The molecular basis of early onset cardiovascular disease
Sivapalaratnam, S.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
GENERAL DISCUSSION

SUMMARY AND FUTURE PERSPECTIVES
GENERAL DISCUSSION

The genetics of cardiovascular disease (CVD) - The bigger picture

Cardiovascular disease (CVD) remains the major cause of worldwide morbidity and mortality with an estimated number of 18 million deaths per year; it represents 31% of all global deaths (www.who.org). CVD is not only a problem in developed countries, but also increasingly in developing countries, illustrated by the fact that over 80% of all CVD deaths takes place in low- and middle-income countries. Moreover, whereas cardiovascular mortality in Europe is slowly declining, a rise is expected in less prosperous regions.

Since the phenotype of CVD often shows a marked heritable pattern, it is likely that genetic factors play an important role. Twin studies indicate that heritability of CVD, defined as the proportion of the inter-individual differences resulting from genetic factors, is estimated to be 30-60% \(^1\). This heritability is considered to be the net result of common and rare genetic variants with small and large effects on disease expression, respectively (figure 1).

**Figure 1.** The genetic landscape of disease risk

This figure depicts the genetic architecture of CVD. It consist on one hand of rare variants with high effect sizes, which can be discovered by family studies and common variants with low effect, which can be discovered by large GWA studies. FH, discussed in Part I of the thesis and the families described in Part III belong to the top left circle. The common variants tested in chapter two and those identified chapter 9 belong to the bottom right circle (with personal permission of Maniolo).
A large number of studies have been performed to decipher the molecular mechanisms underlying the heritability of CVD. Recently, due to major technical advances, much progress has been booked. For many years, studies were performed in families since they have been shown to be instrumental in the identification of pivotal genes. The identification of the gene encoding the LDL-receptor, for example, and the unravelling of its role in lipid metabolism arose from a family study. In other early attempts, case-control studies focused on the differences in carrier frequency of genetic variations in “candidate genes” amongst cases and controls. Genes were marked as “candidate” based on the fact that they were known to influence traditional risk factors (i.e. genes involved in lipid metabolism) or were part of pathways with a putative role in atherogenesis.

Currently, a whole genome approach can be applied to find genetic variations without a prior hypothesis. Initial attempts included classical linkage studies in families with premature CAD using microsatellite markers and later SNP arrays for linkage. These first family based studies have thus far resulted in the identification of two interesting candidate genes, \textit{MEF2A} and \textit{LRP6}.

To increase the power to detect more commonly occurring shared genetic loci, linkage analysis has been performed with data derived from several families. These studies have resulted in the identification of genetic variations in three genes involved in atherosclerosis: arachidonate 5-lipoxygenase-activating protein (ALOX5AP), leukotriene A4 hydrolase (LTA4H), both pro-inflammatory cytokines and LDL-receptor-related protein 8 (LRP8) also known as apoE receptor 2 (APOER2). The relative contribution of these loci to atherosclerosis in replication studies remains modest with odds ratios not exceeding 1.5.

In recent years, numerous genome wide association (GWA) studies with CVD as outcome have been published. These studies primarily focus on populations and are carried out under the assumption that common variants are shared by subjects with a pre-specified phenotype. So far, twenty-five loci comprising common genetic low-risk variants associated with Coronary Artery Disease (CAD) were identified by Genome Wide Association (GWA) studies (figure 2). These variants combined explain up to 10% of the heritability of CAD. Our AMC PAS-cohort has been part of most of these studies.

