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.
A Rare Variant in \textit{MCF2L} Identified using Exclusion Linkage in a Pedigree with Premature Atherosclerosis


Submitted
Abstract

Background
Cardiovascular disease (CVD) is a major cause of worldwide death. CVD risk is largely genetically determined. The molecular pathology is, however, not elucidated in a large number of families suffering from CVD.

Methods
We applied exclusion linkage analysis and next generation sequencing to elucidate the molecular defect underlying premature CVD in a small pedigree, comprising two generations of which 6 members suffered from premature CVD.

Results
A total of three variants showed co-segregation with the disease status in the family. Two of these variants were excluded from further analysis based on the prevalence in replication cohorts, whereas a non-synonymous variant in MCF.2 cell line derived transforming sequence-like protein (MCF2L, c.2066 A>G; p.(Asp689Gly); NM_001112732.1), located in the DH-domain, was only present in the studied family. MCF2L is a guanine-nucleotide exchange factor that potentially links pathways that signal through Rac1 and RhoA. Indeed, in HeLa cells MCF2L-689Gly failed to activate Rac1 as well as RhoA resulting in impaired stress fiber formation. Moreover, MCF2L protein was expressed in human atherosclerotic lesions but not in healthy tissue segments.

Conclusion
A rare functional variant in MCF2L was identified in a small pedigree with premature CVD. The presence of MCF2L in human atherosclerotic plaque specimen lends further support to its potential role in atherosclerosis.
Introduction

Cardiovascular Disease (CVD) is the major cause of morbidity and mortality across the globe. In combination with the strong influence of environmental factors, genetic determinants are pivotal in the pathobiology of atherosclerosis, the underlying disease ultimately resulting in CVD events [1]. This was demonstrated by a 30-60% heritability in twin studies, and the finding that a positive family history of premature CVD is an independent risk factor [2,3]. Forty-six common variants with a minor allele frequency > 5% were found to be associated with CVD in genome wide association studies in recent years [4,5]. Collectively these data showed that only 10% of CVD risk is attributable to common genetic variations, which supports the hypothesis that low frequency variants may help to explain the observed heritability [4-6]. Thus, the complex heritability of CVD may be explained by a combination of common variants with effect sizes so small that they remain undetected in recent GWAS meta-analyses or by rare variants with an intermediate or large effect. Multi-generation pedigrees with a clear Mendelian form of premature CVD will be extremely instrumental in the identification of these novel pedigree-specific mutations with a large effect on CVD. So far, mutations in Low Density Lipoprotein receptor (LDLR), Low-Density Lipoprotein Receptor Related Protein 6 (LRP6) and Myocyte Enhancer Factor 2A (MEF2A), keratocan (KERA), guanylyl cyclase α1 subunit (GUCY1A3) and Chaperonin Containing TCP1, Subunit 7CCT7 have been reported, although conflicting results have been reported for MEF2A [7-15]. Collectively, the identification of informative pedigrees with a dominant form of premature CVD have proven to be instrumental in the identification of novel variants in genes involved in the pathobiology of CVD, which may ultimately result in the new therapeutic targets.

The aim of the current study was to identify the molecular defect in a small pedigree of Caucasian descent with an autosomal dominant form of CVD. By combining data from exclusion linkage analysis and next generation sequencing (NGS) we identified a non-synonymous variant located in the DH-domain MCF2L a gene encoding for MCF.2 cell line derived transforming sequence-like protein also known as the guanine nucleotide exchange factor DBS.

Methods

Recruitment of the Pedigree with Early Onset CVD

A female subject was referred to the outpatient clinic of the Academic Medical Center (Amsterdam, the Netherlands) for analysis after she suffered from an acute myocardial infarction (AMI) at the age of 39 years (Figure 1; index II.6). She was member of a small family with an
A Rare Variant in MCF2L Identified using Exclusion Linkage in a Pedigree with PAS

 autosomal dominant form of early onset CVD. A premature CVD event was defined according to the AHA/ACC criteria as having a documented CVD event before the age of 55 years (male) and 65 years (women) (Figure 1) [16]. The affection status was assessed based on medical records and imaging as extensively described in the Supplemental Methods (Table S1). The study complies with the declaration of Helsinki and the Institutional Review Board of the AMC of the University of Amsterdam approved the study and all participants provided written informed consent.

Blood was collected from the index and her relatives, after an overnight fast, in EDTA-coated tubes. Plasma was isolated after centrifugation at 1600xg, 4°C for 20 minutes and stored at -80°C for further analyses. Plasma cholesterol, low-density lipoprotein-cholesterol (LDL-c), high-density lipoprotein cholesterol (HDL-c) and triglycerides (TG) were analyzed using commercially available assays (Randox, Antrim, United Kingdom and Wako, Neuss, Germany) on a Cobas-Mira autoanalyzer (Roche, Basel, Switzerland).

