Rare genetic variants associated with early onset CVD
Maiwald, S.

Citation for published version (APA):
Maiwald, S. (2015). Rare genetic variants associated with early onset CVD

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.
Chapter 6

Genome Wide Association Studies in Atherosclerosis


Abstract
Cardiovascular disease remains the major cause of worldwide morbidity and mortality. Its pathophysiology is complex and multifactorial. Since the phenotype of cardiovascular disease often shows a marked heritable pattern, it is likely that genetic factors play an important role. In recent years large Genome Wide Association Studies have been conducted to decipher the molecular mechanisms underlying this heritable and prevalent phenotype. The emphasis of this review is on the recently identified 17 susceptibility loci for coronary artery disease. Implications of their discovery for biology and clinical medicine are discussed. A description of the landscape of human genetics in the near future in the context of next generation sequence technologies is provided at the conclusion of this review.
Introduction

Cardiovascular disease (CVD) remains the major cause of worldwide morbidity and mortality, despite the best of currently available therapy. Full appreciation of all risk factors and the pathogenesis of this disease are inevitably required to identify additional targets for therapy. Atherosclerosis, the major underlying cause of CVD, is present in all humans at advancing age and progresses over a lifetime, but its extent and progression is dependent on factors such as dyslipidaemia, hypertension, obesity, diabetes and smoking [1]. In addition, 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% [2]. This translates to the fact that apart from the classical risk factors, a positive family history for premature coronary artery disease (CVD) in a first degree relative also constitutes an additional and independent risk [3].

A large number of studies have aimed to decipher the molecular mechanisms underlying the heritability of CVD, however, in recent years the approach has seen major changes. In 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.

Later, a whole genome approach was applied to find genetic variations without a prior hypothesis. Initial attempts included classical linkage studies in families with premature CVD using microsatellite markers and later attempts single nucleotide polymorphism (SNP) arrays for linkage [4]. These first family based studies have thus far resulted in the identification of two interesting candidate genes, MEF2A and LRP6.

Through linkage analysis in an extended kindred that displayed an autosomal dominant pattern of CVD a large locus containing 93 genes was identified on 15q26 [5]. Within this locus, the gene for Myocyte Enhancer Factor 2 (MEF2A) was considered the best candidate for further investigation, because of its previously described effect on myocardial development and vascular morphogenesis [5]. Interestingly, a number of genetic variants in MEF2A were identified in other small cohorts of patients with CVD, suggesting that MEF2A may be actively involved in atherosclerosis [6]. Recently, however, additional studies have not revealed any evidence for a role of MEF2A as a susceptibility gene for CVD thereby refuting the original hypothesis [7,8]. Novel missense mutations in MEF2A were identified but in-vitro expression of the mutant proteins did not show a difference in transcriptional activity. Thus, the support for a role for MEF2A in CVD remains limited [7].
GWAS in Atherosclerosis

Linkage studies have also identified a locus on 12p13 in a kindred with the combination of extensive CVD and elevated plasma LDL-c levels and osteoporosis. The linkage interval contained 6 annotated genes including LDL-receptor Related Protein 6 (LRP6). The latter is an integral component of the Wnt signaling pathway, which is involved in vascular development and remodelling, but also plays a role in the regulation of glucose and lipid metabolism [9]. Common variants in LRP6 have been associated with LDL metabolism in humans [10,11].

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 (APOE:R2), mainly expressed in the brain [12,13]. The relative contribution of these loci to atherosclerosis in replication studies remains modest with typical odds ratios no larger than 1.5 [12].