The inability to unravel the full heritability of CAD is merely a reflection of its genetic heterogeneity and of our limited ability to classify CAD cases into subgroups that are more homogeneous in their molecular pathology. Mendelian forms of disease are considered the most ultimate form of homogeneity in human disease, and as such, hold great promise to identify monogenetic causes of disease. Furthermore it is expected that sequencing at a higher resolution, rather then genome wide SNP typing will result in even more disease associated variants.
The high heritability of CVD translates into the fact that apart from the classical risk factors, a positive family history for premature coronary artery disease (CAD) in a first degree relative also constitutes an additional and independent risk\textsuperscript{10}. Still, the current 25 common variants associated with CVD risk have little impact on routine clinical medicine, since to date they cannot be directly translated into a risk profile for an individual patient. We expect that a substantial part of the missing heritability will be explained by rare variants with severe effects on the phenotype. Some rare variants causing CVD could even be \textit{de novo} mutations, as recently found in mental retardation\textsuperscript{11}. Others, might be detected by studying families with severe disease, defined by a combination of very early onset CVD and multiple affected individuals per family, for example by a combination of exclusion linkage studies and next generation sequencing as done in the final chapters of this thesis. Hopefully with the advent of risk scores integrating both common and rare yet to be identified variants risk prediction will improve in the near future.

Furthermore, even though, current treatment modalities to prevent CVD have been successful, the majority of cardiovascular events are not prevented. Improved understanding of the molecular mechanism, will ultimately lead to novel targets for therapy, which are desperately needed to further decrease the worldwide CVD burden.

SUMMARY

The main aim of this thesis was to unravel the molecular basis of early onset CVD. We chose to study subjects with early onset CVD, because it has been shown that the heritability of this specific phenotype is larger than events occurring at late age. Environmental factors have a lesser role in the development of early onset disease. The enrichment for heritability increases the a priori chance of finding a causal gene and thus an attractive disease phenotype to study CVD genetics.

PART I: Familial Hypercholesterolemia, a known cause of early onset CVD

The focus of Part one is FH, the major cause of early onset CVD, characterized by elevated LDL-C levels resulting in lipid accumulation in the arterial wall and as a consequence accelerated atherosclerosis and early onset CVD\textsuperscript{12, 13}. The onset and severity of CVD varies considerably between FH patients, even among individuals who share an identical gene defect\textsuperscript{14}.

In chapter one we hypothesized that part of the phenotypical variability later in life might be due to differences in maternal or paternal inheritance of FH and its effect pre birth. It was unknown whether elevated maternal low-density lipoprotein cholesterol (LDL-C) levels lead to dyslipidaemia in the offspring. Since this could have important consequences for cardiovascular prevention in mother and child, we explored the relationship between maternal familial hypercholesterolemia (FH) and lipids in adult offspring. In a large cohort of both Dutch and Canadian origin we compared lipid profiles between patients, aged 18 – 85 years, who inherited FH maternally (n=1069) and those who inherited FH paternally (n=1270). This relationship was evaluated using multivariate regression analyses. Levels of total cholesterol (TC), LDL-C, and apolipoprotein B (ApoB) were significantly elevatred in patients who inherited FH maternally compared with patients who inherited FH paternally (adjusted differences in TC: 0.156 mmol/L, p=0.037; LDL-C: 0.187 mmol/L, p=0.012; ApoB: 0.064g/L, p=0.022). These data show that maternal hereditary hypercholesterolemia slightly increases TC, LDL-C and ApoB levels in their offspring later in life. Although the molecular mechanisms underlying these observations still require elucidation, our data suggest that maternal hypercholesterolemia during
pregnancy may program lipid metabolism to a certain extent in the foetus. We carefully speculate that these slightly elevated lipids levels over a long period of time might have consequences for the occurrence of cardiovascular events.

In recent years, multiple loci dispersed on the genome have been shown to be associated with coronary artery disease (CAD). In chapter two we investigated whether these common genetic variants also hold value for CAD prediction in a large cohort of patients with Familial Hypercholesterolemia (FH). A total of 146 SNPs in 1701 FH patients were genotyped, of whom 482 patients (28.3%) had at least one cardiovascular event during 112,943 person-years follow-up. The association of each SNP with event-free survival time was calculated with a Cox proportional hazard model. In CVD risk adjusted analysis, the lead SNP at the well-known 9p21 locus rs1333049 near CDKN2B-AS1 had a HR for CAD risk of 0.82 (95%CI 0.77-0.87; p-value 0.000945). None of the other tested CAD-associated SNPs were significantly associated with CAD risk (p > 8.9*10^-5 for each). In conclusion, of all the loci analyzed, the 9p21 locus had the strongest negative association with CAD in this high-risk FH cohort. None of the SNPs at neither this 9p21 locus, however, nor any of the other tested CAD-associated SNPs were significantly associated with risk of CAD according to a priori defined significance threshold that took into account multiple testing.