Exclusion linkage analysis and exome sequencing

Genomic DNA was extracted from whole blood on an Autopure LS system according to the manufacturer’s protocol (Gentra, Systems, Minneapolis, USA). Human CytoSNP-12 DNA analysis BeadChip kits (Illumina, Eindhoven, The Netherlands) were used for genome wide single nucleotide polymorphism genotyping in 5 affected and 2 unaffected relatives (Figure 1). The data quality was checked using Graphical Representation of Relationship errors (GRR) and PedCheck packages [17 18]. Multipoint parametric linkage analyses was performed using a fully penetrant autosomal dominant model and mutant locus frequency (MAF) of 0.001 using the Allegro program including 110,000 high-quality SNPs throughout the genome [19]. The maximum theoretical LOD score for linkage analysis in the core pedigree was 1.5. In parallel, exome sequencing using the DNA of the index case was performed using the Agilent SureSelect 38Mb exome on the Illumina GAII platform (Illumina, Little Chesterford, United Kingdom) and used for exclusion linkage.[20] Three missense mutations in respectively MCF2L (NM_01112732.1), ZC3HC1 (NM_016478.3) and CAMSAP1 (NM_015447.2) were identified and confirmed with Sanger sequencing as previously described using the following primer pairs: MCF2L forward 5’-TGC TTT TGC TTT GAT GGA TG-3’ and reverse 5’-CAT TCC AGC CCC CTG AAG-3’; ZC3HC1 forward: 5’-GAG AAA ACT CTC TTT TTC ATT CC-3’ and reverse 5’-CAC CCA AAT AAG CTA AGT GAA TAC-3’; CAMSAP1 5’-AAA CAG ATG CTA CCA ATC CCT TAC-3’ and reverse 5’-CCT CTT CCA AAG ATG CCA AC-3’.[21] The
data are registered on the LOVD database under screening number 00027156 (http://database.lovd.nl/shared/screenings/0000027156).

Validation cohorts

Premature AtheroSclerosis (PAS) Cohort: this cohort comprises 935 patients with early symptomatic atherosclerosis (CVD) before the age of 51 years. CAD was defined as myocardial infarction, coronary revascularization, or evidence of at least 70% stenosis in a major epicardial artery [22]. Patients were recruited at the cardiology and vascular outpatient clinics of the AMC, Amsterdam, The Netherlands [23].

Sanquin Blood Bank common Controls: DNA samples from 1,440 healthy volunteers were recruited from a large cohort of blood donors, who were free of CVD, at one of the collection sites of the Sanquin Blood Bank covering the northwest section of the Netherlands, which geographically overlaps the PAS case cohort [23].

Cambridge Bioresource Collection: NHS Blood and Transplant enrolled DNA samples of 8,946 healthy volunteers in a resource for genotype-phenotype association studies [24].

MCF2L Constructs and Cell Transfections

A vector containing human MCF2L (pENTR221™/MCF2L; clone IOH23111, Invitrogen, Bleiswijk, The Netherlands) was used. A human MCF2L-Asp689Gly vector was generated by site-directed mutagenesis with the Quick QuikChange kit (Stratagen) using the primer pairs: forward 5'-CGC CGC GGA GAT GG TAA CCC ACT GAT GG-3' and reverse 5'-GCG GCC CCT CTA CCC ATT GGG TGA CTA CC-3' (QuikChange primer design tool (https://www.genomics.agilent.com) [25]. Sequences were checked with Sanger Sequencing using the following primers: M13 forward 5'-GTT GTA AAA CGA CGG CCA GT-3' and reverse 5'-CAC AGG AAA CAG CTA TGA CC-3'. Next, the wild type and mutant constructs were inserted into the destination vector pcDNA™-DEST40 (Invitrogen) using the Gateway® LR Clonase II® enzyme mix (Invitrogen) according to the manufacture’s protocol.

Rac1-GTP pull down assay

HeLa cells were cultured in Iscove’s Modified Dulbecco’s medium (IMDM; Invitrogen) supplemented with 10% (v/v) heat-inactivated fetal calf serum, 1% glutamine and antibiotics and kept at 37 °C at 5% CO₂ as previously described [26]. Cells were transfected with pcDNA™-DEST40/MCF2L or pcDNA™-DEST40/MCF2L-D689G using TransIT®-LT1 Transfection Reagent (Mirus, Ochten, The Netherlands). After 24h incubation, cells were lysed in RIPA and subsequently centrifugation at 14,000 g for 5 min. GTP-bound Rac1 (Ras-related C3 botulinum
toxin substrate 1) was isolated with biotinylated CRIB-peptide coupled to streptavidine agarose beads during a 30 min. incubation at 4°C [27]. Beads were washed 4 times in 50mM Tris, pH 7.4, 0.5 mM MgCl₂, 150 mM NaCl, 1% (v/v) Triton X-100, supplemented with protease and phosphatase inhibitor cocktails (Roche, Woerden, The Netherlands). Rac1 was visualized by Western blotting using a mouse-anti-human Rac1 antibody (cone 102, BD bioscience, Breda, The Netherlands).