In recent years, numerous genome wide association (GWA) studies with CVD as outcome have been published. These studies primarily focus on populations, rather than on families, and are carried out under the assumption that common variants are shared by subjects with a pre-specified phenotype. In this review we will provide an update of the loci identified in GWA studies for CVD. Of note, a total of 9 loci identified in the 6 GWA studies focusing on stroke have been poorly replicated and will not be discussed. Only the 9p21 locus has been found in all the stroke GWA studies. The other eight loci (2q33, 4q25, 8q12, 11q12, 12q13, 14q22, 16q22 and 22q13) show inconsistent results between the different studies and warrant further validation [14]. To date, no GWA study on peripheral artery disease (PAD) has been conducted. For these reasons this review will focus on loci discovered in GWA studies with CVD as outcome (Figure 1). In addition, we will discuss the possible current implications for biology and clinical medicine as a result of the discovery of these loci.
Proposed site of action for the proteins identified so far to be involved in the pathogenesis of atherosclerosis: liver, vessel wall, plaque and thrombus.

Apo: apolipoprotein; LDL: low-density lipoprotein; VLDL : very low-density lipoprotein; LDL-R: low-density lipoprotein receptor.

**GWA Studies**

The rationale behind GWA studies is the “common disease, common variant” hypothesis in which a limited number of genetic variants with a high frequency (typically above 5%) in the general population contribute to susceptibility for disease [15]. GWA studies became possible because of two major advances in genetics and technology. First, the identification of almost 10 million common SNP s with a minimum minor allele frequency of 5% in the hapmap project and the notion that a large number of these SNPs are in linkage disequilibrium (LD) [15,16]. Because of these LD blocks, a reduced number of SNPs provide complete whole genome coverage [16]. Second, the development of DNA microarrays for genotyping large numbers of SNPs and the great reduction in their price have facilitated the investigation of large cohorts of patients and controls. The first generation GWA studies used a microsatellite marker approach covering the genome by 400 unique microsatellite markers [17]. Currently, microarrays are available with more than a million SNPs. In addition, microarrays were developed based on a candidate gene approach, such as the IBC CardioChip. These have the advantage of a denser SNP profile per locus facilitating fine mapping, but the disadvantage that these often fail to find signals in loci not previously associated with the phenotype [18].
Major GWA Studies in CVD and Implications for Vascular Biology

Until recently GWA studies with CVD as the major outcome variable comprising more than 28,000 CVD cases and 65,000 controls have altogether identified 13 unique loci [19,20]. The recently published Global Lipids Consortium study in 24,607 cases with CVD and 66,197 controls added another four unique loci associated with CVD [21]. These 17 loci are summarised in table 1. As expected, the effect sizes are modest and in the same order of magnitude as the results obtained in GWA studies for other disorders.

The 9p21 locus was the first novel locus associated with CVD risk in a large case-control study in six independent samples compromising more than 23,000 participants from four Caucasian populations [22]. Two copies of the 9p21 risk alleles are present in about 25% of the population and convey a 60% increased risk for coronary artery disease as compared to the common genotype [22,23]. The minor risk alleles of SNPs in 9p21 region are found at a higher frequency in patients with early compared to late onset CVD [22,24-26]. Variations in the 9p21 locus might mediate their effect through modulation of platelet reactivity and vascular stiffness [27,28]. Additionally, the locus has also been found to be associated with other diseases associated with CVD such as stroke, abdominal aneurysms and type 2 diabetes [29]. But variants in this locus are also associated with ovarian cancer, acute lymphoblastic leukaemia, poor physical function in the elderly, glioma, malignant melanoma, breast cancer, and pancreas cancer [29], suggesting an interesting chromosomal area harbouring genes involved in a number of different physiological processes.

Although the chromosomal region of 9p21 contains a number of genes such as cyclin-dependent kinase inhibitor 2A and 2B (CDKN2A, CDKN2B), methylthioadenosine phosphorylase (MTAP) and ANRIL, a long non-coding RNA, the disease associated SNPs are located far away from these genes, which makes it more difficult to find the causal gene responsible for the observed signal in CVD [30]. CDKN2A, CDKN2B and MTAP are involved in the regulation of the cell cycle by interfering with cell cycle progression but additionally they are key players in physiological processes including replicative senescence, apoptosis and stem cell renewal. Direct sequencing of the chromosomal region did not reveal any causal variants that could explain the observed association. Interestingly, no association could be detected between the causal SNPs and differential expression of these genes [33]. CDKN2A, CDKN2B and MTAP are expressed in human vasculature and atherosclerotic plaques. In mice, deletion of the risk interval of 9p21 was shown to result in decreased expression of Cdkn2a and Cdkn2b [31]. Moreover, aortic smooth muscle cells harvested from mice deficient in 9p21 were found to have increased proliferative capacity [31].
## Table 1: Genetic loci identified in GWA studies which are associated with CVD.

