGFs are potent mitogens for vascular endothelial cells (Ferrara et al. 2003, Roy et al. 2006). VEGF-A (hereafter VEGF) is a potent angiogenic growth factor that may have a dual role in atherosclerosis. It is possible that VEGF can serve as
a protecting factor by providing relief for tissue ischemia caused by atherosclerosis by activating neovascularization, but on the other hand, plaque angiogenesis induced by VEGF could lead to the growth of atherosclerotic plaques and these neovascularized plaques could be vulnerable and unstable making the coronary arteries potentially susceptible to AMI. It has been argued that plasma VEGF levels are not correlated with the extent of coronary artery disease in humans (Alber et al. 2005). However, it has been shown that injections of DNA containing VEGF gene into the myocardium, improve the symptoms of angina pectoris and the contractile function of the myocardium (Gyongyosi et al. 2005, Reilly et al. 2005). On the other hand, the neovascularization of atherosclerotic lesions in human aorta is correlated with their risk of rupture (Moreno et al. 2004) and there is evidence that VEGF expression is increased in human endothelial cells collected from advanced atherosclerotic lesions compared with endothelial cells taken from early atherosclerotic lesions, whereas endothelial cells from normal arteries do not express VEGF (Morsi et al. 2006). In addition, plasma VEGF level was elevated in patients with multivessel CHD (Nakajima et al. 2004) and in a population-based prospective, case-cohort study, baseline levels of VEGF showed a significant independent association with the risk of CHD death (Eaton et al. 2008).

Lower VEGF levels have also been linked to beneficial effects of statins, providing a possible mechanism of plaque stabilization in patients with CHD. High-dose atorvastatin reduced plasma VEGF levels in patients with diabetes and CVD. This effect was independent of LDL-C and total cholesterol changes but was associated with HDL-C changes (Jaumdally et al. 2010). Also rosuvastatin therapy decreased VEGF serum levels significantly (Sergienko et al. 2007). On the other hand, if VEGF was a protecting factor, it could also play a part in ethanol-induced protective effects against CHD, since HDL particles containing phosphatidylethanol have been suggested to bind to CLA-1 receptor and thereby increase the secretion of VEGF from endothelial cells (Liisanantti & Savolainen 2009).

2.4 Structure and analysis of the human genome

“Humans are much more than simply the product of a genome, but in a sense we are, both collectively and individually, defined within the genome. The mapping, sequencing and analysis of the human genome is therefore a fundamental advance in self-knowledge; it will strike a personal chord with
many people. And application of this knowledge will, in time, materially benefit almost everyone in the world.” (Editors of Nature 2001)

2.4.1 Organization of the human genome

The human nuclear genome is organized into 23 chromosome pairs. Each individual receives one chromosome of each type from each parent so that an individual’s genome consists of 22 autosomal chromosome pairs and a sex chromosome pair, either XX (female) or XY (male). These chromosomes contain an estimated 20,000 – 25,000 protein-coding genes and 3 billion base pairs of double-stranded DNA. DNA is the chemical name for the molecule that carries genetic instructions in all living things. Each strand of DNA has a backbone made of alternating sugar (deoxyribose) and phosphate groups and attached to each sugar is one of four bases - adenine (A), cytosine (C), guanine (G), and thymine (T). Genes are specific sequences of bases that encode instructions on how to make proteins. Human genes consist of exons that determine the amino acid sequence of the nascent proteins and introns (intervening sequences) that are located between the exons. Genes comprise only about 2% of the human genome; the rest consists of noncoding regions, whose functions remain still largely uncharacterized but which may provide chromosomal structural integrity and regulate protein production.

The sequence of the human genome encodes the genetic instructions for human physiology, as well as rich information about human evolution. The Human Genome Sequencing Consortium reported the first draft sequence of the euchromatic portion of the human genome in 2001 and the result of the finishing process was published in 2004, covering 99 percent of the euchromatic genome, having an error rate of only 1 event per 100,000 bases and containing 2.85 billion nucleotides (International Human Genome Sequencing Consortium 2004).

Variation in the human genome

Different kinds of variations and polymorphisms in the genome sequence constitute the basis of human inherited diseases. In addition to single nucleotide polymorphisms (SNPs), which by definition differ between individuals only by one base pair (bp), there are multiple types of variations of various sizes: insertions, duplications, deletions, translocations and repetitive elements such as copy-number variants (may range from thousand bases to several million bases in
size) and simple sequence repeats (SSRs). SSRs with a short repeat unit \((n=1-13\) bases) are termed microsatellites, whereas those with longer repeat units \((n=14-500\) bases) are often termed minisatellites. These sequence repeats are highly prone to DNA replication errors due to replication slippage and tend to be duplicated in tandem. Therefore, the mutation rates of microsatellite loci are substantially higher compared to other loci and they are highly polymorphic with up to 10 alleles per locus. There is approximately one SSR every 2 kb and SSRs comprise about 3\% of the human genome, with the greatest single contribution coming from dinucleotide repeats \((0.5\%)\) (Lander et al. 2001). Biallelic SNPs are the most common variation in the human genome. SNPs are historical point mutations that have accumulated and persisted over the course of evolution either because they have had some beneficial effect in some environment at some point in time and have been a target for natural selection or because they have been selectively neutral and have persisted simply due to chance. SNPs are very frequent: It was suggested in 2001 that there are approximately 8 million common SNPs in European populations, estimated in every 300 bp (Kruglyak & Nickerson 2001). By 2008 the public catalogue of variant sites \((\text{dbSNP 129})\) contained approximately 11 million SNPs (Sachidanandam et al. 2001, International HapMap Consortium. 2005, Frazer et al. 2007), and in 2010 the location, allele frequency and local haplotype structure of approximately 15 million SNPs was described (Durbin et al. 2010). SNPs in coding regions of a gene are thought to be involved in causing genetic diseases because they may change the structure and function of the encoded protein, but in addition, it has been realized that SNPs in non-coding regions may also affect the regulation, splicing, and messenger ribonucleic acid \((\text{mRNA})\) stability of a gene (Pagani & Baralle 2004, Capon et al. 2004). These sequence and length polymorphisms can also be used as markers of the sequence in hunting for susceptibility genes. SNPs are commonly used in association studies, whereas microsatellite markers are rarer (there are more than 300,000 known microsatellites in the human genome \((\text{NCBI Build 36.1})\) but more informative and are commonly used in linkage studies.

Instead of using a single polymorphism, one can use haplotypes, which are combinations of alleles at multiple loci that are transmitted together on one chromosome, and therefore, contain more information than individual unorganized SNPs. Recent studies have shown that the human genome has a haplotype block structure. The genome can be divided into discrete blocks that have limited haplotype diversity. A typical feature of a haplotype block is that only three to five common haplotypes cover more than 90\% of all haplotypes
observed (Gabriel et al. 2002). In each block, a small fraction of SNPs, referred to as "tag SNPs," can be used to distinguish a large fraction of the haplotypes. These tag SNPs are useful for association studies and it has been reported that these haplotype-based analyses can be much more powerful than marker-by-marker analysis (Zhang et al. 2002). In fact, because SNPs that lie close together on chromosomes often tell quite similar stories, researchers are able to explore the variation through analyzing a subset of all the SNPs; half a million is usually used in genome-wide association studies (GWAS).

2.4.2 Linkage and linkage disequilibrium

Alleles that exist close to each other in a chromosome tend to be inherited together more often than would be expected by chance. This phenomenon is called linkage. Linkage disequilibrium (LD) is very tight linkage and synonymous with allelic association. It means that even on the population level, certain alleles tend to co-occur together in a chromosome more often than expected based on their overall population frequencies. Polymorphisms that are in LD with each other have not been separated by recombinations or new mutations and thus share a common ancestry (Gabriel et al. 2002). Therefore, one should keep in mind that in population-based association analysis one is actually assuming a very large family where all the affecteds descend from a common ancestor and have the same mutation in the same ancestral chromosome.

2.5 Genetic analysis of complex diseases

2.5.1 Selection of the study population

There are multiple challenges complicating the mapping of complex diseases such as late onset of the disease, unknown mode of inheritance, phenocopies, difficulties in diagnosis, low penetrance, genetic heterogeneity and common disease-predisposing alleles with minor effects. Focusing on well-characterized extended families with multiple cases having an early onset of the disease, should decrease the probability of phenocopies (i.e. an environment-induced phenotype that closely resembles a phenotype produced by a genotype) and genetic heterogeneity, compared with affecteds selected from the general population.
In isolated populations, such as Finland, genetic and environmental heterogeneity are reduced. The history of the Finnish population can briefly be described as having a small number of original founders, isolation for the past centuries for cultural, linguistic and geographical reasons, multiple bottlenecks due to wars and diseases and finally an exponential growth of the population since the 18th century. These factors have made the Finnish population homogeneous and ideal for genetic studies. Considering the length of the LD maps only, it has been suggested that association analyses in samples from Finland and some other population isolates, would require at least 30% fewer markers than studies in outbred populations (Service et al. 2006). Accordingly, population isolates that have experienced marked recent population growth, starting from a small founding population, have higher overall levels of LD compared to outbred populations, and they have far fewer regions of very low LD.

2.5.2 Traditional linkage and modern association?

Traditional parametric linkage analysis with a few well-characterized families with a monogenic disease running in the family created the basis for non-parametric genome-wide linkage studies of polygenic diseases, which typically consisted of around 400 microsatellite markers covering the whole genome with 10 cM intervals. These linkage analyses demanded hundreds of samples from well-characterized families and were usually followed by fine mapping of the discovered linkage regions using additional microsatellites and SNPs. Today the trend is to carry out GWAS using thousands or tens of thousands of unrelated samples and around 300–500,000 SNPs straight away. This is the number of SNPs that is needed to cover most of the LD or allelic variation in European descent populations, and large sample sizes are needed because each of the loci has a small size effect. There has been a dramatic change in the nature of the analysis of complex diseases and new approaches have been enabled by a decrease in the financial costs and a remarkable increase in the human genome sequence data available, thanks to the prominent human genome sequencing project. However, in the age of GWAS one should not underestimate the importance of the “old” linkage studies, since the approaches adopted in linkage and association studies have different strengths and weaknesses: linkage analysis can provide greater power to capture a rare variant with a large effect, whereas association studies have a greater power to detect common variants with small effects. Association studies provide greater resolution of location than linkage
studies, but unlike population-based studies, family-based designs are robust against population substructure. In addition, family-based studies offer a solution to the problem of multiple-hypothesis testing, in which many genetic markers are tested against many different phenotypes, which is an important issue in association studies (Laird & Lange 2006). Numerous GWAS have recently been performed producing substantial amount of data. However, it has been estimated that all the susceptibility loci discovered in these extensive GWAS together would only explain about 10 percent of the variability of HDL-C levels (Sabatti et al. 2009, Teslovich et al. 2010). Population and family designs should be viewed as complementary, not competitive approaches. Sometimes studying families with multiple members that have the same extreme phenotype may help us to identify a disease locus with a less frequent, more highly penetrant variant, which is rare in the general population, but important in understanding the mechanisms behind the disease. A susceptibility locus identified by family-based methods may also have different variants that contribute to more common forms of the disease or the general variation of the trait. Furthermore, mutations identified in even more rare monogenic forms of diseases sharing trait components with common diseases may expose critical new pathways involved in the pathogenesis of common health problems (Peltonen et al. 2006). Therefore, linkage studies provide relevant additional information when hunting for disease susceptibility loci.

2.5.3 Non-parametric linkage analysis

The basic idea behind the linkage analysis is that in families where there is an inherited disease the affected individuals have a higher probability of sharing the part of the genome where the susceptibility loci lies on; in other words, affected individuals will tend to be alike at genetic markers closely linked to a disease locus. The goal in genetic mapping is to find statistically significant relationships between an individual’s phenotype and genetic variables (Figure 6). In complex disease studies NPL (non-parametric linkage) analysis has to be used, because we do not know parameters such as the mode of inheritance (dominant, recessive etc.) needed for parametric linkage analysis. In NPL analysis no disease model is assumed and the question is: “Is the marker inherited with the disease more often than expected by chance?”