Finally, in chapter three we investigated the effectiveness of statins in daily practice in reducing the arterial wall thicknesses by comparing the carotid intima media-thickness (cIMT) between statin-treated Familial Hypercholesterolemia (FH) patients and their unaffected spouses. Clinical data and carotid intima-media thickness (cIMT) as surrogate marker for atherosclerosis were acquired. In total 40 FH patients, age 48.4 ± 4.2 years, and their 40 unaffected spouses, age 47.4 ± 3.9 years, were included. Pre treatment total cholesterol levels of FH patients were on average 9.3 ± 2.0 mmol/L. Treated FH patients and unaffected spouses exhibited similar LDL-c (3.8 ± 1.5 vs. 3.5 ± 1.1 mmol/L; p=0.25) and total cholesterol levels (5.8 ± 1.6 vs. 5.6 ± 1.1 mmol/L; p=0.56). Also, in a multivariate model cIMT adjusted for age and sex did not differ between affecteds and spouses (95% CI: -0.032 to 0.092 mm; p=0.34). In conclusion, after long-term statin treatment cIMT in severe FH patients has normal values and therefore it is likely that they are no longer at high risk for CVD, reaffirming the need for identification and treatment of those at high risk for early onset CVD.

**PART II: Heritability of CVD, Risk prediction and prevention**

In Part two we examined the role of family history in CVD risk prediction and whether apparently healthy individuals with a family history of CVD and subclinical atherosclerosis benefit from primary prevention.
In **chapter four** we tested whether adding family history of premature CHD in first degree relatives improves CVD risk prediction compared to the well established Framingham risk score (FRS) alone. This study comprised 10,288 men and 12,553 women aged 40 to 79 years participating in the EPIC-Norfolk cohort who where followed for an average of 10.9 ± 2.1 years (mean ± SD). The Framingham risk score as well as a modified score taking into account family history of premature CHD were computed. In our study, family history of CHD was indeed associated with an increased risk of future CHD, independent of established risk factors (FRS-adjusted hazard ratio of 1.74 (95%CI 1.56-1.95) for family history of premature CHD). However, adding family history of CHD to the Framingham risk score resulted in a negative net reclassification of 2%. In the subgroup of individuals estimated to be at intermediate risk, family history of premature CHD resulted in an increase in net reclassification of 2%. The sensitivity increased with 0.4% and the specificity decreased 0.8%. In conclusion, although family history of CHD was an independent risk factor of future CHD, its use did not improve classification of individuals into clinically relevant risk categories based on the FRS. Among study participants at intermediate risk of CHD, adding family history of premature CHD resulted in, at best, a modest improvement in reclassification of individuals into a more accurate risk category.

In **chapter five** we tested whether a positive family history for CVD identifies patient who benefit from primary prevention. A post-hoc analysis on the database of the St. Francis trial was performed, to assess efficacy of treatment with atorvastatin 20mg/d in those with CCS >80th percentile and presence (n=543) or absence (n=462) of a positive family history for premature CAD. All participants received aspirin 81 mg/d. The primary outcome measure was a composite, included coronary death, myocardial infarction, coronary revascularization, stroke and arterial surgery. In this trial 1005 individuals participated; with a mean age 59.0±5.9 years. After a mean follow-up of 4.3 years, 7.2% of the treated individuals with a positive family history had a cardiovascular event versus 12.5% of the placebo group (hazard ratio (HR) 0.55 (95% confidence intervals (CI) (0.31-0.97; p=0.040). In individuals without a family history, events were minimally reduced: 6.6% in the treated versus 6.8% in the placebo group (HR 1.04 ; ( 95% CI 0.51-2.13; p=0.912). We concluded that the combination of a positive family history and CCS >80th percentile identifies a subgroup within the primary prevention population, who benefit from statin treatment to a greater extent than the population at large. These results might have important implications for future guidelines concerning individuals with a positive family history for premature CAD. We suggest that all apparently healthy first degree family members of individuals with early onset CAD patients undergo non-invasive imaging and treatment where appropriate. This is idea is confirmed by the editorial written in these same issue of the journal15.
Part III: Identification of novel causes of early onset CVD

