Atherosclerosis 284 (2019) 11–17

Contents lists available at ScienceDirect

Atherosclerosis

journal homepage: www.elsevier.com/locate/atherosclerosis

Regulation of *MFGE8* by the intergenic coronary artery disease locus on 15q26.1

Sébastien Soubeyrand[a,*], Majid Nikpay[b], Adam Turner[a], Ann-Thu Dang[a], Mikayla Herfkens[a], Paulina Lau[a], Ruth McPherson[a,b,**]

[a] Atherogenomics Laboratory, University of Ottawa Heart Institute, Ottawa, Canada
[b] Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

HIGHLIGHTS

- A chromosome region associated with heart disease influences expression of neighboring genes.
- The region is located in a gene desert and consists of multiple variants in tight linkage.
- MFGE8 is identified as a likely intermediate bridging the variants to heart disease.

ARTICLE INFO

Keywords:
MFGE8
Coronary artery disease
Single nucleotide polymorphism
Smooth muscle cells
Expression quantitative trait loci

ABSTRACT

Background and aims: A recently identified locus for coronary artery disease (CAD) tagged by rs8042271 is in a region of tight linkage disequilibrium (LD) between 2 genes (*MFGE8, ABHD2*) previously linked to atherosclerosis. Here we have explored the regulatory framework of this region to identify its functional relationship to CAD.

Methods: The CAD Associated Region between *MFGE8* and *ABHD2* (CARMA) was investigated by bioinformatic approaches and transcriptional reporter assays to prioritize target genes and identify putative causal variants. Findings were integrated with publicly available gene expression datasets. *MFGE8* silencing was performed in cell models relevant to CAD.

Results: The regulatory potential of CARMA is disseminated sparsely over the entire region. CARMA contains multiple eQTL that regulate *MFGE8* in coronary artery and coronary artery smooth muscle cell (CoSMC). SNPs that predict the expression of *MFGE8* in artery are concordantly associated with higher risk of CAD ($p$val = 0.0014). Targeting CARMA by CRISPR/Cas9 in a cellular model increased *MFGE8* expression. *MFGE8* silencing was found to reduce CoSMC and monocyte (THP-1) but not endothelial cell proliferation.

Conclusions: These findings support a mechanistic link between a GWAS identified CAD risk locus and atherosclerosis. The intergenic locus CARMA regulates *MFGE8* in a haplotype dependent manner. Individuals genetically susceptible to increased *MFGE8* expression exhibit greater CAD risk. Suppressing *MFGE8* expression reduced SMC and THP-1 proliferation. These data support an atherogenic contribution of CARMA/*MFGE8* that may be linked to cell proliferation and/or improved survival of CAD relevant cell types.

1. Introduction

Genome-wide association studies (GWAS) have identified over 300 chromosomal single nucleotide polymorphisms (SNPs) that associate with CAD risk [1]. Together, these loci explain approximately 25% of CAD heritability, suggesting that additional small effect size variants remain unidentified. Identifying disease associated SNPs is only the first step in a validation process that can ultimately uncover unanticipated contributors to disease. The process of linking GWAS SNPs to disease is a challenging one, in large part because most GWAS variants are located remotely from coding genes, which renders the identification of their direct targets difficult [2]. In addition, causal SNPs in out-bred populations are typically part of a larger haplotype region encompassing dozens of additional “passenger” variants that may share

* Corresponding author. University of Ottawa Heart Institute, 40 Ruskin St – H4203, Ottawa, K1Y 4W7, Canada.
** Corresponding author. University of Ottawa Heart Institute, 40 Ruskin St – H4203, Ottawa, K1Y 4W7, Canada.
E-mail addresses: ssoubeyrand@ottawaheart.ca (S. Soubeyrand), rmcpherson@ottawaheart.ca (R. McPherson).

https://doi.org/10.1016/j.atherosclerosis.2019.02.012
Received 16 November 2018; Received in revised form 24 January 2019; Accepted 8 February 2019
Available online 22 February 2019

significant association with a trait or disease but no causal link [3]. This complexity is further compounded by the fact that distinct SNPs within a given haplotype may have shared and unique targets. Thus extracting causal SNPs from sentinel or tagging SNPs, via a fine-mapping process, is important to adequately translate GWAS findings into disease pathogenesis and prevention [4].

With the recent advent of large-scale expression databases (e.g GTEx [5,6]) and novel bioinformatic approaches, new tools are now at our disposal to help prioritize potential targets. Under the assumption that GWAS variants alter the expression of genes implicated in disease etiology, these approaches integrate GWAS genotypes and phenotypes with RNA expression levels. Tools such as Metaxcan and Summary Mendelian Randomization aim to identify plausible causal genes by correlating genotype and disease by leveraging rich expression datasets [7,8]. Others (such as coloc) directly integrate SNPs associated with changes in RNA levels with GWAS signals. Co-localized signals are then interpreted as evidence of a plausible causal link between SNPs and disease [9].

Using joint association analysis of GWAS data, Nikpay et al. previously identified rs8042271 as the top associated CAD associated SNP within 15q26 [10]. rs8042271 is situated between two protein coding genes of relevance to atherosclerosis and its sequelae, *MFGE8 (*Milk Fat Globule-EGF factor 8*)* and *ABHD2* (abhydrolase domain-containing protein 2). MFGE8 has anti-inflammatory properties and in a small data set, CAD was associated with reduced circulating MFGE8 [11]. An antibody-based approach had previously shown reduced MFGE8 expression in plaque *vis-a-vis* healthy tissue in a small set of samples [12]. Conversely, *ABHD2* expression was found to be elevated in vulnerable plaque and has been shown to function as a triglyceride lipase and ester hydolase [13,14]. In addition, the region contains additional genes (*ACAN1, HAPLN3*) that might also represent regulatory targets. Here we explore the functional relationship of rs8042271 and its proximal region on 15q26 to surrounding genes using a combination of *in silico* and *in vivo* approaches. We show rs8042271 to be part of a larger haplotype with evidence of compounded regulatory potential and single out *MFGE8* as the likely intermediate underpinning the statistical association of this haplotype with CAD.

2. Materials and methods

See Supplementary Methods for additional details, clarifications and online resources.