To image stress fibers, HeLa cells were plated on fibronectin- (FN) (R&D systems Europe, Abingdon, United Kingdom) coated glass cover slips and transfected as described above. Next, cells were fixed with 3.7% formaldehyde and permeabilized with 0.5% Triton X-100, stained with rabbit-anti-human MCF2L, and visualized with Phalloidin-Texas Red (Invitrogen). F–actin was made visible with ALEXA488-labeled anti-rabbit IgG (Invitrogen). The nuclei were stained with DAPI (Invitrogen) [28]. Images were recorded with LSM510 META microscope (Carl Zeiss MicroImaging, Inc., Germany).

**Human plaque quantification**

Specimens of tonsil (n = 2), which is used as a negative control tissue, atherosclerotic artery specimen (n = 9), and non-diseased artery specimens (n = 9) were collected from patients at autopsy. Coronary arteries were classified according to different stages of atherosclerosis using hematoxylin (Fluka Biochemica, Buchs, Switzerland) and eosin (Sigma Aldrich, St. Louis, MO, USA) stains according to standard protocols. Tissue specimen were stained with antibodies against MCF2L (rabbit polyclonal anti-MCF2L, WH0023263M1, Sigma, Munich, Germany), smooth muscle cell α-actin (SMA) (mouse 1A4 antibody, DAKO, Glostrup, Denmark), macrophage CD68 (mouse monoclonal anti human CD68 PG-M1, DAKO, Heverlee, Belgium), endothelial cell marker CD34 (mouse anti human CD34Q, Bend10, ThermoFischer Scientific, Waltham, MA, USA), CD11c as marker for antigen presenting cells (mouse monoclonal 5D11; Monosan, Uden, The Netherlands), CD8 as marker for cytotoxic T cells (mouse monoclonal anti-human CD8, CD8/144B; DAKO), and CD3 as marker for type 1 helper T-cells (anti-CD3, IgG monoclonal SP7; ThermoFischer Scientific). Secondary antibodies: swine polyclonal anti rabbit immunoglobulins/HRP 1:3000 (DAKO) and goat polyclonal anti mouse immunoglobulins/HRP 1:1000 (DAKO). Additional information is available in the Supplemental Method section.
**Statistical analysis**

Results are expressed as mean ± standard deviation unless otherwise stated. Statistical differences were tested using Students t-Test statistics. Two-sided p-values of < 0.05 were considered statistically significant. All statistical analyses were performed with SPSS version 18.0.

**Results**

**Identification of the mutation in MCF2L (Figure 1)**

The index case suffered from an AMI at the age of 39 years, who had been smoking 20 cigarettes/day before she had the event. Plasma lipids were all within the normal range (Table 1). Upon expansion of the family it was noted that a total of five relatives did suffer from premature CVD, whereas 2 relatives were classified as “unaffected” according to the definition presented in Supplemental Methods (Figure 1A, Table S1).

<table>
<thead>
<tr>
<th>Table 1: Characteristics of The 7 Relatives Used in Exclusion Linkage.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No CVD</strong></td>
</tr>
<tr>
<td>MCF2L mutation carrier</td>
</tr>
<tr>
<td>Sex (N = male/female)</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Age CVD</td>
</tr>
<tr>
<td>Smoking (N = Yes)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
</tr>
<tr>
<td>Total cholesterol (mmol/l)</td>
</tr>
<tr>
<td>HDL cholesterol (mmol/l)</td>
</tr>
<tr>
<td>LDL cholesterol (mmol/l)</td>
</tr>
<tr>
<td>Triglyceride (mmol/l)</td>
</tr>
</tbody>
</table>

Data are expressed as mean ± standard deviation or as number (N). Lipid values are expressed as median with interquartile range (IQR). Subjects were considered smokers if they were current smokers or when they quitted smoking within the last 5 years. N/A = not applicable; BMI = body mass index; LDL = low-density lipoprotein; HDL = high density lipoprotein. * = Pedigree member I-1 was not included since no clinical data was available.

Multipoint linkage analysis for exclusion linkage was performed (Figure S1), followed by exome sequencing using DNA from the index case. A total 453,100 variants were identified that met the
commonly described quality criteria of which 4,510 were non-synonymous, indels or splice site variants [29]. Of these, 961 variants were not found in the large public database (dbSNP129), 8 Hapmap project exomes [30], and the 1000 Genome pilot project SNPs and indels [31].
Figure 1: The identification of a mutation in MCF2L.

A. Co-segregation of the MCF2L variant in a small pedigree with premature CVD. Squares represent males and circles represent females. Right half-filled symbols represent cases with premature CVD. The arrow indicated the index case. Left upper forth-part-filled symbols represent MCF2L mutation carrier. For each individual the Arabic number is an identifier within their generation, whereas the generation is marked with a roman number. DNA of all living 12 individuals was included in the genetic analysis. B. Schematic overview of the gene finding strategy. C. DNA Sanger sequencing chromatogram showing the heterozygote c.2066 A > G; p.(Asp689Gly). D. Schematic overview of the position of the identified heterozygote variant in MCF2L on gene level.