<table>
<thead>
<tr>
<th>Loci</th>
<th>Chromosomal Localization</th>
<th>SNP</th>
<th>RAF (%)</th>
<th>Odds ratio per risk allele (95% CI)</th>
<th>Genes of interest</th>
<th>Putative Mechanism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1p13</td>
<td>rs599839</td>
<td>77</td>
<td>1.13 (1.08-1.19)</td>
<td>CELSR2, PSCR1, SORT1</td>
<td>LDL mediated</td>
<td>[20]</td>
</tr>
<tr>
<td>1p13</td>
<td>rs646776</td>
<td>81</td>
<td>1.19 (1.13-1.26)</td>
<td>CELSR2, PSCR1, SORT1</td>
<td>LDL mediated</td>
<td>[20]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2p32</td>
<td>rs11206510</td>
<td>81</td>
<td>1.15 (1.10-1.21)</td>
<td>PCSK9</td>
<td>LDL mediated</td>
<td>[24]</td>
</tr>
<tr>
<td>3</td>
<td>1q41</td>
<td>rs3008621</td>
<td>72</td>
<td>1.10 (1.04-1.17)</td>
<td>MA3</td>
<td>Unknown</td>
<td>[20]</td>
</tr>
<tr>
<td>1q41</td>
<td>rs17465637</td>
<td>72</td>
<td>1.14 (1.10-1.19)</td>
<td>MA3</td>
<td>Unknown</td>
<td>[24]</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2q33</td>
<td>rs6725887</td>
<td>14</td>
<td>1.17 (1.11-1.19)</td>
<td>WDRI2</td>
<td>Unknown</td>
<td>[24]</td>
</tr>
<tr>
<td>5</td>
<td>2q36</td>
<td>rs2972416</td>
<td>37</td>
<td>0.46</td>
<td>IRS1</td>
<td>Defective Insulin signaling and NO production</td>
<td>[21]</td>
</tr>
<tr>
<td>6</td>
<td>3q22</td>
<td>rs9818870</td>
<td>15</td>
<td>1.15 (1.11-1.19)</td>
<td>MRAS</td>
<td>Unknown</td>
<td>[49]</td>
</tr>
<tr>
<td>7</td>
<td>6p21.31</td>
<td>rs2814982</td>
<td>16</td>
<td>0.49</td>
<td>C6orf106</td>
<td>LDL mediated</td>
<td>[21]</td>
</tr>
<tr>
<td>8</td>
<td>6p24</td>
<td>rs12526453</td>
<td>65</td>
<td>1.12 (1.08-1.17)</td>
<td>PHACTR1</td>
<td>Unknown</td>
<td>[24]</td>
</tr>
<tr>
<td>9</td>
<td>6q26-6q27</td>
<td>rs2048327</td>
<td>18</td>
<td>1.20 (1.13-1.28)</td>
<td>LPA,LPA.L2, SLC22.A3</td>
<td>Coronary calcification</td>
<td>[60]</td>
</tr>
<tr>
<td>6q26-6q27</td>
<td>rs3127599</td>
<td>18</td>
<td>1.20 (1.13-1.28)</td>
<td>LPA,LPA.L2, SLC22.A3</td>
<td>Promotes atherothrombosis</td>
<td>[60]</td>
<td></td>
</tr>
<tr>
<td>6q26-6q27</td>
<td>rs7767084</td>
<td>18</td>
<td>1.20 (1.13-1.28)</td>
<td>LPA,LPA.L2, SLC22.A3</td>
<td>LDL mediated</td>
<td>[44]</td>
<td></td>
</tr>
<tr>
<td>6q26-6q27</td>
<td>rs1075578</td>
<td>18</td>
<td>1.20 (1.13-1.28)</td>
<td>LPA,LPA.L2, SLC22.A3</td>
<td>LDL mediated</td>
<td>[44]</td>
<td></td>
</tr>
<tr>
<td>6q26-6q27</td>
<td>rs3798220</td>
<td>2</td>
<td>1.47 (1.35-1.60)</td>
<td>LPA,LPA.L2, SLC22.A3</td>
<td>LDL mediated</td>
<td>[44]</td>
<td></td>
</tr>
<tr>
<td>6q26-6q27</td>
<td>rs10455872</td>
<td>7</td>