If two individuals have the same alleles of the marker they can share these alleles as identical by descent (IBD), which means that the alleles are demonstrably copies of the same ancestral (usually parental) chromosome or the
alleles can be identical by state (IBS), which means that they look the same, and may have the same DNA sequence, but they are not derived from a known common ancestor. Shared segment analysis can be conducted using either IBS or IBD data. IBD is more powerful, but it requires parental data. Multiallele microsatellites have more information than two-allele markers such as SNPs for defining IBD. The linkage analyses in this thesis work have been conducted using IBD methods and microsatellite markers.

Fig. 6. Linkage analysis. In linkage analysis the correlation of the observed phenotype and the observed genotype is tested. However, these “observations” may not always be the exact “truth”. Mistyping due to errors in laboratory work or data handling for example, brings errors to the genotype data, and misdiagnosis due to a complex phenotype and inadequate accuracy of the diagnosis may result in false phenotypes. Both of these possible sources of errors make it harder to detect true linkage. However, the basic idea of linkage analysis is to find a putative trait-locus (disease-locus) that has a genetic effect on the phenotype (disease) using markers that are in linkage or linkage disequilibrium (LD) with it. However, the true phenotype is usually also affected by many other factors, which makes hunting for the disease-locus more difficult. IBD, identical by descent; IBS, identical by state.
2.5.4 Heritability

Heritability allows a comparison of the relative importance of genes and environment to the variation of traits within and across populations. It is important in predicting disease risk in medicine and a parameter that determines statistical power in gene-mapping studies that use pedigree information (Visscher et al. 2008). Heritability is defined as the proportion of phenotypic variance that is explained by genetic factors. The broad-sense heritability, $H^2$, reflects the proportion of all genetic effects ($V_G$) (additive, dominant and epistatic) influencing phenotypic variation ($V_P$):

$$H^2 = \frac{V_G}{V_P}$$

Whereas, the genetic effects are further decomposed when defining the proportion of phenotypic variance attributable to additive genetic effects, the narrow-sense heritability:

$$h^2 = \frac{V_A}{V_P},$$

where the denominator contains the total phenotypic variation ($V_P$), usually excluding variation that is due to known fixed factors and covariates such as sex and age, and the numerator contains variation that is due to additive genetic values ($V_A$) in the population.

Heritability estimate is always specific to the sample where it is estimated. A high heritability means that most of the observed variation in the present population is caused by variation in genotypes. However, it does not mean that the phenotype is determined once we know the genotype. A low heritability, on the other hand, means that of all observed variation, a small proportion is caused by variation in genotypes.

2.5.5 Variance component linkage analysis

Some interesting traits should not and cannot be dichotomized as healthy and affected, and dichotomizing quantitative phenotypes generally leads to loss of information. Quantitative phenotypes can be studied using variance component linkage analysis (VCLA), which is an extension of the strategy developed by Amos (Amos 1994). The idea in VCLA is the same as in the analysis of qualitative traits: The pattern of phenotypic similarity among pedigree members
should be reflected by the pattern of IBD sharing among them at chromosomal loci influencing the trait of interest. The main difference is that in quantitative analysis the trait is generally assumed to follow multivariate normal distribution, whereas in qualitative analysis the trait is dichotomous.

Quantitative genetic models are based on the decomposition of the phenotypic value of an individual to genetic and environmental components. In the simplest model, phenotypic variance can be described as:

$$\sigma_p^2 = \sigma_g^2 + \sigma_e^2,$$

where $\sigma_p^2$, $\sigma_g^2$ and $\sigma_e^2$ are phenotypic, genetic and environmental variance, respectively.

In the basic variance component model $\sigma_p^2$ is modeled as a function of the additive genetic effect $\sigma_q^2$ of the marker location (putative quantitative trait locus, QTL), the aggregate additive genetic effect $\sigma_g^2$ of all other QTL in the genome (polygenes) and the aggregate environmental effects $\sigma_e^2$. Statistical evidence for linkage is then evaluated with a likelihood-ratio test which is typically represented as a logarithm-of-odds (LOD) score.

### 2.5.6 Bivariate linkage analysis

The basic variance components approach has been extended to a multivariate framework. In the multivariate linkage model, the phenotype covariance is further decomposed to include the genetic correlation between traits caused by additive genetic effects and the shared effects of the QTLs. In the bivariate analysis, trait-specific means, variance components relating to major gene effects $\sigma_q^2$, residual additive genetic effects $\sigma_g^2$, random environmental effects $\sigma_e^2$, covariate effects, as well as three associated correlations; correlation caused by a major gene $\rho_q$, correlation caused by residual additive genetic effects $\rho_g$ and correlation caused by random environmental effects $\rho_e$ are estimated simultaneously using maximum-likelihood estimates. To test complete pleiotropy (the same major gene in the chromosomal region of interest affects both traits) or coincident linkage (no shared major gene effects in the chromosomal region of interest on the two traits), likelihood for the linkage model in which $\rho_q$ is estimated is compared with a model in which $\rho_q$ is constrained to 1 or 0, respectively.
2.5.7 Association and haplotype analysis

In addition to GWAS, association analysis is used in candidate gene studies, where a selected amount of SNPs are genotyped on the region of interest and the frequency of the alleles or genotypes is compared between unrelated cases and controls. Differences in the allele/genotype frequencies between cases and controls may indicate that the SNP itself or a variation in LD with the SNP is involved in the development of the studied trait. If the studied variable is quantitative, the differences in the variable are compared between allele or genotype groups. The same idea can be applied to haplotype analysis, with the exception that haplotypes cannot be determined like genotypes, and they have to be estimated. After statistical estimation of haplotypes their estimated frequencies can be compared between cases and controls, and if the variable is quantitative, the differences in the variable can be compared between the different haplotypes.

2.6 Genetics of HDL-cholesterol (HDL-C), adiponectin and VEGF

2.6.1 HDL-C

Low plasma HDL-C level is an important risk factor for atherosclerosis and low plasma HDL-C is the most common dyslipidemia associated with premature and familial CHD (Genest et al. 1992). Genetic factors account for about 50% of the variability in plasma HDL-C levels, and the inheritance of HDL-C levels is complex, influenced by multiple genetic and environmental factors and by interactions between these single factors. The genetic heterogeneity also involves the general link between low plasma HDL-C levels and increased CHD risk, since some genetic variants are associated with low HDL-C levels and decreased CHD risk (e.g. apoA-I (Milano)) (Sirtori et al. 2001) and some with high HDL-C levels and increased CHD risk (e.g. a HL gene variant) (Jansen et al. 1997).

The most common inherited form of low HDL-C is familial hypoalphalipoproteinemia, defined as an HDL-C level below the 10th percentile, without secondary cause, and associated with a family history of low HDL-C levels. Some of the cases are due to mutations in HDL structural genes while others are related to accelerated catabolism of HDL and its apolipoproteins, but the genetic cause is not fully characterized, however. There are also other mutations that occur in genes that play integral parts in HDL-C metabolism and predispose towards extreme HDL-C levels, such as mutations in ABCA1, APOA1,
CETP, HL, LCAT and LPL. The functions of these gene products in HDL-C metabolism are shown in Figure 2. Monogenic disorders due to mutations in some of these genes cause extreme HDL-C levels, but they are very rare (Table 1). There is evidence that both common variants and rare mutations contribute to the inter-individual variation in HDL-C levels (Klos & Kullo 2007).

| Table 1. Rare monogenic disorders causing extreme HDL-C levels |
|-----------------|-----------------|-----------------|
| **Gene**        | **Locus**       | **Disorder**    | **Phenotype and variants** |
| ABCA1           | 9q31.1          | Tangier disease | Low HDL-C, premature CAD (in some people), very large, yellow-orange tonsils, enlarged liver, spleen and lymph nodes, hypocholesterolemia, and abnormal chylomicron remnants, peripheral neuropathy, hemolytic anemia, corneal opacities<sup>1</sup> |
| ANGPTL3         | 1p31.1-p22.3    | Combined hypolipidemia | Extremely low plasma levels of LDL-C, HDL-C and TG. Two nonsense mutations (E129X and S17X) were recessive with respect to HDL-C (Musunuru et al. 2010). |
| APOA1           | 11q23-q24       | ApoA-I deficiency | Low HDL-C levels and decreased CHD risk in ApoA-I Milano, and corneal clouding, CHD, amyloidosis and neuropathy in Detroit and Iowa type<sup>1</sup>. 26 allelic variants reported in OMIM<sup>1</sup>. ApoA-IFin: dominantly inherited hypoalphalipoproteinemia (Miettinen et al. 1997). |
| CETP            | 16q21           | CETP deficiency  | High HDL-C levels |
| HL              | 15q21.3         | HL deficiency    | Abnormally triglyceride-rich HDL and LDL particles and increased risk for CHD |
| LCAT            | 16q22.1         | LCAT deficiency (fish-eye disease and Norum disease) | Low HDL-C levels, corneal opacities, anemia, renal failure<sup>1</sup> LCAT<sub>Fin</sub>: 5% of the cases with HDL-C lower than 0.7 mmol/l (Miettinen et al. 1998) |

Apoe, apolipoprotein E; ABCA1, ATP binding cassette sub-family A, member 1; CETP, cholesteryl ester transfer protein; HL, hepatic lipase; LCAT, lecithin-cholesterol acyltransferase. 1Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD). http://omim.org/.

**Previous findings of genome-wide methods**

“Traditional” genome-wide linkage analyses have been followed by large-scale GWAS, which have provided us with overwhelming amounts of data about the genetics of HDL-C and atherosclerosis. So far, genome-wide scans for the loci regulating HDL-C levels have been published with significant results reported

### 2.6.2 Adiponectin

Adiponectin gene (*ADIPOQ*) is located on chromosome 3q27, and SNPs close to or at *ADIPOQ* have been associated with adiponectin levels. However, also other genes and chromosomal regions have shown evidence of linkage or association with total adiponectin levels in genome-wide studies (Table 2), suggesting that the levels may be influenced by multiple genetic factors. On the other hand, some large GWAS and meta-analyses support the role of polymorphisms at the *ADIPOQ* locus as major predictors of circulating adiponectin levels (Richards et al. 2009, Heid et al. 2010), and in addition, predictors of insulin sensitivity and atherosclerosis (Menzaghi et al. 2007). It is possible that certain *ADIPOQ* variants may determine lower *ADIPOQ* expression, causing in turn an increased risk of developing insulin resistance and CHD. There is also evidence that the genetic architecture of adiponectin overlaps with the genetics of metabolic
syndrome related traits, such as HDL-C and insulin (Butte \textit{et al.} 2005, Comuzzie \textit{et al.} 2007). The genetic heritability of total plasma adiponectin levels has been estimated to be 39–93\% (Comuzzie \textit{et al.} 2001, Lindsay \textit{et al.} 2003, Chuang \textit{et al.} 2004, Pollin \textit{et al.} 2005, Butte \textit{et al.} 2005). Linkage studies or estimations of heritability for the HMW adiponectin had not been published prior to our study.

\textbf{Table 2.} Chromosomal loci showing significant linkage/association to adiponectin in previous genome-wide studies.