The vast majority of these variants (899 = 93%) were located within the genomic regions with LOD scores below -2 and were excluded from further analysis. From the remaining 62 variants (Table S2) 15 were predicted to have deleterious effects based on SIFT or Polyphen prediction software [32 33], from which 6 variants were proven to be false positive by Sanger sequencing, 3 variant were not located near s reported gene and from the remaining 6 variants only 3 were confirmed by Sanger sequencing to co-segregate with the phenotype within the pedigree: p.(Ser1216Leu) (c.3647C>T; rs149524209) in calmodulin regulated spectrin-associated protein 1 (CAMSAP1), p.(Ile305Val) (c.913A>G; rs139253452) in nuclear-interacting partner of anaplastic lymphoma kinase (ZC3HC1) and p.(Asp689Gly) (c.2066A>G) in MCF.2 cell line derived
A Rare Variant in MCF2L Identified using Exclusion Linkage in a Pedigree with PAS

transforming sequence-like (MCF2L, NM_001112732.1). Next we tested whether the 3 variants were present in different replication cohorts (Table S3). The CAMSAP1 variant was found in 13 healthy volunteers of the Dutch Sanquin Blood Bank cohort (0.9%) and the ZC3HC1 variant was found in 6 individuals of the PAS cohort (0.6%), in 3 healthy volunteers of the Sanquin Cohort (0.2%) and in 5 volunteers of the Cambridge Biobank Resource (0.06% (Ras homolog gene family, member A); (Table S4). Thus, both variants were eliminated for further studies. MCF2L p.(Asp689Gly) variant was not found in any of the replication cohorts. MCF2L is located on chromosome 13q34 and harbours a highly conserved Rho-GEF Dbl-homology (DH) domain encoding the GEF activity specific for Rho family members and a pleckstrin homology domain (PH) that binds inositol phosphates (Figure 1D). The p.(Asp689Gly) variant is located in a highly conserved region of the DH-domain. MCF2L is a guanine-nucleotide exchange factor from the Rho family that are signaling molecules of Rho proteins. MCF2L is ubiquitously expressed in circulating blood cells. Total white blood cell MCF2L mRNA expression in mutation carriers was lower as compared to the control (Figure S2; p < 0.05). Noteworthy, mRNA expression of Ras- related C3 botulinum toxin substrate 1 (RAC1) and Ras homolog gene family, member A (RHOA) was decreased, but this did not reach statistical significance (Figure S2).

**MCF2L-689Gly variant is impaired in activating the Rac1 pathway**

As a guanine-nucleotide exchange factor (GEF), MCF2L has been described to activate the small GTPases Rac1 and RhoA [34]. We used a classical Rac1 pull down assay to assess the ability of MCF2L to activate Rac1. HeLa cells were transfected with either wild-type MCF2L-689Asp or mutant MCF2L-689Gly and Rac1.GTP was measured (Figure 2). MCF2L expression resulted in an activation of Rac1, whereas the mutant MCF2L-689Gly failed to activate Rac1. Next to RAC1, MCF2L may also activate RhoA, which is specifically involved in the formation of stress fibers by bundling the F-actin filaments. In the next experiment we tested whether MCF2L expression resulted in RhoA activation by studying stress fiber formation. Indeed expression of wild-type MCF2L in HeLa cells resulted in increased stress fiber formation and cell rounding, both characteristics of RhoA activation (Figure 3). In contrast, MCF2L-Asp689Gly expression failed to induce the formation of stress fibers or cell rounding (Figure 3). Together, these experiments show that MCF2L-Asp689Gly is impaired in activating both Rac1 and RhoA.

Mutant MCF2L-689Gly does not activate the small GTPase Rac1. Hela cells were transiently transfected with wild-type or mutant MCF2L constructs. The expression of Rac1 was measured in cell-lysates. Rac1-GTP was measured as described in methods section. pcDNA™-
DEST40/MCF2L-689Asp but not pcDNA™-DEST40/MCF2L-689Gly activates endogenous Rac1. Second panel from above shows Rac1 protein loading in the cell lysates.
A Rare Variant in MCF2L Identified using Exclusion Linkage in a Pedigree with PAS

Figure 2: MCF2L-Asp689Gly does not activate Rac1.

![Image of Western Blot](image)

Mutant MCF2L-689Gly does not activate Rac1. HeLa cells were transiently transfected with wild-type or mutant MCF2L constructs. The expression of Rac1 was measured in cell-lysates. Rac1-GTP was measured as described in methods section. MCF2L-689Asp but not MCF2L-689Gly activates endogenous Rac1. Beta-actin was used as a loading control.