2.1. Cell culture

Coronary artery smooth muscle cells (CoSMC), endothelial cells (CoENDO) and Human Umbilical Vein Cells (HUVEC) were obtained from ATCC and grown using the recommended media. Thoracic artery cells (HITC6) were kindly provided by Dr. J.G. Pickering [15].

2.2. CARMA editing

HuH7iCAS9 were generated by infection with a lentivirus using standard procedures. Briefly, 293FT cells were transfected with Lenti-iCas9-neo [16] (Addgene plasmid 85400) together with psPAX2 and pMD2.G. Infected HuH7iCAS9 Cells were selected with G418 (0.4 mg/ml) for 5 days, followed by induction with Doxycycline (5 μg/ml) and cell sorting (FACS Aria III) using the GFP tag. Sorted cells were maintained in low glucose DMEM supplemented with amphoterin/penicillin/streptomycin. Gene editing was initiated by the addition of doxycycline to induce clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 prior to the transfection of plasmids. Single clones were obtained by seeding the cells in 96 well plates. Cells were allowed to proliferate and editing was assessed by Resolvase assay (Guide-It, TAKARA), as recommended by the supplier.

2.3. Luciferase assay

Fragments were obtained by PCR (CloneAmp; Takara) on a sample heterozygous at rs8042271 and rs8023801. Fragments were inserted using either HiFi assembly (New England Biolab) or restriction digest cloning (Fragment 6) in pGL3promoter (Promega).

2.4. Silencing experiments

For silencing experiments, cells were seeded onto 24 well plates 24 h prior to siRNA treatment. Adherent cells were at 40–60% density at the onset of treatment while monocytic (undifferentiated) THP-1 were at 5×10^5 cells/ml. CoSMC, THP-1 and endothelial cells were incubated with respectively 20 nM and 5 nM final siRNA using Lipofectamine RNAiMax as transfection reagent. CoSMC, media was changed 24 h after addition of siRNA while treatment was continuous for the other cell types. All silencing were performed for 72 h.

2.5. Conditional analysis

To identify SNPs in CARMA region that are independently associated with CAD, we did conditional analysis using Genome-wide Complex Trait Analysis (GCTA) [17], CAD-GWAS summary statistics from CARDIoGRAMplusC4D Consortium [10] and Howson et al. [18] and genotype data from 6681 unrelated subjects of Caucasian ancestry from the Ottawa Heart Genetic Study [19].

2.6. Multi-Omic data analyses

To study the expression of genes in CARMA region with regard to CAD risk, we obtained eQTL data (GTEx) and GWAS data for CAD from the Cardiogram Consortium and jointly analyze them using SMR and MetaXcan programs. SMR tests whether the effect of a SNP on a phenotype is mediated by gene expression whereas MetaXcan is a gene-based test that determines whether change in expression of a gene is causally associated with disease risk (by aggregating the effect of several eQTLs). Both SMR and MetaXcan differentiate between a pleiotropic effect (where GWAS and eQTL signals overlap) and linkage effect (where GWAS and eQTL signals are close but independent) by doing co-localization analysis.

3. Results

3.1. The MFGE8-ABHD2 CAD associated region spans 18 kb and contains multiple linked variants

The intergenic region comprising rs8042271, which we term CARMA (CAD Associated Region between *MFGE8* and *ABHD2*), contains a cluster of ~50 SNPs forming a haplotype of ~18 kb in significant Linkage Disequilibrium (LD) ($r^2 > 0.5$), displaying nominally significant association with CAD (Supplementary Table 1). This indicates that multiple SNPs within the region have the potential to influence CAD risk. Indeed a distinct SNP (rs2083460) within CARMA, linked with rs8042271 ($r^2 = 0.57$) but located ~ 0.26 kb away, was identified as the top associated CAD SNP in a recent study [20]. Interestingly rs2083460 is in tighter LD ($r^2 = 0.8–0.95$) than rs8042271 with the majority of the CAD SNPs in the CARMA region. As such, to identify SNPs that independently increase the CAD risk, we performed conditional analysis using GCTA software. The underlying algorithm (COJO) re-evaluates the association of SNPs with CAD after including the lead SNP (rs8042271) as a covariate in the association model (conditional analysis). SNPs that remain significant after conditional analysis are considered to be independently associated with CAD. We found that the effect of rs2083460 on CAD is no longer significant ($pC = 0.2$, Table 1) after including rs8042271 as a covariate in the association model. This indicates that rs2083460 is a proxy for

Table 1
Association results before and after conditional analysis.

SNP	bp	refA	Initial association results					Results after conditional analysis		
			b	se	*p*	n	PMID	bC	bC_se	*p*C
rs2083460	89574484	T	7.16E-02	1.36E-02	1.41E-07	3.08E + 05	28530674	1.70E-02	1.36E-02	2.10E-01
rs11073826	89571424	A	5.98E-02	1.52E-02	8.02E-05	2.34E + 05	28530674	4.55E-02	1.52E-02	2.70E-03
rs2083460	89574484	T	7.97E-02	1.48E-02	6.93E-08	1.60E + 05	26343387	1.46E-02	1.48E-02	3.22E-01
rs11073826	89571424	A	8.89E-02	1.91E-02	3.24E-06	1.17E + 05	26343387	6.89E-02	1.91E-02	3.08E-04

Statistical association of rs2083460 with CAD was lost after conditioning for rs8042271 but not rs1107382. Two distinct datasets were interrogated: top 2 results are from Howson et al. and the bottom 2 are from Nikpay et al. [10,18]. rs11073826 was the only SNP (1) not in LD with rs8042271/rs2083460 and (2) showing a $p < 10^{-4}$ link to CAD. Reference dataset parameters: SNP ID (SNP), physical position (bp; hg37/19), effect allele (refA), effect size (b), standard error (se), *p*-value from the CAD meta-analysis (*p*), estimated effective sample size (n). Conditional analysis parameters: effect size (bC), standard error (bC_SE) and *p*-value from conditional analyses (*p*C).

rs8042271 and is not associated with CAD independently. This contrasts with rs1107382, a proximal (~2.8 kb away from rs8042271), infrequent SNP (MAF = 0.14, EURO) in equilibrium ($R^2 = 0.05$) with both rs2083460 and rs8042271 which showed independent (*p*C < 0.05) albeit weak CAD (8.02×10^{-5}) association. In summary, these findings point to the presence of two distinct CAD signals in this region: a strong one associated with the rs8042271/rs2083460 haplotype and a weaker one arising from rs11073826.