<td>1.68 (1.43-1.98)</td>
<td>LPA,LPA.L2, SLC22.A3</td>
<td>LDL mediated</td>
<td>[44]</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>7q32</td>
<td>rs4731702</td>
<td>48</td>
<td>0.59</td>
<td>KLF14</td>
<td></td>
<td>[21]</td>
</tr>
<tr>
<td>11</td>
<td>8p22</td>
<td>rs1495741</td>
<td>22</td>
<td>2.85</td>
<td>NAT2</td>
<td>LDL mediated</td>
<td>[21]</td>
</tr>
<tr>
<td>12</td>
<td>9p21</td>
<td>rs1333049</td>
<td>52</td>
<td>1.20 (1.16-1.25)</td>
<td>ANRIL, CDKN2A, CDKN2B, MTAP</td>
<td>Unknown</td>
<td>[20]</td>
</tr>
<tr>
<td>9p21</td>
<td>rs4977574</td>
<td>56</td>
<td>1.29 (1.25-1.34)</td>
<td>ANRIL, CDKN2A, CDKN2B, MTAP</td>
<td>Increased proliferation of smooth muscle cells</td>
<td>[24]</td>
<td></td>
</tr>
<tr>
<td>9p21</td>
<td>rs10757274</td>
<td>48</td>
<td>1.25 (1.15-1.36)</td>
<td>ANRIL, CDKN2A, CDKN2B, MTAP</td>
<td>LDL mediated</td>
<td>[61]</td>
<td></td>
</tr>
<tr>
<td>9p21</td>
<td>rs28383206</td>
<td>51</td>
<td>1.25 (1.15-1.36)</td>
<td>ANRIL, CDKN2A, CDKN2B, MTAP</td>
<td>LDL mediated</td>
<td>[61]</td>
<td></td>
</tr>
<tr>
<td>9p21</td>
<td>rs2383207</td>
<td>51</td>
<td>1.22 (1.13-1.33)</td>
<td>ANRIL, CDKN2A, CDKN2B, MTAP</td>
<td>LDL mediated</td>
<td>[61]</td>
<td></td>
</tr>
<tr>
<td>9p21</td>
<td>rs107572378</td>
<td>47</td>
<td>1.25 (1.15-1.36)</td>
<td>ANRIL, CDKN2A, CDKN2B, MTAP</td>
<td>LDL mediated</td>
<td>[61]</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>10q11</td>
<td>rs501120</td>
<td>84</td>
<td>1.11 (1.05-1.18)</td>
<td>CXCL12</td>
<td>Neointima formation after arterial injury, platelet activation in atherosclerotic lesions</td>
<td>[20]</td>
</tr>
<tr>
<td>10q11</td>
<td>rs1746048</td>
<td>84</td>
<td>1.17 (1.11-1.24)</td>
<td>CXCL12</td>
<td>LDL mediated</td>
<td>[20]</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>12q24</td>
<td>rs2259816</td>
<td>37</td>
<td>1.08 (1.05-1.11)</td>
<td>HNF1A</td>
<td>apoM mediated HDL modification</td>
<td>[49]</td>
</tr>
<tr>
<td>15</td>
<td>12q24</td>
<td>rs3184504</td>
<td>40</td>
<td>1.13 (1.11-1.19)</td>
<td>SH2B3</td>
<td>reduced anti-inflammatory activity</td>
<td>[56]</td>
</tr>
<tr>
<td>12q24</td>
<td>rs11065987</td>
<td>34</td>
<td>1.14 (1.10-1.19)</td>
<td>SH2B3</td>
<td>LDL mediated</td>
<td>[24]</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>16p13</td>
<td>rs1122608</td>
<td>75</td>
<td>1.15 (1.10-1.21)</td>
<td>KCNHE2, MRPS6, SLC5.A3</td>
<td>LDL mediated</td>
<td>[24]</td>
</tr>
<tr>
<td>17</td>
<td>21q22</td>
<td>rs9882601</td>
<td>13</td>
<td>1.20 (1.14-1.27)</td>
<td>LDL mediated</td>
<td>[24]</td>
<td></td>
</tr>
</tbody>
</table>