Figure 3: MCF2L-689Asp but not mutant MCF2L-689Gly induces the RhoA phenotype.

![Image of Immunofluorescence](image)

Wild-type and mutant versions of MCF2L were expressed in HeLa cells. Immunofluorescence imaging showed that MCF2L-689Asp gives a RhoA phenotype, meaning round cells and pronounced stress fibers (F-Actin, green, white square). MCF2L-689Gly did not show an effect on the cell phenotype. The cells stay flat and do not show stress fibers (F-actin, green). F-Actin expression and MCF2L wild-type are co-localized as shown in the merged images (DAPI, blue). ROI=region of interest.
**MCF2L is present in atherosclerotic segments in humans**

MCF2L was profoundly expressed in human coronary artery segments from patients with different grades of atherosclerosis (Figure 4A) but not in non-diseased arteries (Figure 4B). Human tonsil tissue served as a positive control (Figure S3A). Staining without the secondary antibody in either atherosclerotic lesions or tonsil tissue was negative (Figure 4C and Supplementary Figure 3B). MCF2L did not co-localize with CD68, a marker for macrophages, CD34 (endothelial cell marker) or smooth muscle cells actin (SMA) (Figures 4D - F) but was co-localizing with antigen presenting cells such as CD11c, CD3 type I helper cells and CD8 cytotoxic T-cells within the atherosclerotic plaque (Figures 4G - 2I).

**Figure 4: MCF2L is expressed in atherosclerotic but not in non-diseased coronary arterial segments and co-localized with CD3 and CD8 cells.**

A. Single staining of MCF2L (brown) in human atherosclerotic tissue. B. Absence of MCF2L in a normal healthy coronary artery tissue segments. C. Negative control in atherosclerotic tissue using only the secondary antibody No background signal is observed. D. No co-localization of MCF2L (blue) with CD68 (red); E. No co-localization with smooth muscle alpha-actin (SMA
Discussion

The aim of the current study was to identify the genetic defect underlying premature CVD in a small pedigree with an autosomal dominant form of CVD. We identified a rare variant in MCF2L at position 689 in the DH-domain of the protein where an aspartic acid was changed to a glycine. The variant was neither present in 10,376 healthy controls nor in 935 patients with premature CVD. The MCF2L-p.(Asp689Gly) variant is located in the DH-domain. Interestingly, MCF2L was absent in healthy artery segments, but abundantly expressed in atherosclerotic lesions and co-localized with CD3 type I helper T-cells and CD8 cytotoxic T-cells.

MCF2L belongs to the Dbl family, which are guanine-nucleotide exchange factors (GEFs) in particular Rho guanosine triphosphatase (Rho-GTPases). MCF2L contains a Dbl homology-(DH-), a pleckstrin homology- (PH-) domain and a CRAL-TRIO domain (Figure 1D) [35]. The DH-domain is crucial for MCF2L function. The specific function of the PH-domain is largely unknown, but initial studies suggest a supportive function for the binding to the plasma membrane [34]. MCF2L provides the shift of GDP to GTP, resulting in the activation of Rho-GTPases. It is the MCF2L-DH-domain that has been shown to play an important role in the interaction with and activation of the small GTPases RhoA, Rac1 and Cdc42 [36,37]. RhoA functions as a molecular switch to regulate the assembly of actin stress fibers [38,39]. Here we showed that the MCF2L-689Gly mutant protein was defective in activating Rac1 and consequently no actin stress fibers were formed, which is highly indicative for an impaired RhoA activation of the MCF2L variant. These defects may result in increased permeability and perturbed cell migration in the presence of the mutant MCF2L protein. Noteworthy, abnormal activity of Rho-family GTPases has been reported in cardiovascular disorders [40]. Endothelial dysfunction reflected by reduced vasodilation is a critical factor contributing to the pathogenesis of cardiovascular disease.

GEFs together with Rho kinases play a pivotal role in leukocyte migration [41]. The first step of leukocyte attachment involves rolling of the leukocytes over the endothelium mediated by different selectins. The subsequent clustering to adhesion molecules like ICAM-1 and VCAM-1 induces intracellular signaling of the endothelium involving RhoA and Rac1 activation. Although it is unclear if abnormal RhoA or Rac1 activation, as observed in the presence of the MCF2L mutant protein, may interfere with leukocyte transendothelial migration, we observed that RAC1
and RHOA mRNA expressions in circulating blood cells was altered in MCF2L mutation carriers as compared to healthy controls. Furthermore we observed increased MCF2L protein expression in atherosclerotic tissue segments that co-localized with CD3 and CD8, markers for T-cells. Further studies are required to study to what extent circulating cell behaviour is underlying the pathogenesis that is observed in this family.