3.2. Epigenetic landscape of CARMA

Both CAD SNPs are in strong linkage with several other SNPs in CARMA, any one of which could theoretically be causal. In order to prioritize functional SNPs, the architecture of the region was examined for overlaying chromosomal marks that could point to regulatory functions. Using the ENCODE dataset covering 7 cell lines of various origins, H3K27 acetylation levels, indicative of enhancer potential, were low and variable across cell lines (Supplementary Fig. 1). Since epigenetic marks tend to be cell specific, data from vascular cell models were analyzed. CoSMC, human aortic endothelial cells and THP-1 (monocyte model) acetylation profiles were obtained via the cistrome browser (http://cistrome.org/) [21–23]. In coSMC H3K27 acetylation was distributed unevenly over the CARMA region, with evidence of local enrichment. Across the three samples available, H3K27 acetylation was generally higher on the first half of the CARMA region (Supplementary Fig. 2). ENCODE arterial SMC data were consistent with this pattern. Endothelial cells showed a more dispersed acetylation pattern, which was affected by various pro-inflammatory triggers. By contrast, THP-1 showed a local enrichment over the rs8042271/rs2083460 region that was retained upon differentiation into macrophages. Thus CARMA harbors low H3K27 acetylation with some local enrichment and displays evidence of cell type specificity (coSMC vs THP-1), but no salient candidate enhancer region.

3.3. Fragments derived from CARMA are repressive in SMC

Next an enhancer trapping approach was employed to experimentally examine whether regions within CARMA showed regulatory properties, with the expectation that functional variants therein should differentially affect the region's enhancer potential. Each haplotype, based on 1000 Genome phasing of CARMA was divided into 9 fragments of similar lengths which were cloned upstream of a luciferase reporter and introduced in arterial SMCs of thoracic or coronary origin (see Supplementary Table 1 for positions of SNPs within the fragments). SMCs represent the bulk of the normal vasculature by weight and play a crucial role in the pathogenesis of atherosclerosis [24]. Irrespective of cell type, most fragments reduced reporter activity significantly (Fig. 1). Although both haplotypes affected reporter activity similarly, when data were plotted as a ratio of the Alt over Ref haplotype in each experiment, the findings indicated that in CoSMC, the Alt allele

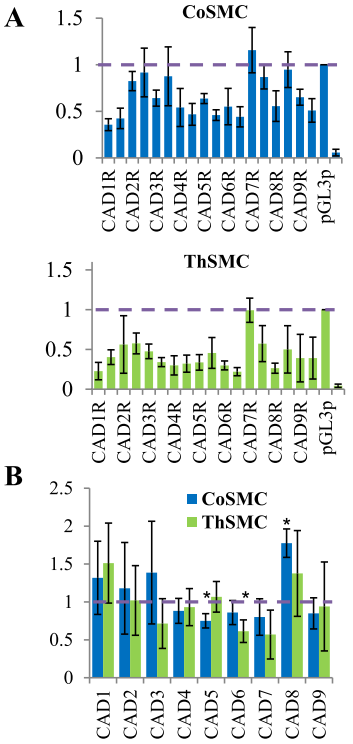

Fig. 1. Enhancer trapping approach to identify candidate regulatory SNPs within CARMA.
Smooth muscle cells of coronary or thoracic origin were transfected with luciferase constructs, each harboring contiguous ~ 2 kb fragments spanning the 18 kb CAD associated region, derived from either reference (R) or alternate (A) haplotypes (CAD1R-CAD9R). (A) Results are normalized to an SV40 *Renilla* internal control and expressed relative to the pGL3promoter construct (± S.D.). (B) Data from A represented as a ratio of Alternate/Reference value, within each experiment, then averaged over 3 experiments. *Statistically significant difference between Alt and Ref in the indicated cell type ($p < 0.05$).

suppressed activity in one fragment (CAD5) and increased it in another (CAD8). By contrast, a distinct fragment (CAD6) contained variants that reduced expression significantly in ThSMC. Thus at least within the context of an episomal system, multiple regions within CARMA can inhibit promoter activity, with some evidence of discrete haplotype and cell type specificity.

3.4. CARMA is part of a topologically associated domain characterized by vascular expression

We then sought direct evidence that the top CAD associated SNPs in this region, rs2083460 and rs8042271, associate with local gene expression. Hi-C data from CAD related tissue (aorta) [25] indicate that

CARMA is located central to a ~800 kb predicted topologically associated domain (TAD), a region characterized by preferential physical interaction, functional integration and conservation (Supplementary Fig. 3) [26,27]. Tissue expression of well annotated (RefSeq) genes located in the CARMA TAD was then examined in the rich GTEx dataset [6]. Among the proximal genes, *ACAN*, *HAPLN3* and *MFGE8* were enriched in arterial tissue, suggestive of a particular function in this vascular bed (Supplementary Fig. 4**).** By comparison, *ABHD2* was more ubiquitously expressed. Lastly, *RLBP1*, *FANC1* and *POLG*, genes more distantly located, exhibited dissimilar expression profiles. Thus these result point to the existence of a functional network encompassing *ACAN*, *HAPLN3*, *MFGE8* and CARMA that may be of particular relevance to vascular tissues.

3.5. CARMA SNPs are eQTLs for MFGE8 in coronary artery and CoSMC

We next investigated the impact of the CAD associated SNPs on the expression of the genes located within the TAD domain. Altered expression of adjacent genes (*MFGE8*, *HAPLN3*, *ABHD2*) but not *ACAN* was evident in most arterial tissues, with *MFGE8* showing significant statistical association specifically in coronary artery (Supplementary Table 2 and Supplementary Fig. 5). Similarly, testing the entire haplotype within CARMA showed that all SNPs within the region are linked to altered *MFGE8* expression in coronary artery, with the protective haplotype associated with reduced *MFGE8* abundance **(**Supplementary Table 3**).** In view of this crucial role, whether CARMA could regulate *MFGE8* in CoSMCs was examined by interrogating expression data derived from a panel of SMCs, focusing on chromosome 15 [28]. Strikingly, among the local genes impacted by rs8042271, only *MFGE8* was significantly affected **(**Table 2). Using cell models aimed at representing normal vascular tissues, our analysis indicates that CoSMC express significantly more *MFGE8* than other cell types suggesting that *MFGE8* may be of particular importance to CoSMC function (Supplementary Fig. 6). By comparison, *ABHD2* showed some enrichment in non-immune models while *HAPLN3* was weakly expressed in CoSMC but enriched in an activated macrophage model.