RAF: risk allele frequency; CI: confidence interval.
Another finding was that the 9p21 knockout mice gained significantly more weight than their control littermates [31]. An association between the 9p21 risk variant and body weight has not been observed in humans. Also, the burden of atherosclerosis measured by fatty lesion formation was similar in mutant and wild type animals.

Risk variants in the 9p21 locus associate with gene expression levels of ANRIL, a non-coding RNA can alter gene expression of nearby genes by RNA interference or gene silencing [32]. Whole blood RNA expression revealed a positive association between expression of ANRIL with CDKN2A and CDKN2B [33]. ANRIL is expressed in 8 different splice variants depending on the tissue type which complicates the analysis of the 9p21 locus [33]. In conclusion, the 9p21 locus illustrates the complexity of gene identification studies using GWA approach.

Another highly significant signal associated with both cholesterol levels and the occurrence of CVD is locus 1p31 [34]. Although several SNPs in this region were associated with these traits, fine mapping of the locus identified the rs12740374, located in the intergenic noncoding region, as the causative SNP [21]. Individuals who are homozygous carriers of the major allele have up to 16mg/dl (0.4 mmol/L) higher LDL-c plasma levels as well as a 40% increased risk of MI when compared to carriers homozygous for the minor allele [24,25]. Three genes: sortilin1 (SORT1), Cadherin (CELSR2) and proline-serine rich coiled coiled1 (PCSR1) are located in the 1p31 region. Tissue expression analysis reveals a strong association with hepatic SORT1 and P5RC1 expression. Rs12740374 creates a predicted binding site for C/EBPalpha transcription factors [34]. As a consequence, the hepatic gene expression of SORT1 will be altered. SORT1 encodes for the sortilin protein, also known as neurotensin receptor, a protein that functions as a multi ligand sorting receptor. Two studies have been published recently investigating the role of sortilin 1 in hepatic lipoprotein metabolism using different mouse models with conflicting results. In one study Sort1 deficient mice showed a reduction in the secretion of apoB100 containing lipoproteins from the liver and a reduction in atherosclerosis, whereas overexpression of Sort1 increased plasma cholesterol levels [35]. In the second study, a different approach using specific hepatic overexpression of Sort1 in the liver using adenovirus-mediated gene transfer in 'humanized mouse models revealed a marked decrease in plasma cholesterol levels by 73%, change in LDL size towards larger particles and consequently a reduction in atherosclerosis [34]. A second approach using siRNA technology revealed similar results. Currently, there is no explanation for the conflicting laboratory data observed.

In the GWA studies highly significant signals at 19p13 at the LDL receptor (LDLR) and 1p32 at the proprotein convertase subtilisin/kexin type 9 (PCSK9) were found. Both well-known
proteins play an important role in LDL metabolism and are strongly associated with atherosclerosis risk and demonstrate that GWA studies are indeed capable to detect genes that are related to CVD [36,37]. Rare variants in LDLR and PCSK9 result in the mendelian forms of familial hypercholesterolemia (FH). Clinically, FH manifests in lipid accumulation in tendons (xanthomas), the cornea (arcus lipoïdes) and arterial wall, resulting in accelerated atherosclerosis and premature CVD [38,39]. In untreated heterozygous FH patients, about 50% of males and 30% of females will develop CVD before the age of 60 years [40].