<table>
<thead>
<tr>
<th>Chromosomal locus or gene</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>5q21, \textit{FER}</td>
<td>Qi \textit{et al.} 2011</td>
</tr>
<tr>
<td>8q12.1-21.3</td>
<td>Tejero \textit{et al.} 2007</td>
</tr>
<tr>
<td>8p23</td>
<td>Ling \textit{et al.} 2009</td>
</tr>
<tr>
<td>9p</td>
<td>Lindsay \textit{et al.} 2003</td>
</tr>
<tr>
<td>11q23.2-24.2</td>
<td>Tejero \textit{et al.} 2007</td>
</tr>
<tr>
<td>11p</td>
<td>Hicks \textit{et al.} 2007</td>
</tr>
<tr>
<td>14q12-13</td>
<td>Comuzzie \textit{et al.} 2001</td>
</tr>
<tr>
<td>16q23.3, \textit{CDH13}</td>
<td>Jee \textit{et al.} 2010</td>
</tr>
</tbody>
</table>

\textbf{2.6.3} \textit{VEGF}

The gene encoding VEGF is located on chromosome 6p12 and comprises a 14-kb coding region with 8 exons and 7 introns (Brogan \textit{et al.} 1999). Polymorphisms of the \textit{VEGF} gene have been associated with CHD severity (Howell \textit{et al.} 2005, Biselli \textit{et al.} 2008, Guerzoni \textit{et al.} 2009) and plasma levels of VEGF. A promoter polymorphism of the \textit{VEGF} gene -2578A>C (rs699947, at position -2578 from the first translated nucleotide, changing A to C) has been associated with VEGF production (Shahbazi \textit{et al.} 2002) and with the degree of coronary atherosclerosis quantified angiographically, but only in subjects without previous AMI (Howell \textit{et al.} 2005). The C allele has been associated with higher VEGF production (Shahbazi \textit{et al.} 2002) and the CC genotype with a milder degree of coronary atherosclerosis (Howell \textit{et al.} 2005). The CC genotype of the -634C>G SNP (rs2010963) has been associated with higher VEGF serum levels (Awata \textit{et al.} 2002) and the T allele of the +936C>T SNP (rs3025039) has been linked with lower VEGF levels (Renner \textit{et al.} 2000, Krippl \textit{et al.} 2003).
3 Aims of the study

The aim of the study was to investigate the genetic background of atherosclerosis and, in more detail, the genetics of HDL-C. The following specific aims were addressed:

1. To identify loci that regulate HDL-C levels and predispose to CHD by performing a genome-wide scan in families with low HDL-C levels and premature CHD.
2. To study the relationship between conventional CHD risk factors and the adiponectins (total and HMW adiponectin), to investigate genetic and environmental determinants, heritability and genetic regulation of the adiponectins, and to find out whether the adiponectins have a common genetic background with HDL-C.
3. To investigate whether polymorphisms of the VEGF gene, previously associated with VEGF production, are pro- or anti-atherosclerotic and whether they affect intima-media thickness and the risk of AMI.
4 Subjects and methods

4.1 Study subjects

All the studies were approved by the Ethics Committee of Oulu University Hospital, and all the subjects provided informed consent.

4.1.1 Northern Finnish families with early onset CHD and low HDL-C (I, II)

Probands (men and women) with premature CHD (i.e. AMI, a coronary artery bypass graft operation or percutaneous transluminal coronary angioplasty before the age of 55 years) and low HDL-C (<1.1 mmol/l) were selected by screening the hospital records of all the young CHD-patients in Oulu University Hospital during the years 1990–1996. In addition, the probands were required to have normal to moderately elevated levels of TG (<3.5 mmol/l) and total cholesterol (<7 mmol/l), no diabetes and an entry in the hospital records indicating a family history of CHD.

Altogether 35 extended families were recruited from Northern Finland in the first stage. There were on average three generations (minimum 2, maximum 5) and the average pedigrees size was 19 subjects (minimum 5, maximum 89). Of the 644 family members 375 were genotyped. The age of the genotyped subjects used in the analyses ranged from 16 to 88 years.

Blood samples were obtained from each subject and lipid measurements were performed in the case of the CHD patients before or at least three months after AMI or coronary bypass operation. Information about medication, past medical history and smoking was elicited using a questionnaire.

Furthermore, in the later stage of the project (1997–2000), 4 additional families with 27 genotyped subjects were collected from the same geographical area using the same criteria and added to the further analysis stages (finemapping, adiponectin measurements).

4.1.2 OPERA controls (III)

Subjects from the OPERA (the Oulu Project Elucidating Risk of Atherosclerosis) study whose carotid IMT had been measured (259 males and 267 females) were
used in this study as controls for AMI survivors. The subjects were from the city of Oulu, 40–61 years old and had been selected by the Social Insurance Institution as control subjects for hypertensive patients, as described previously (Kauma et al. 1996). All of the subjects made an outpatient visit to the research unit for laboratory tests, a physical examination and interviews concerning their alcohol consumption, smoking and exercise habits, medication and past medical history.

4.1.3 Acute myocardial infarction (AMI) survivors (III)

A multiple risk factor analysis trial (MRFAT) was started in 1996 at the Division of Cardiology, University of Oulu, to determine the prognostic power of several non-invasive risk markers of mortality in patients who survived an AMI. The patients were recruited to participate within the first 7 days after the diagnosis of AMI. The exclusion criteria were death during the hospital stay, unstable angina at recruitment, dementia, alcoholism, drug abuse or any other condition that could impair the patient’s capacity to provide informed consent. The qualifying diagnostic criteria of the patients have been described in detail earlier (Tapanainen et al. 2001). The group of AMI survivors selected for this study comprised 251 AMI patients younger than 65 years and originating from the same geographical region as the (OPERA) control group.

4.2 Methods

The methods used in the studies are shown in Table 3.

4.2.1 Selection and genotyping of the polymorphisms (I, II, III)

The ABI Prism Linkage Mapping Set 1 (Applied Biosystems, Foster City, CA, USA) was used for genotyping microsatellite markers in the initial screening. The set consisted of 358 markers covering the whole genome (except the Y chromosome) at about 10cM resolution. The markers were originally selected from the 1996 Genethon map (Dib et al. 1996) but additional markers on chromosomes 2, 4, 6, 10, 15 and 22 were selected for the regions showing the highest evidence of linkage. Genotyping was done by ABI 377 and ABI 3100 automatic sequencers as recommended by the manufacturer (Applied Biosystems) and the sizes of the alleles were determined using the Genescan 3.1, Genotyper 2.0 and GeneMapper 3.5 programs.
Table 3. Methods and software used in the publications (in alphabetical order).

<table>
<thead>
<tr>
<th>Methods and software used</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory procedures and measurements</td>
<td></td>
</tr>
<tr>
<td>Carotid ultrasound examinations and measurements</td>
<td>III</td>
</tr>
<tr>
<td>DNA extraction (by salting-out method)</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Lipid and lipoprotein measurements</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Measurement of total and HMW-adiponectin</td>
<td>II</td>
</tr>
<tr>
<td>PCR (polymerase chain reaction)</td>
<td>I, II</td>
</tr>
<tr>
<td>TaqMan SNP genotyping</td>
<td>III</td>
</tr>
<tr>
<td>Analysis and data-handling software</td>
<td></td>
</tr>
<tr>
<td>Downfreq</td>
<td>I, II</td>
</tr>
<tr>
<td>Genescan 3.1, Genotype 2.0 and GeneMapper 3.5</td>
<td>I, II</td>
</tr>
<tr>
<td>GRR (Graphical Representation of Relationships)</td>
<td>I, II</td>
</tr>
<tr>
<td>Haplopainter</td>
<td>I, II</td>
</tr>
<tr>
<td>Haploview</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Linkagepar</td>
<td>I, II</td>
</tr>
<tr>
<td>Makedata</td>
<td>I, II</td>
</tr>
<tr>
<td>Makeped</td>
<td>I, II</td>
</tr>
<tr>
<td>Mega2</td>
<td>I, II</td>
</tr>
<tr>
<td>Mendel</td>
<td>I</td>
</tr>
<tr>
<td>Merlin</td>
<td>I</td>
</tr>
<tr>
<td>PAWE</td>
<td>III</td>
</tr>
<tr>
<td>PedCheck</td>
<td>I, II</td>
</tr>
<tr>
<td>PedStats</td>
<td>I, II</td>
</tr>
<tr>
<td>SimWalk2</td>
<td>I</td>
</tr>
<tr>
<td>SOLAR bivariate analysis</td>
<td>II</td>
</tr>
<tr>
<td>SOLAR univariate analysis</td>
<td>I, II</td>
</tr>
<tr>
<td>SPSS</td>
<td>I, II, III</td>
</tr>
<tr>
<td>THESIAS: testing haplotype effects in association studies</td>
<td>III</td>
</tr>
</tbody>
</table>

Haploblock structures of the functional and positional candidate genes, peroxisome proliferator-activated receptor delta (PPARD) and retinoid X receptor, beta (RXRB), were constructed using the data from the HapMap project (HapMap LD Map, CEU population, 2006) and Haplovieview software. This information was then utilized to select SNPs (minor allele frequency at least 5 percent) that were representative of most of the variation and the haploblocks of the genes, and available as TaqMan® SNP Genotyping Assays. SNPs -2578A>C, -634C>G and +936C>T of the VEGF gene (rs699947, rs2010963 and rs3025039, respectively) were selected based on the previously reported associations with serum levels of VEGF and VEGF production. The SNPs were located on the promoter region (at
position -2578 from the first translated nucleotide), 5’untranslated region (5’UTR) (at position -634 from the first translated nucleotide) and 3’UTR (at position +936 from the first translated nucleotide). Genotypings were done using TaqMan® SNP Genotyping Assays and the Applied Biosystems 7000 Real-Time PCR (polymerase chain reaction) System.

4.2.2 Laboratory methods

Lipid and lipoprotein measurements (I, II)

Blood samples were obtained after an overnight fast. Venous blood was drawn into EDTA tubes. Plasma was separated by centrifugation at 2000 rpm for 10 min and kept at 4°C until further lipoprotein analysis. The very low-density lipoprotein (VLDL) fraction (d < 1.006 g/ml) was separated from plasma by ultracentrifugation in a Kontron TFT 45.6 rotor at 105,000 xg for 18 h. The plasma HDL-C concentration was determined by mixing 1 ml of the VLDL-free fraction with 25 µl of 2.8% (weight/volume) heparin and 25 µl of 2M MnCl₂ and by measuring the cholesterol concentration in the supernatant after centrifugation at 1000 xg and 4°C for 30 min. The plasma LDL-C concentration was then calculated by subtracting the cholesterol concentration in HDL from that in the VLDL-free fraction. The concentrations of cholesterol and TG were determined in the plasma and lipoprotein fractions by enzymatic colorimetric methods (kits of Boehringer Diagnostica, Mannheim GmbH, Germany, catalogue nos. 236691 and 701912, respectively) using a Kone Specific analyzer (Kone Instruments, Espoo, Finland) (Kakko et al. 1998).

Measurement of total and HMW-adiponectin concentration (II)

Total adiponectin and HMW adiponectin were measured using Human adiponectin ELISA kit (Cat # EZHAPD-61K) and Human HMW adiponectin ELISA (Cat # EZHMWA-64K) supplied by Linco Research Inc, Missouri, US. The assays were performed strictly according to manufacturer instructions. The intra- and inter-assay variations for total adiponectin assay were <7.4 and <8.4 %, and for HMW adiponectin <3.4% and <9.1%, respectively. The HMW adiponectin assay specifically measures only the HMW form of adiponectin and does not recognize hexameric or trimeric adiponectin (Sinha et al. 2007).
4.2.3 Measurement of intima-media thickness (IMT) (III)

The carotid ultrasound examinations and measurements were performed by a trained radiologist as described previously (Kauma et al. 1996). IMT was measured by B-mode ultrasonography from the far wall and defined as the distance between the media-adventitia interface and the lumen-intima interface. The thickest point of the IMT was measured at five points on each side on the far wall (internal carotid artery (ICA), bifurcation enlargement (BIF) and proximal, middle and distal locations of the common carotid artery (CCA)), i.e. at a total of 10 sites, though avoiding sites with plaque. In the case of an atheromatous plaque, a combined IMT plaque thickness (CombIMT) was measured and if there were no plaques at the site of the measurement, the IMT value alone was used as CombIMT. The overall IMT values (later IMT without plaques and CombIMT) were defined as the mean of the measurements in CCA, BIF and ICA.