3.6. MFGE8 expression levels correlate with the incidence of CAD

Next, we studied the expression of genes in the region surrounding CARMA with regard to CAD risk by carrying out Multi-Omic data analyses. We obtained eQTL data (GTEx) and GWAS data for CAD from Cardiogram Consortium and jointly analyzed them using SMR and MetaXcan programs [7,8,29]. Both methods use summary statistics to correlate expression levels with GWAS association data, under the assumption that genetic variants that affect both expression levels and phenotype are more likely to be causal [8]. The approaches differ: SMR leverages individual eQTL information to test correlations while Metaxcan offers an integrative approach using multiple eQTL signals per gene.

SMR analysis was performed on CARMA SNPs associated with CAD. Sufficient data were available on only three polymorphisms. While correlations were observed between CAD risk and expression of genes in the vicinity (*MFGE8* and *HAPLN3*), they approached but did not reach genome-wide significance ($p_{SMR} < 8.4 \times 10-6$) (Supplementary Table 4). Interestingly the identified SNPs were located within fragments distinct from those identified earlier (see Supplementary Fig. 1 for positions of these SNPs). In addition SMR incorporates a colocalization test called HEIDI (HEterogeneity In Dependent Instruments), an additional step that tests SNPs for possible linkages. Implementing HEIDI reveals that the 3 identified SNPs are likely to contribute individually to both CAD and *MFGE8/HAPLN3*.

By contrast, Metaxcan analysis identified only *MFGE8* as significantly associated with both CAD/MI and gene expression in a select subset of arterial tissues (Supplementary Table 5). Colocalization analyses revealed that eQTLs associated with higher expression of *MFGE8* in Aorta concordantly (coloc_prob4 > 0.85**)** increase the risk of CAD (*p*val = 0.00014) and MI (*p*val = 0.0075). This implies that genetic susceptibility to increased expression of MFGE8 confers greater CAD risk.

Table 2
MFGE8 is an eQTL for rs8042271 in CoSMC.

Gene name	*p*val	Position
ANPEP	0.001109843	90,786,872
MFGE8	0.001886597	89,464,652
RP11-97O12.6	0.013201012	88,502,947
HMGB1P8	0.048342644	90,664,718
HAPLN3	0.379658563	90,227,956
ABHD2	0.726121708	89,666,056
ACAN	0.93748799	90,325,184

Genes were tested for eQTL association with rs8042271 within a 2 Mb window. The analysis generated 62 distinct gene-rs8042271 association results. With the exception of *HAPLN3, ABHD2* and *ACAN* included for consistency, only genes reaching nominal significance ($p < 0.05$) are shown. RP11-97O12.6 is a lncRNA of unknown function, *MFGB1P8* is a pseudo-gene while *ANPEP* is an intestinal and renal aminopeptidase. A weak ($4.3E^{-4}$) association between *ANPEP* (rs2305443) and heart failure in the African American population has been reported (dbGaP: phs000226.v6.p1).

3.7. Deleting the CARMA region preferentially affects MFGE8 and HAPLN3 expression

To directly gauge the ability of CARMA to differentially regulate local gene expression, and in particular *MFGE8* expression, CARMA was deleted using a CRISPR/Cas9 approach. While a primary vascular cell model would have been ideal, the limitations inherent to primary cell proliferation required an alternate strategy. Deletions were performed in a transformed model system (HuH-7) transduced with a doxycyclin inducible CRISPR/Cas9. This model was found in pilot experiments to permit efficient editing while having growth properties amenable to follow-up experimentation. Two sets of cgRNAs, each one editing either extremity of CARMA were first identified and the best cgRNAs of each set were combined and introduced alongside CRIPSR/Cas9 (Supplementary Fig. 7). Single clones were obtained, amplified, genotyped by PCR and analyzed for RNA expression of gene neighbors and *SRP14,* an unrelated control housekeeping gene located further down on chromosome 15. Loss of one allele was associated with a modest albeit non-significant increase in both *MFGE8* and *HAPLN3* while loss of both alleles led to a strong increase in both *MFGE8* and *HAPLN3* (Fig. 2A). By contrast levels of *ABHD2* and *SRP14* remained unchanged. With the limitation that a single knock-out clone was available for study, these results indicate that under basal conditions the CARMA region has a net repressive on both *MFGE8* and *HAPLN3*, but not on *ABHD2*.

The large variability in expression level observed across clones sharing a common genotype may reflect clonal drift and/or adaptation. To address this, the impact of CARMA deletion was examined in a cellular pool. While the stochastic nature of gene editing limits editing to a subset of cells (Supplementary Fig. 7), CARMA deletion may still be sufficiently widespread to measurably affect transcription at the population level. Interestingly, when expression was analyzed in the continued presence of CRISPR/Cas9, inhibition of the whole locus (but not *SRP14*) was observed (Fig. 2B), which could reflect the presence of editing and/or repair complexes that impede proximal gene expression. Removal of doxycycline, allowing CRISPR/Cas9 to return to a near-baseline level resulted in increased *ABHD2*, *HAPLN3* and *MFGE8* expression, although only *MFGE8* reached statistical significance. Taken together, these findings further suggest that CARMA represses neighboring genes and further prioritize *MFGE8* and *HAPLN3* as *bona fide* regulatory targets.