Another highly significant signal in GWA studies was on locus 6q25.3-26. This region harbours the gene Lipoprotein (a) (LPA), which was also previously shown to be implicated in atherosclerosis, but also solute carrier family 22 member 3 (SLC22A3) and lipoprotein(a)-like 2 (LPA-L2). Increased plasma Lp(a) levels are causally related to CVD [41]. Lp(a) is a cholesterol-rich LDL particle with one apoprotein(a) molecule attached to the apoB100 protein [42]. Due to homology of the repeated kringle structure of apo(a) with plasminogen, it has been suggested that Lp(a) intervenes with the process of atherosclerosis through involvement of fibrinolysis and as a consequence the atherothrombotic processes. Plasma levels of Lp(a) are strongly genetically determined by the number of kringle IV repeats in the apo(a) protein, which has an heritability between 21 and 27% [43]. SNPs identified by GWA studies are strongly associated to the kringle IV repeats and thus to plasma Lp(a) levels [44]. A recent meta-analysis revealed that plasma Lp(a) concentration is a stronger biomarker for risk than the gene score based on above mentioned SNPs [44].

Other loci that have been identified in the GWA studies are less frequently studied. In some cases there is suggestive evidence for a role in atherosclerosis. In the next section we will discuss the genes located in proximity of these identified loci.

The gene of interest on locus 10q11 is CXCL12/SDF1 (Chemokine (C-X-C-motif) ligand 12/stromal cell-derived factor-1). In patients with stable and unstable angina plasma levels of CXCL12 are lower compared to healthy controls [45]. Interestingly, CXCL12 is expressed in atherosclerotic lesions and is of importance for vascular repair and remodelling [45]. Furthermore it has a role in neo-intimal formation after injury [46]. Blocking its receptor Cxcr4 in Apoe−/− or Ldlr−/− mice aggravates diet induced atherosclerosis by affecting neutrophil recruitment machinery[47]. CXCL12 may promote atherothrombosis by its effect on platelets. It stimulates platelet adhesion to immobilized collagen and fibrinogen under flow, induces the release of various platelet chemokines in particular P-selectin. P-selectin plays an essential role in the initial recruitment of leukocytes to the site of injury during inflammatory processes such as
Atherosclerosis [48]. Additional evidence that genetic variation in CXCL12 is associated with CVD risk is eagerly awaited.

The locus on 3q22 encodes the Ras-related protein M-Ras (MRAS) gene. The lead SNP, rs9818870, is located in the 3' UTR of MRAS in close proximity to a cluster of miRNA binding sites[49]. MRAS is widely expressed in all tissues, with high expression in the cardiovascular system and brain [49,50]. MRAS is a member of the RAS family of small molecular weight GTP binding molecules and is involved in the PI3K-MAPK signaling pathways. The authors who reported the locus suggest a role for MRAS in adhesion signaling, which is important in the atherosclerotic process. No direct clues can be found in the literature to bolster this statement. Recently, rs9818870 was shown to be significantly associated with MRAS gene expression in aortic tissue of patients, confirming the previous findings [49,51]. Additional functional studies are lacking.

Phosphatase and actin regulator 1 (PHACTR1) is an inhibitor of protein phosphatase 1 (PP1), an enzyme that dephosphorylates serine and threonine residues on a range of proteins and resides near 6q26-27 [52]. PP1 is an ubiquitously expressed protein and involved in a number of physiological processes such as Protein Kinase-A (PKA) and Protein Kinase C (PKC) phosphorylation [52]. It is interesting, that in an independent GWAS for coronary artery calcification in >10,000 participants from six prospective cohort studies, PHACTR1 SNPs (along with chromosome 9p21 SNPs) were associated with coronary artery calcification at genome-wide significance (C.J. O'Donnell, NHBL Institute, personal communication).