Part three describes our efforts to identify novel candidate genes and genetic variation related to early onset CVD.

In chapter six, we reviewed the landscape of human genetics at present, and speculate about the near future in the context of next generation sequence technologies.

In chapter seven and eight we applied genome wide expression studies in monocytes to identify candidate genes for CVD. We specifically chose to study monocytes since these cells are involved in all sequels of atherosclerosis.

Atherosclerosis is considered a chronic inflammatory state. In chapter seven we hypothesized that activation of inflammatory pathways in monocytes would lead to proatherogenic changes in the monocyte transcriptome. To activate these pathways, endotoxin, lipopolysaccharide (LPS), or saline control was infused in healthy volunteers. All subjects who received LPS (n=11) experienced the anticipated clinical response, including high fever, indicating successful stimulation. One hour after LPS infusion, 11 genes were identified as being differentially expressed. Four hours after LPS infusion, 28 genes were identified as being differentially expressed. No genes were significantly differentially expressed following saline infusion (n=5). Comparison with results obtained in in vitro experiments lead to the identification of 6 strong candidate genes (BAT, BID, C3aR1, IL1RN, SEC61B and SLC43A3).

In chapter eight we compared the expression profiles in monocytes from twenty two young male patients with early onset familial CAD with unknown aetiology with those from twenty two controls matched for age, sex and smoking status, without a family history of CVD. Since all patients were on statins and aspirin treatment, potentially affecting the expression of genes in monocytes, twelve controls were subsequently treated with simvastatin and aspirin for 6 and 2 weeks, respectively. By whole genome expression arrays six genes were identified to have differential expression in the monocytes of patients versus controls; ABCA1, ABCG1 and RGS1 were downregulated in patients, whereas ADRB2, FOLR3 and GSTM1 were upregulated. Differential expression of all genes, apart from GSTM1, was confirmed by qPCR. Aspirin and statins altered gene expression of ABCG1 and ADBR2. All finding were validated in a second group of twenty four patients and controls. Differential expression of ABCA1, RSG1 and ADBR2 was replicated. Recent, in another cohort of early onset CVD cases, ABCA1 was differentially expressed (unpublished, personal communication Dr. Zeller).

Dyslipidemia accounts for approximately 50% of the population attributable risk of developing CVD. Genome wide association studies (GWAS) have identified many single-nucleotide polymorphisms (SNPs) underlying variations in plasma lipid levels.
In chapter nine we set out to explore whether additional loci associated with plasma lipid phenotypes, such as high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), total cholesterol (TC) and triglycerides (TG) can be identified by a dense gene-centric approach. The meta-analysis of 32 studies in 66,240 individuals of European ancestry was based on the custom ~50,000 SNP genotyping array covering ~2,000 candidate genes (the ITMAT-Broad-CARE (IBC) array). SNP-lipid associations were replicated in a cohort comprising an additional 24,736 samples or within the Global Lipid Genetic Consortium. A total of 4, 6, 10 and 4 new SNPs were identified in established lipid genes for HDL-C, LDL-C, TC and TG respectively. In addition several lipid-related SNPs in previously unreported genes were identified: DGAT2, GPR109A, LOC338328, PPARG, and FTO for HDL-C; SOCS3, APOH, SPTY2D1, BRCA2 and VLDLR for LDL-C; SOCS3, UGT1A1, BRCA2, UBE3B, FCGR2A, CHUK, and INSIG2 for TC; and SERPINF2, C4B, GCK, GATA4, INSR and LPAL2 for TG. The proportion of phenotypic variance explained in the subset of studies providing individual-level data was 9.9% for HDL-C, 9.5% for LDL-C, 10.3% for TC and 8.0% for TG. The explained phenotypic variance using this approach was comparable to meta-analysis of GWAS data suggesting that a focused genotyping approach can further increase the understanding of heritability of plasma lipids.