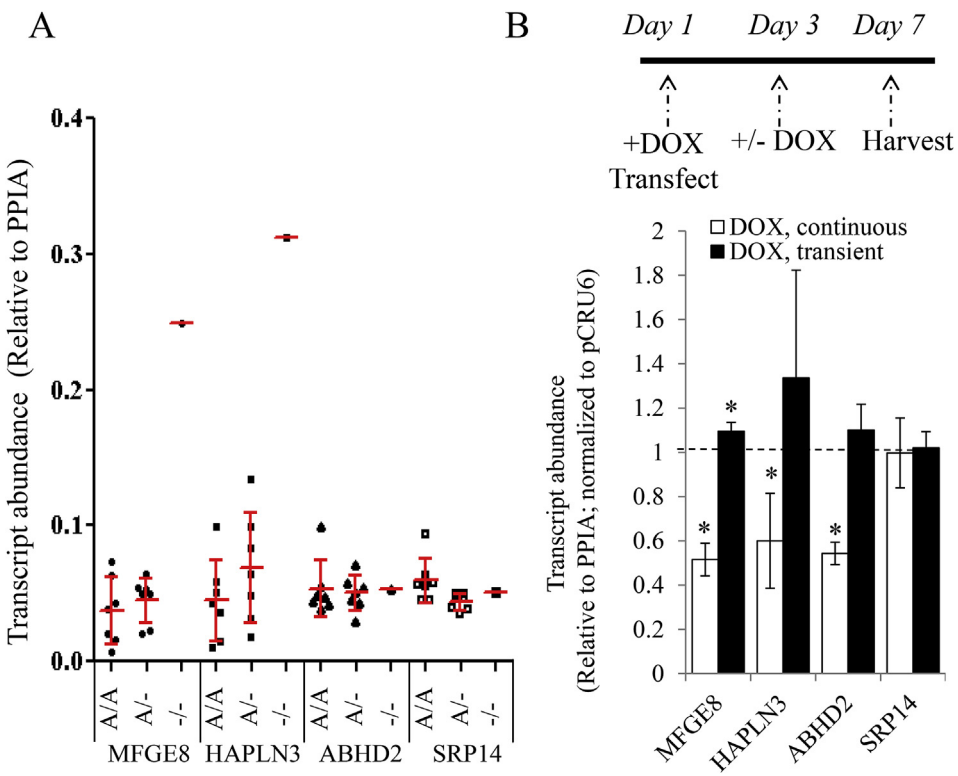

Fig. 2. Targeting CARMA by CRISPR/Cas9 affects local gene expression.
(A) Deletion of CARMA in single clones affects *MGFE8* and *HAPLN3*. HuH-7 stably transduced with a doxycline (DOX) inducible Cas9 (HuH-7iCas9) were transfected with two cgRNA targeting the ends of CARMA. Single clones were amplified and analyzed for expression by qRT-PCR. To simplify visualization and comparison across transcripts, values shown are adjusted to fit on the same scale (MFGE8, X 40; HAPLN3, X 900, ABHD2, X 2, Y; SRP14, X 1). Each data point corresponds to the average value of duplicate determinations from a unique clone. Error bars represent the average value of individual clones ± SD. Data are representative of 3 independent determinations performed over 3 passages. (B) Targeting CARMA in a pool of cells affects local expression. HuH-7iCas9 were transfected as above except that 2 days post transfection cells were exposed to fresh media containing doxycycline or vehicle alone, and left for an additional 4 days. Results from 4 biological replicates are expressed relative to matched DOX treated HuH-7iCas9 cells transfected with pCRU6 control plasmid. Error bars represent 95% CI ($\alpha = 0.05$).

3.8. *MFGE8* is required to support CoSMC and THP-1 proliferation

The protective allele at this locus correlates with lower *MFGE8* expression in arterial tissues. As arterial tissue consists of several anatomically distinct layers, *MFGE8* regulation via CARMA could occur in a number of cell types that constitute normal artery. In view of previous data suggesting a role for MFGE8 in cell proliferation, knock-down experiments were performed in CoSMC, CoENDO and HUVEC and changes in cell number assessed. HUVEC were included as controls given their previously demonstrated requirement for *MFGE8* [30]. In view of the importance of monocyte infiltration and proliferation in CAD, a human monocyte model (THP-1) was also included in this analysis. *MFGE8* silencing reduced CoSMC and THP-1 proliferation while neither CoENDO cells nor HUVEC were affected (Fig. 3 and Supplementary Fig. 8).

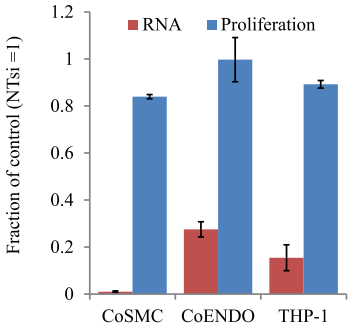

Fig. 3. *MFGE8* silencing reduces coronary smooth muscle proliferation and THP-1 cells.
CoSMCs, CoENDO and THP-1 cells were treated with a control siRNA (NTsi) or a cognate *MFGE8* siRNA. For RNA measurements, RNA was isolated and converted to cDNA and *MFGE8* and *PPIA* were quantified by qRT-PCR followed by *MFGE8* normalization to *PPIA*; values are expressed relative to NTsi. Proliferation was assessed by an end point assay (adherent cells) or cell counting (THP-1). Values are normalized to the matching Non-Target siRNA control. Values represent the average of 3 independent determinations ± S.D. All changes, except for CoEndo proliferation, were statistically significant ($p < 0.05$).

4. Discussion

The majority of common CAD associated variants exhibit small individual effects on disease risk, and many are likely to exert their effects by altering promoter or enhancer activity of adjacent or distal genes. Using DNAse hypersensitivity sites as a measure of (positive) regulatory elements, it has been estimated that each gene is on average under the influence of 15–150 regulatory elements [31]. In addition, pleiotropy, the ability of a variant to independently affect several traits or targets is prevalent. In this context, it is not surprising that progress in deciphering the causal effects of some 300 CAD associated loci has been slow.

Here we have interrogated the functional link of an intergenic locus on 15q26.1, CARMA, to CAD risk. The rare allele (MAF = 0.06) in Europeans, associated with protection from CAD, is much more frequent in East Asians (MAF = 0.35) and is the common allele in African (MAF = 0.57) populations. We hypothesized that the protective effect is mediated in *cis*, consistent with the observation that local regulation is prevalent and that linked SNPs within this locus were found to be significant *cis*-eQTLs for *MFGE8* in coronary artery. Moreover the protective alleles associate with reduced *MFGE8* expression in both coronary artery and isolated coronary smooth muscle cells.