4.2.4 Statistical methods

Linkage analysis (I, II)

Knowledge of true family relationships is crucial in genetic mapping since all genetic analyses are performed conditional on the family structure. Therefore, prior to the linkage analysis, the familial relationships were verified using GRR (Graphical Representation of Relationships) (Abecasis et al. 2001) which visually helps to detect potential errors. GRR plots the mean and standard deviation of genome-wide IBS sharing between all pairs of individuals in a given sample, and different relationships (sibling pairs, half-sibs, parent-offspring and unrelated pairs) are shown in different colors and form distinct clusters when relationships are correct. If suspicious relationships were detected, the original patient questionnaires were re-checked and relationship was confirmed from the Population Register Center if needed. Mendelian errors were checked using the PedCheck program (O’Connell & Weeks 1998). If errors were observed, the genotypings were double-checked from the original data, the microsatellites were re-genotyped or the erroneous data was deleted.

Qualitative linkage analysis was performed (original article I) using SimWalk2 (version 2.83) (Sobel & Lange 1996) and MERLIN (version 0.9.12b) (Abecasis et al. 2002) software. The subjects with their measured HDL-C levels in the lowest 10th percentile of the sex-specific population HDL-C levels were
coded as being affected in the qualitative linkage analysis, others were coded as unknown. The limit was 0.864 mmol/l for men and 1.08 mmol/l for women. The allele frequencies for each marker were estimated from all individuals using the Downfreq program (Terwilliger & Göring 2000). Mega2 (Mukhopadhyay et al. 2005) was used to construct all the input files for the SimWalk2 and MERLIN programs. An exact analysis using the Lander-Green algorithm was performed by MERLIN on the pedigrees of small or intermediate size (9 pedigrees) and the more complex pedigrees (10 pedigrees) were analyzed by SimWalk2 using Markov chain Monte Carlo and simulated annealing algorithms. SimWalk2 was used to combine the pre-computed scores for the smaller pedigrees with the estimates obtained for the large pedigrees and then to compute empirical p-values for both individual pedigrees and the overall dataset.

Quantitative univariate (I, II) and bivariate (II) linkage analyses were performed by means of SOLAR software (Almasy & Blangero 1998) using a variance component method, which is based on specifying the expected genetic covariances between pairs of relatives as a function of their IBD at a marker linked to a QTL (Amos 1994). The total observed phenotypic variance was split into components attributable to QTL, residual polygenic effects and non-genetic effects, and the presence of a putative QTL was tested by means of a likelihood ratio statistic (LOD score). Plasma HDL-C levels (I) and adiponectin levels (total and HMW adiponectin) (II) were used as a continuous variable in the analysis, and the subjects’ sex, age (I) and BMI (only when stated), being statistically significant covariates in the dataset (p<0.001), were also considered in the analysis. Ascertainment correction by conditioning for the probands was performed and parameters with skewed distributions were normalized using log transformation. To deal with kurtosis, LOD adjustment was performed by using a simulation to build up the distribution of LOD scores that one could expect to observe under the null hypothesis of no linkage. This consisted of 10 000 trials, in each of which a fully informative marker completely unlinked to the trait was simulated and trait linkage was then tested at that simulated marker. The LOD adjustment regressed the observed LOD scores against the LOD scores expected for a multivariate normal trait. The inverse of the slope of the regression line was the LOD adjustment. Empirical P-values for the LOD scores in the sample were also determined using the pedigree data by simulation of 10,000 replicates of a fully informative marker completely unlinked to the trait. In order to determine the level of suggestive evidence of linkage, we compared the unadjusted empirical p-values from the simulations to the point-wise levels calculated using
the equation presented in the article of Lander and Kruglyak (Lander & Kruglyak 1995). The average crossover rate used in the equation, given the types of relatives in our pedigrees, was 2.94.

As a post hoc analysis (I), to control for the possibly confounding effect of statin use, all the chromosomes with LOD scores > 1.5 were re-analyzed after subtracting a constant (6% of the sex-specific mean) from the HDL-C values of all the statin users (also performed for the qualitative analysis). The daily doses of statin taken by our subjects were small and their effect on plasma HDL-C levels was estimated to be only modest, with less than a 6% increase (Jones et al. 2003).

In order to investigate the shared genetic contributions between individual traits, a bivariate quantitative genetic analysis (II) (Almasy et al. 1997), an extension of the univariate approaches, was performed for the quantitative phenotypes measured (HDL-C, LDL-C, TG, VLDL-C, VLDL-TG, BMI, total adiponectin, HMW adiponectin and HMW adiponectin/total adiponectin ratio) and adjusted for sex (and also for BMI or statin use only when stated). Genetic and environmental correlations were estimated between pairs of the studied phenotypes, based on maximum likelihood ratio and variance component decomposition. Genome-wide bivariate linkage analysis was performed for HDL-C and total or HMW adiponectin. In addition, heritabilities of the studied quantitative traits were estimated.

**Associations, correlations and haplotype analysis (I, II, III)**

SPSS-package (Statistical Package for Social Studies version 16.0, SPSS Inc., Chicago, IL, USA) was used to study associations and correlations between studied traits (I, II and III). Differences in the basic characteristics between the cases and controls or different study samples (I, II and III) were tested using Student’s T-test (continuous variables) and chi-square test (categorical variables). Differences in adiponectin levels between men and women and CHD-patients and healthy subjects were tested using Student’s T-test. Partial correlations were calculated between the measured variables (the adiponectins and conventional atherosclerosis risk factors) and controlled for sex, and in addition, only when indicated, for BMI or waist diameter (II). Correlations of lipid and lipoprotein parameters were controlled also for statin use. Parameters with skewed distributions were normalized using log transformation. To control for the relatedness of the study subjects, the partial correlations of the adiponectins and
the conventional atherosclerosis risk factors were analyzed using probands and their spouses.

SPSS was also used when the differences in IMT between the different genotype groups were evaluated by ANOVA (with Bonferroni’s correction) (III). Logarithmic transformation was used to normalize IMT (without plaques) distribution, but non-parametric Kruskal-Wallis test was required for CombiIMT (not normally distributed after logarithmic transformation). The differences in the genotype and allele frequencies between the controls and AMI survivors were analyzed by chi-square test (III). The general factorial procedure of the General Linear Model (GLM) with Type III sums of squares was used to construct a model to explain the variation in carotid IMT (III). The genotypes of the VEGF polymorphism were added as an independent factor in the model without any assumption of the mode of inheritance (i.e. without any combination of genotypes) and covariates (BMI, pack-years of smoking, age, systolic blood pressure, LDL cholesterol and HDL cholesterol) were selected based on their proposed roles in atherosclerosis. The eta-squared statistic was used to estimate how much of the variation each factor or covariate explained. The PAWE (Power for Association with Errors) (Gordon et al. 2002) software was used for the power analysis (III).

Mendel software was used to test family-based quantitative association. QTL association was tested using variance component models, treating genotypes at the marker locus as predictors modifying the mean for a quantitative trait (QTL association, model 1). This “measured genotype” approach controls for random environment and polygenic backgrounds while interpreting the probabilities using maximum likelihood estimation and likelihood ratio tests (Lange et al. 2001).

Haplotype analysis was performed using THESIAS software (Tregouet & Garelle 2007). THESIAS allows one to simultaneously estimate haplotype frequencies and their associated effects on the phenotype of interest. In addition to qualitative (AMI vs. control) also quantitative trait (IMT) was studied and covariate-adjusted (age, BMI, systolic blood pressure, LDL-C, pack-years of smoking, only when indicated) haplotype effects were investigated.
5 Results

5.1 Genetic loci regulating HDL-C levels in Northern Finnish families with early onset CHD and low HDL-C levels (I)

5.1.1 Basic characteristics, heritability and significance thresholds

CHD was diagnosed in 42% of the men and 12% of the women in the whole sample, the average age at the first CHD event being 49 and 53 years, respectively. This relatively young age at CHD onset emphasizes the potentially strong genetic component in CHD in our pedigrees. The average plasma HDL-C levels were 1.07 mmol/l for men and 1.39 mmol/l for women, and the CHD patients had lower plasma HDL-C levels than the healthy subjects among both the men (0.96 vs. 1.15 mmol/l, p<0.001) and the women (1.27 vs. 1.43 mmol/l, p=0.06) (Table 4). CHD-status and HDL-C values are shown as examples for four pedigrees in Figure 7.

The additive genetic heritability of HDL-C was 43%. Based on the empirical p-value estimate, the empirical p-value for LOD score 3.1 (our highest finding on 4p12) was 0.0004 (unadjusted) and the threshold for suggestive evidence of linkage was LOD score 2.0, p=0.0011.

5.1.2 Quantitative and qualitative linkage analysis

VCLA revealed four suggestive QTLs for HDL-C levels, with the highest LOD score, 3.1, at the chromosomal locus 4p12. Other suggestive LOD scores were 2.1 at 2q33, 2.1 at 6p24 and 2.0 at 17q25 (Figure 8). Inclusion of BMI did not alter the results, except in the case of chromosome 6, where the two-point LOD scores increased to 2.7 on 6p24 and a two-point LOD score of 2.1 was found on 6p22.

The first stage of the qualitative linkage analysis (Figure 9) revealed one suggestive locus with a multipoint NPL score of 2.6 at 10p15.3. For the chromosomes (6, 10, 22) with NPL scores >1.5, we selected nine new microsatellite markers, and four new families with 27 genotyped persons were added to the analysis. Thus, genotyping of the additional markers and families resulted in three suggestive loci for the qualitative low HDL-C trait, with a multipoint NPL score of 2.6 at the chromosomal locus 10p15.3, 2.5 at 22q11 and 2.1 at 6p12. When only overweight subjects (BMI>25) were included in the
analysis, a multipoint NPL score of 1.8 was found on 6p22. After correction for statin use, the strongest evidence of linkage was shown on chromosomes 4p12, 6p24, 6p12, 15q22 and 22q11. In conclusion, we identified seven chromosomal regions for HDL-C regulation exceeding the level for suggestive evidence of linkage.

Fig. 7. Pedigrees of four studied families illustrating the manifestation of CHD (indicated with *) and low HDL-C phenotype (indicated with black coloring), which was defined as the lowest 10th percentile of the sex-specific population HDL-C levels (under 0.864 mmol/l and 1.08 mmol/l for men and women, respectively). Some of the studied families did not show co-inheritance of the low HDL-C phenotype and CHD (families A and C), whereas in other families these phenotypes seemed to correlate more strongly (B). In some of the families the young age of some family members possibly prevented observation of the potential underlying co-inheritance (D).
Table 4. Characteristics of the genotyped subjects used in the quantitative and qualitative linkage analyses for HDL-C regulation (original publication I), in the adiponectin study (II), and in the study of the VEGF gene polymorphisms (III).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Quantitative analysis¹</th>
<th>Qualitative analysis²</th>
<th>Adiponectin study³</th>
<th>OPERA⁴</th>
<th>AMI°</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
</tr>
<tr>
<td>Number of subjects</td>
<td>212</td>
<td>163</td>
<td>137</td>
<td>104</td>
<td>223</td>
</tr>
<tr>
<td>Age</td>
<td>48.6 (13.1)</td>
<td>49.3 (16.5)</td>
<td>48.0 (13.9)</td>
<td>49.9 (16.7)</td>
<td>50.9*</td>
</tr>
<tr>
<td>CHD patients</td>
<td>90 (42%)</td>
<td>19 (12%)</td>
<td>60 (44%)</td>
<td>12 (12%)</td>
<td>96 (43%)</td>
</tr>
<tr>
<td>Age at first CHD event</td>
<td>48.8 (13.1)</td>
<td>52.9 (16.5)</td>
<td>49.5 (13.9)</td>
<td>54.5 (16.7)</td>
<td>50.9*</td>
</tr>
<tr>
<td>Low HDL-C patients</td>
<td>55 (26%)</td>
<td>29 (18%)</td>
<td>49 (36%)</td>
<td>24 (23%)</td>
<td>n.d.</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>26.9 (3.9)</td>
<td>25.5 (4.7)</td>
<td>26.5 (5.1)</td>
<td>25.5 (4.6)</td>
<td>26.9*</td>
</tr>
<tr>
<td>TC (mmol/l)</td>
<td>5.37 (0.87)</td>
<td>5.49 (0.95)</td>
<td>5.44 (1.07)</td>
<td>5.46 (1.08)</td>
<td>5.38</td>
</tr>
<tr>
<td>HDL-C (mmol/l)</td>
<td>1.07 (0.29)</td>
<td>1.39 (0.33)</td>
<td>1.00 (0.35)</td>
<td>1.35 (0.35)</td>
<td>1.06</td>
</tr>
<tr>
<td>LDL-C (mmol/l)</td>
<td>3.33 (0.87)</td>
<td>3.30 (0.95)</td>
<td>3.38 (1.07)</td>
<td>3.36 (1.08)</td>
<td>3.33</td>
</tr>
<tr>
<td>Triglycerides (mmol/l)</td>
<td>1.78 (0.97)</td>
<td>1.49 (0.95)</td>
<td>1.90 (1.07)</td>
<td>1.55 (1.08)</td>
<td>1.81</td>
</tr>
<tr>
<td>Current smokers</td>
<td>49 (23%)</td>
<td>23 (14%)</td>
<td>34 (25%)</td>
<td>12 (12%)</td>
<td>49 (22%)</td>
</tr>
<tr>
<td>Lipid-lowering drug in use</td>
<td>34 (16%)</td>
<td>6 (4%)</td>
<td>21 (15%)</td>
<td>3 (3%)</td>
<td>38 (17%)</td>
</tr>
<tr>
<td>IMT without plaques (mm)</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>0.88</td>
</tr>
<tr>
<td>CombiIMT (mm)</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Values shown for the subjects genotyped are expressed as means (standard deviations) or numbers of subjects (percentages) except for IMT values, which are expressed as medians (interquartile range).