The application of a combined exclusion linkage analysis and next generation sequencing approach has been proven instrumental to identify novel mutations [20]. In the past, due to technical constraints, most of the pedigree studies have focused on identification of molecular defects in pathways with an established role in atherosclerosis. This resulted in the identification of MEF2A and LRP6 [7-10]. Next generation sequencing has provided us with an opportunity to sequence large numbers of genes in an unbiased approach without knowledge of the underlying biological process. This unbiased approach is likely to give a great impetus to our understanding of the pathology of complex diseases such as CVD. The identification of the novel mutation in MCF2L has proven that our approach is a successful method to identify rare variants in smaller families.

**Limitations**

A major limitation of the current study is, however, that only circumstantial evidence for the role of this variant in the CVD phenotype could be gathered. With the use of exclusion linkage we originally identified 3 rare variants in 3 different genes, but only one of them was uniquely present in the family. We also did not identify this variant in other cases with premature atherosclerosis.

However, the presence of MCF2L protein in diseased tissue and the observation that the variant led to impaired MCF2L-DH-domain dependent stress fibres formation does suggest that MCF2L may play a role in the underlying pathobiology of premature atherosclerosis.

In conclusion, we have identified a functional variant in MCF2L using exclusion linkage analysis and next generation sequencing in a pedigree with premature CVD. Further studies are warranted to confirm our findings and to establish whether MCF2L may be an interesting target for therapy.

**Acknowledgements**

We would like to acknowledge all participants. We thank C. Holtkamp for the genetic fieldwork and J. Sambrook, C. Aalbers, S. Nurnberg and J. Peter for technical and laboratory assistance.
We gratefully acknowledge the participation of all NIHR Cambridge BioResource volunteers. We thank the Cambridge BioResource staff for their help with volunteer recruitment.

**Contributors**

All authors listed on the manuscript fulfill the criteria for authorship. Specific contributors are SM: experiments, data analysis, writing of the manuscript; MMM: genetic analysis; JC, SS, patient clinical diagnosis, study sample providers. WHO: genetic replication; JJK, GKH, MDT: clinical data, manuscript reviewer; JB: experimental supervisor, manuscript reviewer; GD data analysis and writing manuscript.

**Funding sources**

SM has been supported by the Marie-Curie NetSim ITN grant [grant number EC-215820] from the European Commission 7th Framework Program and Ipse Movet. SS is supported by a grant from ATHEROS. MMM was supported by Fondation LeDucq (Transatlantic Network, 2009-2014). JJPK is a recipient of the Lifetime Achievement Award of the Dutch Heart Foundation (2010T082). Research in WHO group is supported by program grants from the National Institute for Health Research (NIHR) (RP-PG-0310-1002), from the British Heart Foundation [RG/09/12/28096], from the European Commission and the Medical Research Council. The project made use of NHS Blood and Transplant donors from the Cambridge BioResource (http://www.cambridgebioresource.org.uk/). This local resource for genotype-phenotype association studies is supported by a grant from the NIHR to the Cambridge Biomedical Research Centre. GKH is holder of a Veni grant [91612122] from the Netherlands Organisation for Scientific Research (NWO). This work is supported by CardioVascular Research Initiative [CVON2011-19; Genius] and the European Union [Resolve: FP7-305707 and TransCard: FP7-603091-2].

**Conflict of interest**

We declare that none of the authors has any disclosures to report for the current manuscript.
References


A Rare Variant in MCF2L Identified using Exclusion Linkage in a Pedigree with PAS


[41] Fernandez-Borja M, van Buul JD, Hordijk PL. The regulation of leucocyte transendothelial migration by endothelial signaling events. Cardiovasc Res 2010;86;202-10
Supplementary Methods
Patient selection

CVD was defined by the presence of at least one of the following features: (I) myocardial infarction, proven by at least two of the following criteria: (a) classical symptoms (chest pain that may radiate, oppressive pain, nausea, sweating and absence of chest-wall tenderness on palpation), (b) specific electrocardiographic abnormalities, (c) elevated cardiac enzymes (e.g., troponin and elevated creatine-kinase (CK) and its myocardial band enzyme (CK-MB), levels); (II) percutaneous coronary intervention or other invasive procedures; (III) coronary artery bypass grafting; (IV) angina pectoris, diagnosed as classical symptoms (recurrent attacks of retrosternal pain brought on by effort and emotion and relieved by rest and the administration of nitroglycerin) in combination with at least one unequivocal result of one of the following: (a) exercise test, (b) nuclear scintigram, (c) dobutamine stress ultrasound, (d) a more than 70% stenosis on a coronary angiogram or (f) requiring treatment (V) ischemic stroke, demonstrated by CT- or MRI scan; (VI) documented transient ischemic attack. Blood pressure was measured using an oscillometric blood pressure device (Omron, Hoofddorp, The Netherlands).

Figure S1: The results of linkage analysis using Allegro software.