In addition, we sought to identify causal variants within CARMA that could be of potential value in the generation of actionable targets (e.g transcriptional regulators) and further help clarify causal genes. Using both *in silico* and *in vivo* assays, no single variant could be identified. Rather several regions within CARMA were found to have regulatory potential supporting an additive burden model whereby the net regulatory potential of the region is mediated by the combinatorial contribution of several variants of weaker effect. The number precisely involved remains unclear but the occurrence of multiple causal SNPs at a single GWAS locus has been previously reported (Mohlke et al. and references therein [4]). Experiments in yeast have suggested that

haplotype blocks rather than individual variants may be selected for by evolution [32]. Improved imputation and annotation methods as well as increased statistical power from enriched genome-wide transcription databases should offer a more precise picture of the salient regulatory SNPs within CARMA. Ultimately, as gene editing methods are streamlined, replacements of individual or group of SNPs within the native chromatin environment of the cell may be envisaged.

While we could not pinpoint a causal SNP, the fine-mapping process identified *MFGE8* as a functionally relevant target of CARMA based on several criteria: significant eQTL association in arterial tissues, a history consistent with an involvement in atherogenesis, CARMA perturbation experiments and finally a role for MFGE8 in CoSMC and THP-1 proliferation. A limitation to our findings is that while we focused on protein coding genes the topologically associated domain region contains additional genes that may be regulated by CARMA and/or play a role in CAD, including poorly characterized non-coding RNAs (e.g. RP11-326A19.4) and pseudo-genes for which limited to no further information is available.

Our findings do not rule out a contribution of *ABHD2* and/or *HAPLN3* in atherosclerosis, but rather suggest that any contribution therein is independent of polymorphisms within CARMA. *ABHD2* expression for instance is high in vulnerable plaque [13]. By contrast, examination of past efforts comparing genome-wide expression changes at different stages of plaque development reveals a reduction in *MFGE8*, but not *ABHD2* nor *HAPLN3* in advanced lesions [33]. Similarly in a mouse model, advanced plaque is associated with reduced *Mfge8* [12]. We propose that this apparent reduction in *MFGE8* is in fact secondary to the loss of normal smooth muscle cells, which form the bulk of normal vascular tissue and express most of the arterial *MFGE8*, at the expense of the plaque-specific pathological cell types that express lower levels.

MFGE8 has been reported to promote cell proliferation and neovascularization in murine models. In one study, absence of *Mfge8* (in a $Apoe^{-/-}$ background) led to reduced proliferation of smooth muscle in a post-injury model [34]. In another report, experiments in HUVEC demonstrated that proliferation was impaired in the presence of an MFGE8 specific antibody, a process linked to blunted integrin signaling [30]. In contrast, we show that targeting *MFGE8* fails to alter HUVEC or CoENDO proliferation. It is conceivable that the antibody affected proliferation by achieving greater *de facto* reduction of MFGE8, for instance by sequestering all available protein. Alternatively, via anchoring on MFGE8, the antibody might have affected an unrelated but nearby protein involved in proliferation. Here, our silencing approach reduced *MFGE8* levels by ~70% in HUVEC and CoENDO and by almost 100% in coSMC; however residual *MFGE8* may suffice to promote cell growth. While our data support a role for *MFGE8* in the regulation of human CoSMC proliferation and/or cell survival, the supra-physiological nature of the suppression must be borne in mind.

Whether *MFGE8* plays a role in macrophage polarization and/or foam cell formation and expansion, such as by reducing phagocytic ability of infiltrated macrophages, merits further investigation. Although Viola et al. noted in a restenosis model that neointimal macrophage composition was not affected by *Mfge8* deletion, other functional consequences were not reported [34]. By contrast, bone marrow deletion of *Mfge8* results in increased atherosclerosis [12], an effect that appears specific to accumulation of apoptotic cell corpses and alteration of the protective immune response in bone marrow-derived cells. In addition, data obtained from mouse models demonstrated a requirement for *Mfge8* in phagocytosis [35,36], although the occurrence of a similar mechanism in plaque formation remains is a matter of conjecture. In view of the crucial inflammatory component of plaque etiology, understanding if and how *MFGE8* and CARMA regulate inflammatory processes is important and warrants further investigation. Preliminary experiments using THP-1 cells reveals that macrophage loading with oxidized LDL leads to reduced expression of *MFGE8, ABHD2* and *HAPLN3* (*data not shown*). The contribution of CARMA polymorphisms to these processes, if any, is unknown.

Our findings support a proatherogenic contribution of *MFGE8*, consistent with the fact that the protective allele at this locus associates with decreased *MFGE8* expression in coronary artery but need to be reconciled with prior reports suggesting that *MFGE8* plays a protective role in humans. In a small case control study, circulating protein levels of MFGE8 were found to be inversely correlated to CAD severity, although it was unclear whether the reduction in MFGE8 was causal or secondary to the disease itself [11]. Of note the protective allele at rs80421271 associates with reduced *MFGE8* expression in coronary artery but marginally increased expression in whole blood (Supplementary Table 2). It remains possible that the protective genotype leads to increased secretion and abundance of MFGE8 protein in a cell-type specific manner. Ultimately, clarification of this last point will require the identification of tissues and organs that contribute to circulating concentrations of MFGE8.

VSMCs in the normal vessel wall have a low turnover, with barely measureable proliferation indices. Increased cell proliferation is observed during early atherogenesis and we propose that upon vascular insult, proliferation of naïve CoSMC is modestly reduced as a result of lower *MFGE8* expression in patients harboring the protective haplotype (Supplementary Fig. 9). This model is clearly an oversimplification and proliferation of VSMC of different origins have beneficial effects at different stages of atherosclerosis [37,38]. In addition, or alternatively, lower circulating MFGE8, originating from VSMC harboring the protective allele, may result in dampened local inflammatory processes, perhaps by reducing monocyte proliferation. Consistently, our THP-1 findings point to a role of *MFGE8* in monocyte proliferation.