¹Includes characteristics of all the subjects genotyped for the original variance component linkage analysis (original publication I). ²Includes only the samples used in the qualitative analysis of the low
HDL-C-trait (original publication I).  3Includes the samples used in the adiponectin analyses of the original publication II.  4Includes the control samples used in the association analyses of the VEGF gene polymorphisms (original publication III)  5Includes the case samples used in the association analyses of the VEGF gene polymorphisms (original publication III)  6Subjects having their measured HDL-C levels in the lowest 10th percentile for the general population were coded as affected in the qualitative linkage analysis.  7n=214 (men), n=164 (women). * p<0.05 for the difference between male or female AMI survivors and OPERA cohort subjects, or qualitative and quantitative analysis samples (HDL-C difference in men). CHD, coronary heart disease; CombIMT, combined IMT and plaque thickness, BMI, Body mass index; HDL-C, high-density lipoprotein cholesterol; HMW, high-molecular weight; IMT, intima-media thickness; LDL-C low-density lipoprotein cholesterol, TC, total cholesterol; n.d.; not determined.
Fig. 8. Results of the genome-wide quantitative analysis. All genotyped markers were included in the analysis. The x-axis indicates the distance (cM) from the p-terminus and the y-axis indicates the LOD score. Chr, chromosome (I, published by permission of Nature Publishing Group.).
Fig. 9. Results of the genome-wide qualitative analysis (first stage). The x-axis indicates the distance (cM) from the p-terminus and the y-axis indicates the NPL score. Each multipoint NPL score statistic measures the degree of clustering of the founder.
alleles among the affected cases. BLOCKS is most powerful at detecting linkage to a recessive trait, MAX-TREE is most powerful at detecting linkage to a dominant trait, ENTROPY is a measure of the entropy of the alleles among the affected cases, NPL_ALL is a measure of whether a few founder alleles are over-represented in the affected cases, and NPLPAIR and NPL_ALL are most powerful at detecting linkage to an additive trait and also the two most commonly used statistics incorporated in the most widely used software packages. Chr, chromosome (I, published by permission of Nature Publishing Group.).

5.1.3 Association analysis

The only chromosome showing suggestive evidence of linkage in both, quantitative and qualitative analyses was chromosome 6. In order to search for the underlying gene on chromosome 6, we investigated two functional and positional candidate genes, PPARD and retinoid x receptor beta, RXRB. The highest evidence of association was obtained with marker D6S1713 on 6p25, which showed suggestive evidence of association in quantitative association analysis (p=0.03), but no significant evidence of association was found.

5.2 Genetic and environmental determinants of total and high-molecular weight adiponectin in Northern Finnish families with early onset CHD and low HDL-C levels (II)

5.2.1 Sex difference and association with CHD

Men had lower total adiponectin levels (6.34µg/ml vs. 10.10µg/ml, p<0.001), HMW adiponectin levels (1.93µg/ml vs. 3.82µg/ml, p<0.001) and HMW to total adiponectin ratio (HMW/total ratio) (0.29 vs. 0.38, p<0.001) than women (Table 4). Male CHD patients had significantly lower total (5.50µg/ml vs. 7.12 µg/ml, p<0.001) and HMW (1.69µg/ml vs. 2.13µg/ml, p=0.004) adiponectin levels than men without CHD. The difference between female CHD patients and females without CHD was significant only for HMW adiponectin (total 8.65µg/ml vs. 10.24µg/ml, p=0.16 and HMW adiponectin 2.67µg/ml vs. 3.90µg/ml, p=0.04).
5.2.2 Correlations of the adiponectins and lipids

The genetic and environmental cross-correlations were strong between all the other adiponectin measures, but there was no genetic correlation between the total adiponectin level and the HMW/total ratio (Table 5). HDL-C showed stronger correlation to adiponectins (total 0.57, p=0.001, HMW 0.51, p<0.005) than the other lipids in partial correlation analysis (adjusted for sex, age, BMI and statin use) in probands and their spouses. In the bivariate analysis using all the study samples, the strongest environmental cross-correlation coefficient between adiponectins and lipids was also shown between HDL-C and total adiponectin (0.64, p<0.0001), whereas the strongest genetic correlation (even after adjustment for BMI or statin use) was detected between HMW adiponectin and LDL-C (-0.48, p=0.02) (Table 5). In conclusion, HDL-C failed to show significant genetic correlation with the adiponectins.

Table 5. Genetic and environmental cross-correlations.

<table>
<thead>
<tr>
<th>Genetic correlations</th>
<th>Environmental correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total adiponectin</td>
</tr>
<tr>
<td>Total adiponectin</td>
<td>0.94*</td>
</tr>
<tr>
<td>HMW adiponectin</td>
<td>0.63*</td>
</tr>
<tr>
<td>Adiponectin ratio</td>
<td>0.01</td>
</tr>
<tr>
<td>BMI</td>
<td>-0.06</td>
</tr>
<tr>
<td>HDL-C</td>
<td>-0.01</td>
</tr>
<tr>
<td>LDL-C</td>
<td>-0.03</td>
</tr>
<tr>
<td>VLDL-C</td>
<td>0.28</td>
</tr>
<tr>
<td>VLDL-TG</td>
<td>0.19</td>
</tr>
<tr>
<td>Total TG</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Genetic correlations are displayed below the diagonal and environmental correlations above the diagonal. Correlations are adjusted for sex. *Significant correlations with p-value <0.05 are in italics. Adiponectin ratio, HMW/total adiponectin; BMI, body mass index; HDL-C, high-density lipoprotein cholesterol; HMW, high molecular weight; LDL-C, low-density lipoprotein cholesterol; TG, triglycerides; VLDL-C, very-low-density lipoprotein cholesterol. (II, published by permission of Elsevier).

5.2.3 Heritability and linkage analyses

Our estimates show strong heritability for total (0.53), HMW (0.51) and HMW/total adiponectin ratio (0.68). The highest evidence of linkage in the univariate analyses using only sex as a covariate was shown for total adiponectin.
on chromosome 11p15 with a LOD score of 2.1, and for HMW adiponectin on
chromosomes 6p21 with a LOD score of 2.2. After adjustment for BMI, in
addition to sex in the univariate linkage analysis, the highest evidence of linkage
was shown in regions 11q24.2 with a LOD score of 1.8 (total adiponectin) and in
3q13.2-q24 with a LOD score of 2.0 (HMW adiponectin) (Figure 10a). The
suggestive linkage signals discovered in the bivariate analysis (Figure 10b) were
mostly due to either HMW adiponectin or HDL-C, so the underlying QTLs fail to
explain the variation of both traits; furthermore, the analysis shows no significant
pleiotropy between the traits.

Fig. 10. Results of the genome-wide (a) univariate linkage analyses for loci regulating
total and HMW adiponectin levels and results of the (b) bivariate linkage analyses for
total or HMW adiponectin and HDL-C. The X-axis indicates the distance (cM) from the
p-terminus of the chromosome and the Y-axis indicates the LOD score. The solid line
indicates the results of the multipoint analysis and the filled diamonds the results of
two-point analysis using only sex as a covariate. The dashed line represents the
results of the multipoint analysis and the circles the results of two-point analysis
using sex and BMI as covariates. Chr, chromosome; HMW, high-molecular weight;
HDL-C, high-density lipoprotein cholesterol (II, published by permission of Elsevier).
5.3 VEGF gene in atherosclerosis, quantified as carotid IMT and the risk of AMI (III)

Basic characteristics of the study samples (OPERA cohort and AMI survivors) are shown in Table 4.

5.3.1 Power analysis

Assuming a power of 90% for the study, p-value of 0.05 for statistical significance and the standard deviation of the IMT as 0.21 for the men and as 0.12 for the women (taken from the sample), the difference of IMT between the homozygous AA and CC genotypes should be 0.14 mm for the men (assuming 48 subjects in both groups, taken from the number of samples in the rare allele homozygous group for the polymorphism -2578) and 0.07 mm for the women (assuming 53 subjects in both groups) to be significant. Assuming a study sample of 251 AMI survivors and 515 control subjects (with the rare allele frequency 0.434 for the -2578 polymorphism) and the power of an allelic test as 80%, the difference between allele frequencies should be 0.075 (i.e. the rare allele frequency of 0.359 in AMI survivors) to be detected as statistically significant (p=0.05).