![Linkage Analysis](image)
Table S1: Clinical characteristics of the relatives.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Sex</th>
<th>Type of CVD</th>
<th>Age CVD</th>
<th>Medication</th>
<th>MCF2L Asp689Gly</th>
<th>Hypertension (years)</th>
<th>Smoking</th>
<th>BMI (kg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-1</td>
<td>F</td>
<td>None</td>
<td>N/A</td>
<td>None</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>23</td>
</tr>
<tr>
<td>II:3</td>
<td>F</td>
<td>None</td>
<td>N/A</td>
<td>None</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>22</td>
</tr>
<tr>
<td>II:1</td>
<td>F</td>
<td>PTCA</td>
<td>62</td>
<td>Lipitor 40 mg</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>22</td>
</tr>
<tr>
<td>II:2</td>
<td>M</td>
<td>AP</td>
<td>46</td>
<td>Inegy 40 mg</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>25</td>
</tr>
<tr>
<td>II:4†</td>
<td>M</td>
<td>AMI</td>
<td>43</td>
<td>Lipitor 40 mg</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>23</td>
</tr>
<tr>
<td>II:5</td>
<td>M</td>
<td>CABG</td>
<td>40</td>
<td>Lipitor 20 mg</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>26</td>
</tr>
<tr>
<td>II:6</td>
<td>F</td>
<td>AMI/PTCA</td>
<td>39</td>
<td>Zocor 20 mg</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>24</td>
</tr>
</tbody>
</table>

PTCA = Percutaneous Transluminal Coronary Angioplasty; AP = Angina Pectoris; AMI = Acute Myocardial Infarction; CABG = coronary artery bypass surgery; BMI = body mass index (kg/cm²). † = index case. Subjects were considered smokers if they were current smokers or when they quit smoking within the last 5 years. Hypertension was defined as a systolic blood pressure > 140 mmHg and/or diastolic blood pressure > 80 mmHg or the use of anti-hypertensive lowering drugs. Diabetes mellitus was defined as fasting plasma glucose ≥ 7.0 mmol/l or 2h plasma glucose ≥ 11.1 mmol/l as defined by the World Health Organization (WHO).
### Tabel S2: Linkage intervals with LODscore > -2 (NCBI137/hg19).

<table>
<thead>
<tr>
<th>Chr#</th>
<th>Start</th>
<th>SNP</th>
<th>END</th>
<th>SNP</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr7</td>
<td>64406438</td>
<td>rs10949950</td>
<td>65357776</td>
<td>rs35850374</td>
<td>951338</td>
</tr>
<tr>
<td>Chr7</td>
<td>66689728</td>
<td>rs11761805</td>
<td>67026320</td>
<td>rs4576304</td>
<td>336592</td>
</tr>
<tr>
<td>Chr7</td>
<td>117378355</td>
<td>rs739798</td>
<td>132945754</td>
<td>rs10215367</td>
<td>15567399</td>
</tr>
<tr>
<td>Chr9</td>
<td>136927656</td>
<td>rs1076148</td>
<td>141213431</td>
<td>End</td>
<td>4285775</td>
</tr>
<tr>
<td>Chr11</td>
<td>123289850</td>
<td>rs2156443</td>
<td>131697483</td>
<td>rs11603321</td>
<td>8407633</td>
</tr>
<tr>
<td>Chr13</td>
<td>112349868</td>
<td>rs9560166</td>
<td>113839747</td>
<td>rs515863</td>
<td>1489879</td>
</tr>
<tr>
<td>Chr14</td>
<td>0</td>
<td>start</td>
<td>20709688</td>
<td>rs7156806</td>
<td>20709688</td>
</tr>
<tr>
<td>Chr17</td>
<td>64816464</td>
<td>rs4791032</td>
<td>65343997</td>
<td>rs9904424</td>
<td>527533</td>
</tr>
<tr>
<td>Chr18</td>
<td>429354</td>
<td>rs17564131</td>
<td>4780620</td>
<td>rs6506247</td>
<td>4351266</td>
</tr>
<tr>
<td>Chr20</td>
<td>2340973</td>
<td>rs6082889</td>
<td>9378671</td>
<td>rs1997696</td>
<td>7037698</td>
</tr>
<tr>
<td>Chr21</td>
<td>43448796</td>
<td>rs220229</td>
<td>44715784</td>
<td>rs762391</td>
<td>1266988</td>
</tr>
<tr>
<td>Chr22</td>
<td>0</td>
<td>start</td>
<td>17254399</td>
<td>rs2190742</td>
<td>17254399</td>
</tr>
</tbody>
</table>

SUM (bp)= 82186188

### Table S3: Clinical characteristics of participants in the PAS cohort (n = 935).

<table>
<thead>
<tr>
<th>Male/Female</th>
<th>Age (years)</th>
<th>BMI</th>
<th>LDL-c</th>
<th>HDL-c</th>
<th>TG</th>
</tr>
</thead>
<tbody>
<tr>
<td>708/227</td>
<td>43 ± 5</td>
<td>26.9 ± 1.2</td>
<td>3.11 ± 1.29</td>
<td>1.14 ± 0.32</td>
<td>2.06 ± 3.41</td>
</tr>
</tbody>
</table>