Other roles for MFGE8 have been proposed. Deletion of *Mfge8* in mice was associated with reduced circulating lipids and protection from obesity [39] ascribed to reduced lipid uptake by enterocytes. Increased plasma cholesterol was reported in $Mfge8^{-/-}/Apoe^{-/-}$ relative to $Apoe^{-/-}$ mice fed a high-fat diet [34] . However, we find no association of rs8042271 or other SNPs within the CARMA region or *MFGE8* with human lipid traits (Supplementary Fig. 10) consistent with large GWAS data sets [40]. A recent study in a small Chinese cohort found rs4932450 (intronic to *MFGE8* and in an LD block distinct from CARMA) to associate modestly with several metabolic traits, including HDL [41]. This SNP is not a significant eQTL for *MFGE8* expression in either liver or coronary artery and shows a weak correlation in whole blood ($p = 0.03$) in GTEx but is not associated with CAD. Analyses of blood samples from the Framingham Heart Study [42] (https://preview-ncbi-nlm-nih-gov.proxy.bib.uottawa.ca/gap/eqtl/studies/) revealed a cis-eQTL link to *MFGE8* ($1E^{-12}$) as well as a trans-eQTL with *OXA1L* ($1E^{-4}$) encoding a mitochondrial inner membrane protein. The latter finding is may indicate that despite its presence within *MFGE8*, the variant affects a distal target reminiscent of the *FTO* locus [43]. The effect if any on metabolic traits could be via long distance interactions rather *MFGE8 per se*.

In conclusion our findings reinforce the role of *MFGE8* in CAD by pointing to a genetic contribution of a CAD associated region to *MFGE8* regulation. A limitation to these findings is their reliance on datasets obtained from European populations. The contribution of CARMA/*MFGE8* to CAD in other genetic backgrounds is unresolved. Two smaller CAD GWAS in Japanese and Han Chinese populations failed to detect a genome-wide significant signal at this locus [44,45]. While there is general consistency across different populations in terms of direction of effect for variants linked to common traits, multiple exceptions are reported [46,47]. Whether the absence of correlation reflects a fundamental difference in the relationship of CARMA to CAD or insufficient power will need to be addressed in future studies.

Conflicts of interest

The authors declared they do not have anything to disclose regarding conflict of interest with respect to this manuscript.

Author contributions

Manuscript writing: SS, MN, RM; Study design: SS, AT, RM; Data analysis and interpretation: SS, MN; Experiments: SS, MH, ATD, PL;

Financial support

This work was funded by a Canadian Institutes for Health Research Foundation grant (RM).

Acknowledgments

We are grateful to Curtis Brandt and Tom Lagace for kindly sharing the iCAS9 plasmid.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atherosclerosis.2019.02.012.