Rare variants in hepatocyte nuclear factor 1 homeobox A (HNF1A) on 12q24 locus result in mendelian forms of Maturity Onset Diabetes of the Young (MODY) [53]. In a recent GWA study searching for loci predictive for type 2 diabetes, rs7957197 in HNF1A, reached genome wide significance. HNF1A is transcription factor involved in expression of genes involved in innate immunity, blood coagulation and lipid and glucose metabolism[54].

For Insulin Receptor Substrate 1 gene (IRS1) and SH2B adaptor protein 3 (SH2B3) the evidence for a role in atherosclerosis is much weaker. The rs2972146 SNP on chromosome 2 is close to IRS1, a gene that has been shown to be essential in insulin signaling. Genetic variants in IRS1 impaired insulin signaling in endothelial cells. As a consequence nitric oxide (NO) release by these cells is decreased [55]. However, a relation between endothelial dysfunction and genetic variants of IRS1 has not been published so far. SH2B3 located near 12q24 is expressed in human vascular endothelial cells, where it promotes inflammation. The lead SNP, rs3184504 which results in a missense change, could contribute to the progression of plaque formation in
coronary arteries leading to myocardial infarction by its induction of a pro-inflammatory state. 

*SH2B3* has previously been shown to be associated with type 1 diabetes and celiac disease [56]. Finally, more evidence has to be generated to further identify the suspected candidate genes at the following loci and their association with atherosclerosis: 1q41 locus: melanoma inhibitory activity family member 3 (*MIA3*), 2q33 locus: WD repeat domain 12 (*WDR12*), 7q32 locus: Kruppel-like factor 14 (*KLF14*), 8p32 locus: N-acetyltransferase 2 (*NAT2*) and 21q22 locus: Potassium voltage-gated channel subfamily H member (*KCNH2*), [57].

### Implications for clinical medicine

The identification of the current 17 common variants associated with CVD risk has so far little impact on routine clinical medicine. However, unravelling the biological role of the proteins in the process of atherosclerosis, as has been done for the 1p13 locus, will probably lead to a better understanding of the pathways involved in the pathophysiology of CVD. In addition, these pathways might be targets for development of future therapeutic compounds.

The findings so far cannot be directly translated into a risk profile for an individual patient. In a risk score comprising nine GWA study SNPs associated with early onset MI, including a 9p21 SNPs and three SNPs associated with LDL-c, a two-fold difference in risk for MI between extreme quintiles of scores was found [24]. However, in a prospective cohort there is no improvement in risk discrimination by addition of the risk alleles of 9p21 in men or women [58,59].

### Conclusions

In the last years, efforts to disentangle the genetics of CVD based on GWA studies resulted in the identification of several common variants with moderate effects. This has resulted in the identification of 17 loci. Still, only part of the heritability of CVD is currently explained. More of the missing heritability could be discovered by efforts to identify novel SNPs and efforts to fine map candidate loci [61]. Identification of novel SNPs will occur through future GWA studies and meta-analyses of the current studies. At least three large GWA studies (IBC 50K CVD, Cardiogram and C4D) on CVD are near completion and fine mapping of existing candidate loci is on going.

Nevertheless, even with the outcome of these efforts only a modest part of the risk for CVD will be explained. 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 *de novo* mutations, as recently found in mental retardation [62]. 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.

**Acknowledgement**

S. Maiwald received financial support from NETSIM Marie Curie Initial Training Network, EC (Project Number 215820). M.M. Motazacker was supported by a grant of the European Community (FP6-2005-LIFESCIHEALTH-6; STREP contract number 037631).
Reference list


Excellent state of the art review on the current status of the genetics of CVD.


Combination of GWAS meta analysis and additionally gene expression data to proof the importance for TG metabolism. A novel gene GALNT2 has been describes in more detail in a novel mouse model.


One of the first GWA studies on coronary artery disease.


This paper shows the approach to identify a GWA signal and translate it into a functional mechanism using mouse models, in vitro studies but also large cohorts to identify the QTL signal. All together it enables to understand the role of SORT1 in lipid metabolism and atherosclerosis.


[41] Kamstrup PR. Lipoprotein(a) and ischemic heart disease--a causal association? A review. Atherosclerosis 2010;211:15-23.