5.3.2 The effect of VEGF variation on IMT and on the risk of AMI

None of the single genotyped polymorphisms was significantly associated with overall IMT (without plaques) or with the risk of AMI. However, men with the 936 CT genotype (the +936C>T polymorphism) exhibited higher IMT values than men with CC genotype (0.87 mm vs. 0.78 mm, p=0.01) on ICA (IMT without plaques) and overall CombIMT (1.15 mm vs. 0.98 mm, p=0.02) (Table 6). In GLM modeling, the genotypes did not have any significant effect on IMT (neither overall IMT without plaques nor overall CombIMT) in men or women. There were no significant differences in the risk of AMI between the genotype or allele groups, when the whole study groups were compared or when only male controls without CHD were compared with male AMI survivors. The result was the same for women. Age- and sex-matched groups also failed to show any significant difference, even when they were analyzed separately for current smokers and non-smokers.
Table 6. Genotype CT for VEGF polymorphism +936C>T is associated with higher IMT (mm) in men.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>N</th>
<th>IMT without plaques (mm)</th>
<th>CombIMT (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Overall</td>
<td>ICA</td>
</tr>
<tr>
<td>CC</td>
<td>177</td>
<td>0.87</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.75–1.01)</td>
<td>(0.65–0.90)*</td>
</tr>
<tr>
<td>CT</td>
<td>67</td>
<td>0.93</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.78–1.10)</td>
<td>(0.65–0.95)*</td>
</tr>
<tr>
<td>TT</td>
<td>9</td>
<td>0.78</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.73–0.93)</td>
<td>(0.65–0.75)</td>
</tr>
</tbody>
</table>

Values are expressed as median (25th percentile - 75th percentile). Genotype, genotype for polymorphism +936C>T; N, number of subjects; IMT, intima-media thickness; CombIMT, combined IMT and plaque thickness; Overall, mean of the different segments; ICA, internal carotid artery; BIF, carotid bifurcation; CCA, common carotid artery. *p<0.05 for the difference between the genotype groups by ANOVA with Bonferroni’s correction for the analysis of IMT without plaques and by non-parametric Kruskal-Wallis test for the analysis of CombIMT.
5.3.3 The effect of VEGF haplotypes on IMT and on the risk of AMI

In men, the CCT haplotype (-2578C/-634C/+936C) was significantly associated with higher CombiIMT (haplotypic effect +0.25 mm, p=0.02) on ICA and with higher IMT without plaques (haplotypic effect of +0.11 mm, p=0.001) on ICA, which remained significant after inclusion of covariates (age, BMI, systolic blood pressure, LDL-C, pack-years of smoking). The number of men with haplotype CCT was, however very small (n=4). In women the CCC haplotype was significantly associated with higher IMT on overall IMT without plaques (haplotypic effect of +0.03 mm, p=0.01), IMT without plaques on ICA (+0.03 mm, p=0.02) and CombiIMT on CCA (+0.06 mm, p=0.001) (Table 7). These haplotypic effects remained significant after inclusion of the covariates (age, BMI, systolic blood pressure, LDL-C, pack-years of smoking) into the model. In men, haplotype CCC was not significantly associated with IMT, but the trend was the opposite of that in women.
Table 7. The VEGF haplotypes with statistically significant effects on IMT.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>N (%)</th>
<th>IMT without plaques</th>
<th>CombiMT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Overall</td>
<td>ICA</td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCT</td>
<td>4 (1)</td>
<td>0.053</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p=0.163</td>
<td>p=0.001*</td>
</tr>
<tr>
<td>CCC</td>
<td>61 (12)</td>
<td>-0.058</td>
<td>-0.056</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p=0.078</td>
<td>p=0.133</td>
</tr>
<tr>
<td>AGC ref.</td>
<td>253 (50)</td>
<td>0.470</td>
<td>0.402</td>
</tr>
<tr>
<td>Women</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCT</td>
<td>34 (6)</td>
<td>0.0004</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p=0.991</td>
<td>p=0.872</td>
</tr>
<tr>
<td>CCC</td>
<td>94 (18)</td>
<td>0.034</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p=0.014*</td>
<td>p=0.022*</td>
</tr>
<tr>
<td>AGC ref.</td>
<td>243 (46)</td>
<td>0.403</td>
<td>0.357</td>
</tr>
</tbody>
</table>

The table shows only the haplotypes that showed significant results (p<0.05) in either men or women, their frequencies (only haplotypes with frequencies over 1% were studied) and haplotypic effects on IMT. The haplotypes were reconstructed from the genotyped VEGF polymorphisms (-2578A>C, -634C>G, +936C>T) and the haplotypic effects were estimated separately for both sexes without covariates. *p<0.05. ref, reference haplotype; N, number of subjects; IMT, intima-media thickness; CombiMT, combined IMT and plaque thickness; Overall, mean of the different segments; ICA, internal carotid artery; BIF, carotid bifurcation; CCA, common carotid artery.
When the age- and sex-matched controls were compared with the AMI survivors, the haplotype $\text{AGT}$ was associated with the reduced AMI risk ($\text{OR} = 0.46$, 95% CI [0.25 - 0.86] $p=0.015$).

In conclusion, the haplotype $\text{CCC}$ was associated with higher overall IMT without plaques in women ($p=0.01$, haplotypic effect +0.03 mm), the haplotype $\text{CCT}$ with higher IMT without plaques in the internal carotid artery in men ($p=0.001$, +0.11) and the haplotype $\text{AGT}$ was associated with reduced AMI risk ($p=0.015$, OR=0.46) in the age- and sex-matched sample.
6 Discussion

6.1 Genetic loci regulating HDL-C levels (I)

A low HDL-C level has been found to be the most prevalent dyslipidemia associated with premature CHD (Genest et al. 1992). Given this close relationship, establishment of the genetic background to HDL regulation could illustrate the etiology and pathogenesis of CHD and provide new tools for its prevention and treatment.

We found suggestive evidence of linkage to four regions, 2q33, 4p12, 6p24 and 17q25, in the quantitative linkage analysis of HDL-C levels and three regions 6p12, 10p15.3 and 22q11 in the qualitative linkage analysis for low HDL-C phenotype. Five of these loci, 2q33, 6p12, 6p24, 17q25 and 22q11, are close to or at regions, which have shown significant evidence of linkage or association to CHD or HDL-C or both in genome-wide studies so far (Table 8). In addition, 17q25 has shown suggestive evidence of linkage to total cholesterol/HDL-C levels (Klos et al. 2001), and 22q11-q13 has provided suggestive evidence of linkage to HDL-C (Pollin et al. 2004, Harrap et al. 2006). A novel susceptibility locus, 4p12, yielded a LOD score of 3.1, which is close to the level that is generally considered as significant evidence of linkage. The region 4p15.1-p11 has previously shown suggestive evidence of linkage to HDL-C in a meta-analysis of African Americans (Malhotra et al. 2005). Although we could not show evidence of statistically significant linkage, one would like to speculate on the possible underlying genes near the highest linkage region 4p12. The closest functional candidate gene is peroxisome proliferator-activated receptor gamma, coactivator 1 alpha (PPARGC1A) on 4p15. The protein encoded by PPARGC1A is a transcriptional coactivator that regulates the genes involved in energy metabolism and interacts with peroxisome proliferator-activated receptor gamma (PPARG), which permits the interaction of this protein with multiple transcription factors. For example, PPARGC1A has been shown to regulate the human HL gene, which plays a key role in the metabolism of plasma lipoproteins (Rufibach et al. 2006).
Table 8. Regions showing suggestive evidence of linkage to HDL-C levels in this study compared with regions that have shown significant evidence of linkage or association to CHD or HDL-C or both in genome-wide studies so far.

<table>
<thead>
<tr>
<th>Suggestive evidence of linkage in this study</th>
<th>Significant evidence of linkage or association to CHD or HDL-C in previous genome-wide studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>4p12</td>
<td>-</td>
</tr>
<tr>
<td>6p12</td>
<td>6p12 (Canizales-Quinteros et al. 2003)</td>
</tr>
<tr>
<td>10q15.3</td>
<td>-</td>
</tr>
<tr>
<td>17q25</td>
<td>17q25 (Teslovich et al. 2010), 17q24 (Chasman et al. 2009), 17q21.32 (Schunkert et al. 2011)</td>
</tr>
<tr>
<td>22q11</td>
<td>22q11 (Teslovich et al. 2010)</td>
</tr>
</tbody>
</table>

Although no statistically significant evidence of linkage was observed, it was encouraging that we found suggestive evidence of linkage in seven chromosomal regions. By definition, suggestive evidence of linkage will be found by chance once in each genome scan, and therefore some of these hits should be real and not false positive results. The suggestive QTLs affecting HDL-C variance and suggestive loci for the low HDL-C trait in our sample were mostly located on different chromosomes, suggesting that the general variability in HDL-C levels may be affected by other genes than those causing the lowest HDL-C levels. It is also important to acknowledge the differences between the two statistical methods and the loss of information because of dichotomizing the trait for the qualitative analysis. The lack of replication of the reported linkages between the studies causes concern that some of the results are false positive findings. The genetic basis of HDL-C regulation seems to be largely heterogeneous and different selection criteria in different studies also cause variability in the results. Lipid metabolism is such a complex system that oversimplified research questions may not lead to the same answers in different studies if the environmental background or study population is different, for example. Additionally, different analytical approaches, such as population- and case-control-based association mapping and family-based linkage scans give us different types of information of the same phenotype, for example.
6.2 The adiponectins and lipids in atherosclerosis – genetic and environmental determinants (I, II)

6.2.1 The relationship between the adiponectins

The heritability of the HMW adiponectin had not been estimated by the time when our study (II) publication process was started. Our estimates showed high heritability for total adiponectin, HMW adiponectin and the HMW/total ratio, even higher than the heritabilities of the lipid traits or BMI, which was supported by an Italian study reporting similar heritabilities for total and HMW adiponectin (Menzaghi et al. 2010). The genetic and environmental cross-correlations were strong between all the other adiponectin measures, but interestingly there was no genetic correlation between the total adiponectin level and the HMW/total ratio. Therefore, it seems that there is no single major locus that regulates both the level of total adiponectin and the proportion that is present as “the most functional form” (Wang et al. 2008), HMW isoform.

6.2.2 Genetic and environmental correlations

No significant genetic correlation was detected between HDL-C and the adiponectins, suggesting that although they consistently show metabolic correlations, their levels are according to our results regulated by environmental factors and different genes or more complex genetic pathways. However, there are differing results showing significant evidence of pleiotropy between metabolic syndrome-related traits (such as insulin and HDL-C) and adiponectin (Butte et al. 2005, Comuzzie et al. 2007, Henneman et al. 2010, Menzaghi et al. 2010). In a recent Italian study, significant genetic correlations were observed between HMW adiponectin and traits related to insulin resistance, but in contrast, none were observed between medium-molecular-weight (MMW) and LMW isoform or any trait (Menzaghi et al. 2010). The discrepancy between our results and other findings might be related to differences in the samples’ phenotypes, since our families were ascertained to have low HDL-C levels as the most important risk factor for CHD and the genetic correlation between adiponectin and HDL-C could be more tightly linked to metabolic syndrome overall. Additionally, we may not have had enough power to detect the pleiotropy. However, we detected significant environmental correlations between total or HMW adiponectin and HDL-C when using the whole study sample. A similar,
significant correlation was observed when comparing probands and spouses in the partial correlation analysis, also after adjusting for BMI. Thus, these findings confirm the previous reports of a strong positive correlation between HDL-C and adiponectin levels.

Interestingly, the only significant (negative) genetic correlation between the adiponectins and lipids was detected between HMW-adiponectin and LDL-C. This suggests that there may be a gene affecting both HMW adiponectin and LDL-C regulation simultaneously, but in the opposite directions. No significant correlation between LDL-C and the adiponectins was observed in the partial correlation analysis using probands and spouses, but this might be explained by the finding of the bivariate analysis that the environmental and genetic correlations are opposite and may therefore obscure each other when the correlations are not decomposed into genetic and environmental determinants.

6.2.3 A link between adiponectin, HDL-C and LDL-C?

It can be observed from the Figure 5 that HL could be one possible link between elevated adiponectin levels and lower plasma concentrations of LDL-C and higher levels of HDL-C. HL activity has been shown to be negatively associated with adiponectin (Clarenbach et al. 2007), even independently of other factors such as markers of insulin resistance or inflammation (Schneider et al. 2005). In addition, HL plays an important role in mediating HDL metabolism since it converts larger HDL particles into smaller HDL remnants, pre-β HDL, and lipid-poor or -free apoA-I, and accordingly, lower HL activity is associated with higher HDL-C concentrations (Lewis & Rader 2005). Also VLDL particles can be modified into LDL particles through the action of LPL and HL. HL deficiency in humans is associated with diminished conversion of VLDL remnants into IDL and a near complete absence of IDL to LDL conversion (Zambon et al. 2003). Therefore, lower HL activity could be associated with lower LDL-C concentrations. The region for the HL gene, 15q22, was among the loci showing strongest evidence of linkage to HDL-C levels after statin correction. In addition, the finding on 4p12 showing “close to significant” evidence of linkage to HDL-C could in fact be due to the effect of PPARGC1A, which has been shown to regulate HL gene (Rufibach et al. 2006). However, the observed linkages were not statistically significant and the bivariate analysis failed to show any major genetic loci regulating both adiponectin and HDL-C levels. This does not exclude the possibility that the correlation between these traits could be due to a shared
regulatory pathway that is very complex involving multiple transcription factors, interactions and feedback loops, and cannot be observed by using bivariate analysis. The observed strong and negative genetic correlation between HMW-adiponectin and LDL-C could be due to a shared metabolism-related transcription factor, for example, regulating both of these traits and further activating HDL-C regulatory pathways. However, the correlation requires further studies, including bivariate analysis of LDL-C and adiponectin in an appropriate sample. The bivariate linkage analysis was not performed for the adiponectins and LDL-C in our sample, since the principal interest of the study was to investigate the positive correlation between HDL-C and the adiponectins, and the families were collected based on their CHD and low HDL-C phenotype, stressing the importance of low HDL-C, not high LDL-C in the pathogenesis of CHD in the patients. Therefore, the sample was considered to be too biased for further LDL-C analysis.