Finally, state-of-the-art next-generation-sequencing technologies were applied for family studies in chapter ten and eleven. Here, we set out to identify novel rare genetic variants in well phenotyped large families with Mendelian inheritance of early onset CVD with unknown aetiology.

In chapter ten, in a large four generation family with 11 family members suffering from early onset CVD classical linkage analysis revealed that a 4.4 Mb interval on chromosome 12 was linked to the phenotype with a parametric LOD-score of 3.31. Subsequent capture and sequencing of this region resulted in the identification of only one non-synonymous variant in Kera (NM_007035.3: c.920C>G; p.Ser307Cys), encoding for Keratocan, an extracellular matrix protein. This specific variant was not present in 9000 healthy control individuals nor in 1400 patients suffering from premature CVD. Upon sequencing the gene, we did not identify mutations in an additional 300 premature CVD cases. By immunohistochemistry we showed that Keratocan was not present in healthy arterial walls, but was highly expressed in the lipid rich areas of early atherosclerotic lesions and of fully developed plaques. Furthermore, in a mouse model where atherosclerosis was induced by cuff placement in the vessel wall, Keratocan expression was shown to be linearly associated with plaque size ($r^2=0.7$, $p < 0.001$). The potential biological relevance of Keratocan was exemplified by its presence in atherosclerotic plaques, but absence in healthy vessels and correlation with atherosclerosis progression in humans and mice.
In chapter eleven we studied another family with multiple individuals with early onset CAD. We performed exclusion linkage analysis and identified rare mutations in ZC3HC1 (NM_016478.3, c.913A>G, p.Ile305Val) and MCF2L (NM_001112732.1, c.2066A>G, p.Asp689Gly). It is noteworthy that common variation in ZC3HC1 has been associated to CAD risk in recent large genome wide association studies (GWAS) \(^{17,18}\). For MCF2L, SNP analysis in the Cardiogram Consortium resulted in an association that did not reach GWAS significance levels. Further studies are warranted to unravel the pathophysiological role of these genes in the complex phenotype of early onset CAD.

FUTURE PERSPECTIVES

The future holds a lot of challenges. Firstly, the research represented in this thesis has resulted in numerous candidate genes for further research. However, for most of these genes functionality nor direct causality of the discovered variants has been established. Unfortunately, simultaneously follow up all targets is not feasible. Therefore prioritization based on novelty, prediction models, with all their pitfalls, and current literature is warranted. Depending on the predicted pathway through which variations in the genes of interest result in atherosclerosis intelligent experiments need to be set up. So far, Keratocan has ranked as our top gene for further investigation. We have proven a significant correlation between plaque size and keratocan expression in humans and mice. To investigate the effect of the mutant protein on plaque size and composition, local overexpression of wild-type and mutant KERA in an atherosclerosis mice model is in progress. In this model atherosclerosis is induced by cuff placement in the carotid artery and a western type diet. Also keratocan knock out (KO) mice have been ordered to be crossed with ApoE KO mice. In order to find additional early onset CVD cases, outside the described family, who carry the similar or other rare variants sequencing of keratocan is in full progress in large international cohorts. For the same purpose the regional AMC-PAS cohort has been expended from 1000 to 2000 consenting cases. Secondly, we will be challenged by the wealth of genetic data expected in the near future coming through the discovery pipeline set up during the past years. This thesis contains genetic data on just two families with mendelian autosomal dominant inheritance of early onset CVD. To date, fifteen families have been phenotyped. Of these in four genotype data has been generated. Through these efforts we expect to find even more rare variants associated with CVD in the near future.
REFERENCES