Data are expressed as number (N) and presented as mean ± standard deviation. BMI = body mass index in kg/m²; LDL-c = low-density lipoprotein cholesterol; HDL-c = high-density lipoprotein cholesterol; TG = plasma triglycerides. All lipid data are expressed as mmol/l.
A Rare Variant in *MCF2L* Identified using Exclusion Linkage in a Pedigree with PAS

### Table S4: Genotyping in replication cohort.

<table>
<thead>
<tr>
<th>Cohort name</th>
<th>n</th>
<th>CAMSAP1</th>
<th>ZC3HC1</th>
<th>MCF2L</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAS</td>
<td>935</td>
<td>n/a*</td>
<td>6</td>
<td>1†</td>
</tr>
<tr>
<td>Sanquin</td>
<td>1440</td>
<td>13</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>CBR</td>
<td>8000</td>
<td>n/a*</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

PAS = Premature AtheroSclerosis; Sanquin = Sanquin Blood Bank; CBR = Cambridge Bioresource and n/a = not applicable. * = Genotyping was not performed due to a high frequency in the Sanquin cohort. † = Pedigree Index case, other pedigree members were not included in the database.

**RNA isolation and RT-PCR analysis**

*Isolation of RNA from Circulating White Blood Cells*

Total cellular RNA from the buffy coat fraction was extracted using TriZol (Life Technologies, the Netherlands) according to manufacturer’s protocol and reverse transcribed using random primers and iScript reverse transcriptase (BioRad, Veenendaal, The Netherlands). The conditions were 5 min 25°C, 30 in 42°C, 5 min 85°C. mRNA expression levels were measured using SensiFast Sybr Green master mix (GE-Biotech) on CFX386 system (BioRad). Primers were designed using Primer3 software and were exon-intron boundary crossing (Table S2). RT-PCR conditions were: 10 min 95°C followed by 40 cycles 15’ 95°C 30’ 60°C and finally a melt reaction going from 65°-95°C in 5’ per grade. Gene expression was calculated using the 2△ΔCt method using 36B4 as reference gene.

**Figure S2: mRNA expression of MCF2L, Rac1 and RhoA in total circulating white blood cells in 4MCF2L-689Gly variant carriers and one non-carrier relative.**

mRNA expression was analyzed in three different white blood cell RNA fractions in duplicate. MCF2L expression was significantly decreased (p < 0.05) in MCF2L-689Gly variant carriers.
Immunohistochemistry

Sections were fixed in 4% paraformaldehyde and subsequently embedded in paraffin. For immunostaining, paraffin sections were deparaffinised before endogenous peroxidase quenching. After blocking with Ultra V Block (Thermo Fischer Scientific, Fremont, CA, USA) slides were incubated with anti-MCF2L 1:250 (rabbit polyclonal anti-MCF2L, WH0023263M1, Sigma, Munich, Germany) overnight at 4°C. Staining was performed with anti-rabbit Horse Radish Peroxidase (HRP)-labelled IgG (ImmunoLogic, Duiven, The Netherlands) followed by Bright DAB+ visualization (ImmunoLogic, Duiven, The Netherlands). Counterstaining was performed using hematoxylin and slides were cover-slipped with VectaMount (Vector Laboratories, Burlingame, CA, USA). Positive controls consisted of samples of human tonsil. Negative controls consisted of experimental tissues stained without the addition of primary antibodies following the same protocol. Microscopy pictures were analyzed using Adobe Photoshop CS4.

To explore cellular localization of MCF2L in human atherosclerotic lesions the sequential alkaline phosphatase (AP) double staining method was used as described elsewhere [1]. MCF2L was stained in combination with either smooth muscle cell α-actin (SMA) (1:500; mouse 1A4 antibody, DAKO, Glostrup, Denmark), macrophage CD68 (1:100; mouse monoclonal anti-human CD68 PG-M1, DAKO, Heverlee, Belgium), endothelial cell CD34 (1:1000; mouse anti-human CD34Q, Bend10, ThermoFisher Scientific, Waltham, MA, USA), antigen presenting cells (CD11c; 1:50; mouse monoclonal 5D11, Monosan, Uden, The Netherlands), CD8 lymphocytes (1:50; mouse monoclonal anti-human CD8, CD8/144B, DAKO, Heverlee, Belgium) and rabbit anti human CD3 (1:5000; IgG monoclonal SP7, ThermoFisher Scientific). Visualization was performed with vector blue (Vector Laboratories) for MCF2L and vector red (Vector Laboratories) for SMA, CD68, CD34, CD3, CD11c and CD8 and cover slipped.

Figure S3: Immunohistochemical specifity of MCF2L.

A. Single staining of MCF2L (brown) in control tonsil tissue.  
B. Negative control in tonsil tissue using only the secondary antibody.
Supplementary references