6.2.4 Adjusting for BMI reveals a possible link between total adiponectin and apolipoproteins

After adjustment for BMI in addition to sex in the univariate linkage analysis of the adiponectins, the highest evidence of linkage was shown in regions 3q13.2-q24 (HMW adiponectin) and 11q24.2 (total adiponectin). These regions have recently been associated with adiponectin (3q13.2 and 11q23.1) (Jee et al. 2010), HDL (3q22.3 and 11q23.3) (Chasman et al. 2009), CAD (3q22.3) (Erdmann et al. 2009), APOA1, TG, LDL, APOB and VLDL (11q23.3) (Chasman et al. 2009) in GWAS, and in addition, region 11q24.2 has previously shown linkage to total adiponectin (Tejero et al. 2007). Even if not statistically significant, the linkage to the region 11q24 is interesting because it includes the APOA1 gene, which is closely linked with two other apolipoprotein genes, apolipoprotein A-IV (APOA4) and apolipoprotein C-III (APOC3). ApoA-I is the major protein component of HDL and ApoC-III is a VLDL protein. ApoC-III inhibits LPL and HL and it is thought to delay catabolism of TG-rich particles. Adiponectin levels have previously been negatively associated with HL activity (Clarenbach et al. 2007), positively with LPL activity (von Eynatten et al. 2004) and positively with apoA-I (Verges et al. 2006, Matsuura et al. 2007, Tsubakio-Yamamoto et al. 2008). If the linkage were true and the apolipoprotein gene cluster regulated the total adiponectin levels, it would at least partly explain the strong association between adiponectin and HDL-C levels and provide a link between adiponectin levels, HDL and TG-rich particles, such as LDL and VLDL. Furthermore, it has been
recently shown that increased HDL-C following human apoA-I transfer increases plasma adiponectin levels and adiponectin mRNA expression in adipocytes (Van Linthout et al. 2010). However, if apoA-I did regulate adiponectin levels, one might think that we should have found a genetic correlation between HDL-C and adiponectin levels, since apoA-I is such an important component of HDL. On the other hand, the genome-wide scan for loci regulating HDL-C levels showed no evidence of linkage to the chromosomal region containing the apoA-I gene, suggesting that apoA-I variation is not a major determinant of HDL-C levels in our samples.

6.2.5 Overlapping linkage regions on chromosome 6

Even if the univariate linkage analyses of total and HMW adiponectin suggested different chromosomal regions for their regulations, there was some overlapping between adiponectin and HDL-C scans. Although there were no highly significant linkages, one would like to speculate that the suggestive evidence of linkage in the univariate linkage analysis of HMW-adiponectin reveals an interesting locus, 6p21, since it might connect HMW-adiponectin with HDL-C and atherosclerosis. Adiponectin has been suggested to protect against CHD via positive effects on HDL metabolism by increasing apoA-I synthesis and secretion (Matsuura et al. 2007) and decreasing its catabolism (Verges et al. 2006), by increasing ABCA1 expression (Matsuura et al. 2007), and by increasing apoA-I-mediated cholesterol efflux from macrophages through ABCA1-dependent pathway by the activation of liver X receptor-alpha (LXRA) and PPARG (Tsubakio-Yamamoto et al. 2008). One of the functional candidate genes located in the region 6p21 is RXRB, which is a transcription factor upregulating ABCA1 and ApoA-I-mediated cholesterol efflux (Schwartz et al. 2000). RXRB might upregulate ADIPOQ transcription in addition to HDL-C-related pathways and thus connect adiponectin levels with HDL-C levels and atherosclerosis. There are also other genes associated with atherosclerosis located in that region. One of them is PPARD, which codes for a nuclear transcription factor regulating lipid metabolism (Oliver et al. 2001). Additionally, LDL-associated phospholipase A2 (LDL-PLA2), located on that region, has been reported to be associated with atherosclerosis (Wang et al. 2010b) and suggested to be a reliable marker of risk for cardiovascular events (Madjid et al. 2010). The corresponding region has shown evidence of linkage to adiponectin mRNA levels also in baboons (Tejero et al. 2008) and evidence of linkage or

Since 6p24-22 showed suggestive evidence of linkage in the quantitative univariate linkage analysis of HDL-C, SNPs on functional candidate genes PPARD and RXRB in a nearby region, 6p21 were already genotyped in the original HDL-C genome scan, but they showed no statistically significant evidence of association with HDL-C levels in our sample. However, this result may be due to the lack of power for the family-based association analysis.

The region 6p21 also showed suggestive evidence of linkage in the bivariate analysis of HDL-C and HMW adiponectin and is close to 6p24, which showed suggestive evidence of linkage in the bivariate linkage analysis of HDL-C and total adiponectin. However, the results of the bivariate linkage analyses should be interpreted carefully, because no significant genetic correlation could be shown between HDL-C and the adiponectins. Overall, the bivariate linkage analysis failed to reveal any new linkage regions compared to the univariate linkage analyses and the results of the bivariate analyses most probably reflect QTLs regulating either HDL-C or adiponectin levels.

Adjustment for BMI in addition to sex in the univariate linkage analysis for HMW adiponectin decreased the evidence of linkage on 6p21; this might indicate that the chromosomal region is important in the complex regulation of BMI, which is associated with the regulation of HDL-C and adiponectin. On the other hand, all these variables are very closely associated and may have common regulatory mechanisms, which means that taking BMI as a covariate might in fact obscure some real findings.

6.2.6 The adiponectins and lipids in atherosclerosis - conclusions

It seems that there is no shared major locus regulating both the levels of total adiponectin and its most functional form, HMW adiponectin, or the levels of the adiponectins and HDL-C. However, HDL-C and the adiponectins show strong environmental correlation. Considering the observed genetic correlation between adiponectin and LDL-C, it is conceivable that there is a gene regulating both of them, but in the opposite directions. In the discovered linkage regions there are multiple candidate genes that could take part in the complex interplay between the adiponectins and lipids in atherosclerosis (Figure 12). For comparison of the overlapping linkage regions, the most significant results of the univariate and bivariate linkage analyses are combined in Figure 11.
Fig. 11. Combined linkage results. A human karyotype, i.e. an individual's collection of chromosomes, is presented in the middle of the picture. The results of the linkage analyses are presented in the graphs around the karyotype. Chromosomes showing suggestive evidence of linkage have been circled and the dashed lines connect the chromosomes to their results. The graphs show results from the following analyses:
6.3 VEGF and atherosclerosis (III)

Even if none of the single genotyped polymorphisms was significantly associated with overall IMT (without plaques) or with the risk of AMI, men with the 936 CT genotype (the +936C>T polymorphism) exhibited higher IMT values than men with CC genotype (p=0.01) on ICA (IMT without plaques) and overall CombiIMT (p=0.02). The allelic effect was not logical, since TT genotype seemed to be associated with smaller IMT (not significant), and this trend was the same in most of the segments where IMT was measured. The number of samples in the TT genotype group was very small, which introduces a higher risk of random fluctuations to the results. However, it is possible that the results are due to chance, even if correction for multiple tests was made. Our results for the haplotype analysis suggested logically that haplotype CCC was associated with higher IMT in women, and would therefore be pro-atherogenic, whereas the “contrary” haplotype, AGT was associated with reduced AMI risk in the sex- and age-matched study population. In men, the haplotype CCT was associated with higher IMT, but the number of samples in that haplotype group was very small. On the other hand, the trend for the effect of haplotype CCC was the opposite in men (not significant) to that in women. The difference between the sexes might be related to the possible effect of the 936CT genotype in men. Higher VEGF production could be pro-atherosclerotic, since C-alleles have previously been associated with higher VEGF production, and in our study haplotypes CCC and CCT are associated with higher IMT. Also, AG haplotype has been associated with lower VEGF levels (Lambrechts et al. 2003), and in our study AGT was associated with reduced AMI risk. Our findings are, nevertheless, contradictory to some previous findings suggesting that the AA genotype of the polymorphism -2578 is associated with three vessel-disease, and VEGF could therefore have a protective effect in atherosclerosis (Howell et al. 2005, Biselli et al. 2008, Guerzoni et al. 2009). However, there was no significant difference in allele and genotype distribution between the groups in any of these studies before some further adjustments. The association was only seen for three-vessel disease, and in the study of Howell et al. after excluding patients with AMI, after which the frequency of the AA genotype increased with the number of involved vessels compared with the CC genotype as reference. Additionally, no association between VEGF -2578 polymorphism and atherosclerosis was found in a Chinese population (Lin et al. 2010). The discrepancy between our results and the previously reported results may be due to some false positive findings or
differences in the study populations or quantification of atherosclerosis. It is also possible, that the effect is so weak, that we were not able to detect it with our sample size. Development of atherosclerosis is a complex process with multiple factors and stages, which has to be taken into account when comparing the results of the different studies. It is possible that VEGF has pro- and anti-atherosclerotic properties depending on the environment and other factors, and it is likely that the genetic risk profiles for early atherosclerosis, advanced atherosclerosis and AMI are somewhat different. It is possible that VEGF does not play an important role in early atherosclerosis but is somehow activated in advanced atherosclerosis when there is a severe risk of tissue ischemia. Furthermore, a failure in the activation of VEGF-induced neovascularization could increase the risk of AMI. In most of the studies, some further adjustments towards more extreme phenotypes needed to be done in order to find significant differences between the groups. In our study, none of the single genotyped polymorphisms was significantly associated with overall IMT (without plaques) or with the risk of AMI, which is similar to the findings of Douvaras et al (Douvaras et al. 2009) and Lin et al (Lin et al. 2010), and we can conclude that it is unlikely that the studied variations in the VEGF gene are of major importance in determining the degree of atherosclerosis or the risk of AMI. However, it is possible that some variations in the VEGF gene or its promoter region are involved in the development of atherosclerosis, but in previous studies their effect could only be seen when using certain stages of the disease and in our study when using the information about the haplotypes.

6.4 Combined results

The results of all the studies are combined in the Figure 12.
6.5 Methodological considerations

During the course of this thesis, the available information for genetic epidemiological studies has increased dramatically. Lowering genotyping costs, advances in nanotechnology and increased capacity in data storage and computing have prepared the way for publicly available full sequence of the human genome and more detailed catalogue of polymorphisms and haplotype construction, which have enabled scientists to perform studies that they were only dreaming about ten years ago. These new approaches provide opportunities to revisit families like the ones used in this study and to combine data of high and low throughput to get a better picture of the landscape of the genetic architecture of atherosclerosis.
6.5.1 Genotypings and study populations

Genotypings of studies I and II were started in 1996, when the available tools for laboratory work and data handling were not as automated as today and, accordingly, more prone to risk due to human errors. For example, possible genotyping errors might decrease the power to detect association between a trait locus and a marker locus. The data were, however, checked for Mendelian errors.

Another risk in all the studies is that if the study population is too small, it may lack statistical power to detect true effects. In addition, multiple testing results in a well-known risk of false positive results, especially if the sample size is small compared to the number of tests performed. The estimation of the needed sample size for a certain power, as performed for study III, requires assumptions of some other variables that may not be available at the stage of designing the study, such as the difference of allele frequencies between cases and controls. Therefore, some of these variables have to be estimated based on literature or one has to proceed in the study to find them out, as was done in our study. In the studies I and II, empirical p-values and the level of suggestive evidence of linkage were estimated using simulation, and this estimation confirms that we have used appropriate levels of significance taking into account our study material when reporting the results.

Heterogeneity can make it difficult to detect statistically significant results. To reduce heterogeneity and increase power to detect linkage, we had three strengths in our sample in studies I and II: the sample was collected from a relatively isolated geographical region, it contained large families with multiple genotyped relatives and the probands had a specific phenotype, not only suffering from CHD but also from low HDL-C. In addition, the probands had fewer traditional risk factors of atherosclerosis since they had only moderately elevated levels of TG and total cholesterol and no diabetes and, accordingly, it could be thought that HDL-C could have played a central role in the development of atherosclerosis in the probands.

In study II we used the extended pedigrees in the bivariate linkage analysis to decompose the correlations of the adiponectins and lipids into genetic and environmental determinants. However, we also wanted to investigate the correlations in general, without decomposing and without the genetic relationships. Therefore, we performed an additional analysis, in which only the probands were compared to their spouses when partial correlations (adjusted for sex, age, BMI and statin use) between adiponectins and lipids were calculated.
6.5.2 IMT

Measurement of carotid IMT by ultrasonography provided us with a noninvasive, quantitative tool to estimate the effect of the VEGF gene polymorphisms on atherosclerosis. IMT is associated with the degree of coronary atherosclerosis (Craven et al. 1990) and also with the risk of AMI (van der Meer et al. 2004) and has been considered as a valid surrogate marker of atherosclerosis (de Groot et al. 2004). IMT has a clinically applicable diagnostic accuracy for the presence of angiographically significant CHD, and IMT values help us define zones with high and low probability for the presence of advanced coronary atherosclerosis (Simova et al. 2009): IMT is thus not a completely quantitative measure of atherosclerosis. However, it is considered to reflect increased accumulation of lipids, foam cells and matrix components in the arterial wall. IMT is usually only measured from the CCA, but in this study, it was measured from several different segments, which possibly improves the accuracy of the overall measure (Ibanez et al. 2009). We considered the overall IMT without plaques to be the most reliable values, since they were the mean of 10 site-specific measurements, in each of which the thickest point of the IMT was measured, although avoiding sites with plaque. The original analysis was based on this overall measure, but the site-specific values and CombiIMT values were also analyzed separately as a result of the revision process.

6.6 Future studies

In 2007 the Wellcome Trust Case Control Consortium, the largest ever study of the genetics behind common diseases at that time, published the first groundbreaking results substantially increasing the number of genes known to play a role in the development of some of our most common diseases, such as CHD (The Wellcome Trust Case Control Consortium 2007). It had analyzed half a million genetic variants in each of 17,000 DNA samples from people across the UK, bringing together 50 leading research groups and 200 scientists in the field of human genetics from dozens of UK institutions. The £9 million study was one of the UK’s largest and most successful academic collaborations and showed that it is possible to analyze human variation in health and disease on an enormous scale. It showed the way to “modern genetic studies” and inspired researchers to start to perform GWAS. In 2010 a GWAS for serum lipids (total cholesterol, LDL-C, HDL-C and TG) in >100,000 individuals of European ancestry was published.
(Teslovich et al. 2010). It was not only a meta-analysis of 46 lipid GWAS, but also included an evaluation of the mapped variants in East Asians, South Asians, and African Americans, association testing in individuals with and without coronary artery disease (CAD), evaluation of genetic variants in patients with extreme serum lipid concentrations, and genetic manipulation in mouse models. This enormous study identified 95 loci that showed association with at least one of the four traits tested, and these include all of the 36 loci previously reported by GWAS at genome-wide significance. Altogether 31 novel associations to HDL-C were identified and all the mapped SNPs were estimated to explain 12.1% of the total variance of the trait in the Framingham Heart Study. The study also demonstrated that some of the new loci contain genes of clear biological and clinical importance.

In the future we need more collaboration, large samples and thorough phenotyping. However, it is not possible to increase sample size indefinitely and challenges of population stratification and multiple testing must be taken seriously. When 100,000 samples, which is a huge study population, is compared with the 4 (number of lipid traits) times 2.6 million (number of SNPs) tests being done, one can understand the scale of multiple testing. It has been suggested that the genetic architecture of circulating lipids involves a number of undiscovered variants with very small effects, and that increasing GWAS sample sizes will enable the identification of novel variants that regulate lipid levels (Demirkan et al. 2011). At the population level GWAS are currently identifying potential loci linking lipid metabolism with disease pathogenesis, but in addition to genetic association studies of large scale, we need research in well-characterized data in different populations, “special” family material and well-conducted expressional and functional studies to truly understand the development of atherosclerosis. While whole-genome sequencing is still too expensive for large cohort studies, exome resequencing is increasingly becoming a standard tool for the discovery of genes underlying rare monogenic disease and the discovery of coding variants associated with common disease (Ng et al. 2010, Coffey et al. 2011). In the future, high-throughput omics studies, including transcriptomics, metabolomics (Inouye et al. 2010) and proteomics and also the analysis of tissue, cell and even organelle lipidomes (Meikle & Christopher 2011) will help us unravel the complex relationships between lipoproteins, inflammation, obesity and atherosclerosis.
7 Conclusions

This study investigated the genetic background of atherosclerosis and, in more detail, the genetics of HDL-C. The main findings of the study were as follows:

1. We identified seven chromosomal regions (2q33, 4p12, 6p24, 6p12, 10p15.3, 17q25 and 22q11) for HDL-C regulation exceeding the level for suggestive evidence of linkage, with the highest LOD score of 3.1 in a novel region 4p12. No significant evidence of association was found for two functional and positional candidate genes PPARD and RXRB.

2. Adiponectin was inversely associated with CHD in men. The strongest environmental correlation was shown between HDL-C and total adiponectin, whereas the strongest genetic correlation (even after adjustment for BMI and statin use) was detected between HMW adiponectin and LDL-C suggesting a shared genetic factor for their regulation. The adiponectins showed strong heritability and we found three suggestive chromosomal regions, 11p15 for total adiponectin and, for the first time, 6p21 and 3q13.2-q24 for HMW adiponectin regulation. We could not show evidence of pleiotropy between HDL-C and the adiponectins, so it is possible that their metabolic correlation may be regulated by more complex genetic pathways and environmental factors, or that the genetic architecture of the adiponectins does not strongly overlap with the genetics of atherosclerosis in families with low HDL-C and CHD.

3. Variation in the VEGF gene was weakly associated with IMT and the risk of AMI, but the effect could only be observed when IMT was measured in certain segments of the carotid artery or when plaques were included, or when the information of the SNPs was combined by constructing haplotypes. The haplotype AGT was anti-atherosclerotic, associating with reduced AMI risk in age- and sex-matched subjects, whereas “the contrary” haplotypes CCC and CCT were pro-atherosclerotic, associating with higher IMT in women and men, respectively. In conclusion, higher VEGF production could be pro-atherosclerotic, since alleles associated with higher VEGF production according to the previous findings were associated with higher IMT in our study. However, since none of the single genotyped polymorphisms was significantly associated with overall IMT (without plaques) or with the risk of AMI, it is unlikely that the studied variations in the VEGF gene are of major importance in determining the degree of atherosclerosis or the risk of AMI.
References


Oliver WRJr, Shenk JL, Snaith MR, Russell CS, Plunket KD, Bodkin NL, Lewis MC, 
Winegar DA, Sznaidman ML, Lambert MH, Xu HE, Sternbach DD, Kliewer SA, 
Ouchi N, Kihara S, Arita Y, Maeda K, Kuriyama H, Okamoto Y, Hotta K, Nishida M, 
modulator for endothelial adhesion molecules: adipocyte-derived plasma protein 
Ouchi N, Kihara S, Arita Y, Nishida M, Matsuyama A, Okamoto Y, Ishigami M, 
Kuriyama H, Kishida K, Nishizawa H, Hotta K, Muraguchi M, Ohmoto Y, Yamashita 
S, Funahashi T & Matsuzawa Y (2001) Adipocyte-derived plasma protein, 
adiponectin, suppresses lipid accumulation and class A scavenger receptor expression 
Ouchi N, Kobayashi H, Kihara S, Kumada M, Sato K, Inoue T, Funahashi T & Walsh K 
(2004) Adiponectin stimulates angiogenesis by promoting cross-talk between AMP- 
activated protein kinase and Akt signaling in endothelial cells. J Biol Chem 279: 
1304–1309.
Pagani F & Baralle FE (2004) Genomic variants in exons and introns: identifying the 
Nuotio I, Laakso M, Rotter JI, de Bruin TW, Cantor RM, Lusis AJ & Peltonen L 
(2003) Combined analysis of genome scans of dutch and finnish families reveals a 
susceptibility locus for high-density lipoprotein cholesterol on chromosome 16q. Am J 
Pajukanta P, Cargill M, Viitanen L, Nuotio I, Kareinen A, Perola M, Terwilliger JD, 
Kempas E, Daly M, Lilja H, Rioux JD, Brettin T, Viikari JS, Ronnemaa T, Laakso M, 
Lander ES & Peltonen L (2000) Two loci on chromosomes 2 and X for premature 
coronary heart disease identified in early- and late-settlement populations of Finland. 
M, Knopps A, Xiang AH, Utzschneider KM, Kahn SE, Olefsky JM, Buchanan TA & 
Scherer PE (2004) Complex distribution, not absolute amount of adiponectin, 
correlates with thiazolidinedione-mediated improvement in insulin sensitivity. J Biol 
Chem 279: 12152–12162.
Parthasarathy S, Barnett J & Fong LG (1990) High-density lipoprotein inhibits the 
oxidative modification of low-density lipoprotein. Biochim Biophys Acta 1044: 275– 
283.
Peacock JM, Arnett DK, Atwood LD, Myers RH, Coon H, Rich SS, Province MA & Heiss 
G (2001) Genome scan for quantitative trait loci linked to high-density lipoprotein 
cholesterol: The NHLBI Family Heart Study. Arterioscler Thromb Vasc Biol 21: 
1823–1828.


The Wellcome Trust Case Control Consortium 2007 Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678.


Original articles


*with equal contribution

Reprinted with permission from Nature Publishing Group (I), Elsevier (II) and Informa Healthcare (III).

Original publications are not included in the electronic version of the dissertation.
1106. Juuti, Anna-Kaisa (2011) Sleep disorders and associated factors in 56-73 year-old urban adults in Northern Finland

1107. Marttala, Jaana (2011) First trimester screening and Down syndrome


1109. Paalanne, Niko (2011) Postural balance, isometric trunk muscle strength and low back symptoms among young adults


1111. Pradhan-Palikhe, Pratikshya (2011) Matrix metalloproteinase-8 as a diagnostic tool for the inflammatory and malignant diseases

1112. Kaakinen, Mika (2011) Functional microdomains in the specialized membranes of skeletal myofibres


1114. Vuori, Maire (2011) Pohjoissuomalaisten suurten ikäluokkien tulevaisuudenkuvaatiksinkymmestään, hyvinvoinnissaan ja sosiaali- ja terveyspalveluissaan

1115. Hakalahi, Anna (2011) Human \(\beta_1\)-adrenergic receptor: biosynthesis, processing and the carboxyl-terminal polymorphism

1116. Peltonen, Jenni (2011) TP53 as clinical marker in head and neck cancer


1118. Suorsa, Eija (2011) Assessment of heart rate variability as an indicator of cardiovascular autonomic dysregulation in subjects with chronic epilepsy


1120. Venhola, Mika (2011) Vesicoureteral reflux in children

1121. Naillat, Florence (2011) Roles of Wnt4/5a in germ cell differentiation and gonad development & ErbB4 in polarity of kidney epithelium


Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/
Tiia Kangas-Kontio

GENETIC BACKGROUND OF HDL-CHOLESTEROL AND ATHEROSCLEROSIS

LINKAGE AND CASE-CONTROL STUDIES IN THE NORTHERN FINNISH POPULATION