References

[1] B. Vilne, H. Schunkert, Integrating genes affecting coronary artery disease in functional networks by multi-OMICs approach, Front. Cardiovasc. Med. 5 (2018) 89.
[2] M.D. Gallagher, A.S. Chen-Plotkin, The post-GWAS era: from association to function, Am. J. Hum. Genet. 102 (2018) 717–730.
[3] K.A. Frazer, D.G. Ballinger, D.R. Cox, D.A. Hinds, et al., A second generation human haplotype map of over 3.1 million SNPs, Nature 449 (2007) 851–861.
[4] M.E. Cannon, K.L. Mohlke, Deciphering the emerging complexities of molecular mechanisms at GWAS loci, Am. J. Hum. Genet. 103 (2018) 637–653.
[5] F. Aguet, A.A. Brown, S.E. Castel, J.R. Davis, et al., Genetic effects on gene expression across human tissues, Nature 550 (2017) 204–213.
[6] Gte GTEx Consortium, Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans, Science 348 (2015) 648–660.
[7] A.N. Barbeira, S.P. Dickinson, J.M. Torres, R. Bonazzola, et al., Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics, BioRxiv (2017) 045260https://doi.org/10.1101/045260.
[8] Z. Zhu, F. Zhang, H. Hu, A. Bakshi, et al., Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets, Nat. Genet. 48 (2016) 481–487.
[9] C. Giambartolomei, D. Vukcevic, E.E. Schadt, L. Franke, et al., Bayesian test for colocalisation between pairs of genetic association studies using summary statistics, PLoS Genet. 10 (2014) e1004383.
[10] M. Nikpay, A. Goel, H.-H. Won, L.M. Hall, et al., A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease, Nat. Genet. 47 (2015) 1121–1130.
[11] W. Dai, Y. Li, Y. Lv, C. Wei, et al., The roles of a novel anti-inflammatory factor, milk fat globule-epidermal growth factor 8, in patients with coronary atherosclerotic heart disease, Atherosclerosis 233 (2014) 661–665.
[12] H. Ait-Oufella, K. Kinugawa, J. Zoll, T. Simon, et al., Lactadherin deficiency leads to apoptotic cell accumulation and accelerated atherosclerosis in mice, Circulation 115 (2007) 2168–2177.
[13] K. Miyata, M. Nakayama, S. Mizuta, S. Hokimoto, et al., Elevated mature macrophage expression of human ABHD2 gene in vulnerable plaque, Biochem. Biophys. Res. Commun. 365 (2008) 207–213.
[14] N.K. M, T. V.B.S.C, V. G.K, C.S. B, et al., Molecular characterization of human ABHD2 as TAG lipase and ester hydrolase, Biosci. Rep. 36 (2016) e00358–e00358.
[15] S. Li, Y.S. Fan, L.H. Chow, C. Van Den Diepstraten, et al., Innate diversity of adult human arterial smooth muscle cells: cloning of distinct subtypes from the internal thoracic artery, Circ. Res. 89 (2001) 517–525.
[16] J. Cao, L. Wu, S.-M. Zhang, M. Lu, et al., An easy and efficient inducible CRISPR/Cas9 platform with improved specificity for multiple gene targeting, Nucleic Acids Res. 44 (2016) gkw660.
[17] J. Yang, S.H. Lee, M.E. Goddard, P.M. Visscher, GCTA: a tool for genome-wide complex trait analysis, Am. J. Hum. Genet. 88 (2011) 76–82.
[18] J.M.M. Howson, W. Zhao, D.R. Barnes, W.K. Ho, et al., Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms, Nat. Genet. 49 (2017) 1113–1119.
[19] M. Nikpay, A.F.R. Stewart, R. McPherson, Partitioning the heritability of coronary artery disease highlights the importance of immune-mediated processes and epigenetic sites associated with transcriptional activity, Cardiovasc. Res. 113 (2017) 973–983.
[20] C.P. Nelson, A. Goel, A.S. Butterworth, S. Kanoni, et al., Association analyses based on false discovery rate implicate new loci for coronary artery disease, Nat. Genet. 49 (2017) 1385–1391.
[21] C.L. Miller, M. Pjanic, T. Wang, T. Nguyen, et al., Integrative functional genomics identifies regulatory mechanisms at coronary artery disease loci, Nat. Commun. 7 (2016) 12092.
[22] D.H. Phanstiel, K. Van Bortle, D. Spacek, G.T. Hess, et al., Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development, Mol. Cell. 67 (2017) 1037–1048.e6.
[23] N.T. Hogan, M.B. Whalen, L.K. Stolze, N.K. Hadeli, et al., Transcriptional networks specifying homeostatic and inflammatory programs of gene expression in human aortic endothelial cells, Elife 6 (2017) e22536.
[24] U. Roostalu, J.K. Wong, Arterial smooth muscle dynamics in development and repair, Dev. Biol. 435 (2018) 109–121.
[25] D. Leung, I. Jung, N. Rajagopal, A. Schmitt, et al., Integrative analysis of haplotype-resolved epigenomes across human tissues, Nature 518 (2015) 350–354.
[26] J.R. Dixon, S. Selvaraj, F. Yue, A. Kim, et al., Topological domains in mammalian genomes identified by analysis of chromatin interactions, Nature 485 (2012) 376–380.
[27] N. Harmston, E. Ing-Simmons, G. Tan, M. Perry, et al., Topologically associating domains are ancient features that coincide with Metazoan clusters of extreme noncoding conservation, Nat. Commun. 8 (2017) 441.
[28] B. Liu, M. Pjanic, T. Wang, T. Nguyen, et al., Genetic regulatory mechanisms of smooth muscle cells map to coronary artery disease risk loci, Am. J. Hum. Genet. 103 (2018) 377–388.
[29] E.R. Gamazon, H.E. Wheeler, K.P. Shah, S. V Mozaffari, et al., A gene-based association method for mapping traits using reference transcriptome data, Nat. Genet. 47 (2015) 1091–1098.
[30] J.-S. Silvestre, C. Théry, G. Hamard, J. Boddaert, et al., Lactadherin promotes VEGF-dependent neovascularization, Nat. Med. 11 (2005) 499–506.
[31] D. Hay, J.R. Hughes, C. Babbs, J.O.J. Davies, et al., Genetic dissection of the α-globin super-enhancer in vivo, Nat. Genet. 48 (2016) 895–903.
[32] R. She, D.F. Jarosz, Mapping causal variants with single-nucleotide resolution reveals biochemical drivers of phenotypic change, Cell 172 (2018) 478–490 e15.
[33] J.Y. King, R. Ferrara, R. Tabibiazar, J.M. Spin, et al., Pathway analysis of coronary atherosclerosis, Physiol. Genom. 23 (2005) 103–118.
[34] J. Viola, P. Lemnitzer, N. Paulin, M. Drechsler, et al., Deletion of MFGE8 inhibits neointima formation upon arterial damage, Thromb. Haemostasis 118 (2018) 1340–1342.
[35] R. Hanayama, M. Tanaka, K. Miyasaka, K. Aozasa, et al., Autoimmune disease and impaired uptake of apoptotic cells in MFG-E8-deficient mice, Science (80-.) 304 (2004) 1147–1150.
[36] A. Raymond, M.A. Ensslin, B.D. Shur, SED1/MFG-E8: a bi-motif protein that orchestrates diverse cellular interactions, J. Cell. Biochem. 106 (2009) 957–966.
[37] M.R. Bennett, S. Sinha, G.K. Owens, Vascular smooth muscle cells in atherosclerosis, circ, Res. 118 (2016) 692–702.
[38] S.W. Van Der Laan, Genetic susceptibility loci for cardiovascular disease and their impact on atherosclerotic plaques, Circ Genom Precis Med 11 (2018) 2115.
[39] A. Khalifeh-Soltani, D. Gupta, A. Ha, J. Iqbal, et al., Mfge8 regulates enterocyte lipid storage by promoting enterocyte triglyceride hydrolase activity, JCI Insight 1 (2016) e87418.
[40] R. Do, C.J. Willer, E.M. Schmidt, S. Sengupta, et al., Common variants associated with plasma triglycerides and risk for coronary artery disease, Nat. Genet. 45 (2013) 1345–1352.
[41] L. Liu, C. Cheng, D. Zhang, R. Zhang, et al., MFGE8 polymorphisms are significantly associated with metabolism-related indicators rather than metabolic syndrome in Chinese people: a nested case–control study, Gene 677 (2018) 176–181.
[42] R. Joehanes, X. Zhang, T. Huan, C. Yao, et al., Integrated genome-wide analysis of expression quantitative trait loci aids interpretation of genomic association studies, Genome Biol. 18 (2017) 16.
[43] M. Claussnitzer, S.N. Dankel, K.-H. Kim, G. Quon, et al., FTO obesity variant circuitry and adipocyte browning in humans, N. Engl. J. Med. 373 (2015) 895–907.
[44] F. Takeuchi, M. Yokota, K. Yamamoto, E. Nakashima, et al., Genome-wide association study of coronary artery disease in the Japanese, Eur. J. Hum. Genet. 20 (2012) 333–340.
[45] X. Lu, L. Wang, S. Chen, L. He, et al., Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease, Nat. Genet. 44 (2012) 890–894.
[46] C.S. Carlson, T.C. Matise, K.E. North, C.A. Haiman, et al., Generalization and dilution of association results from european GWAS in populations of non-european ancestry: the PAGE study, PLoS Biol. 11 (2013) e1001661.
[47] W. Ke, K.A. Rand, D. V Conti, V.W. Setiawan, et al., Evaluation of 71 coronary artery disease risk variants in a multiethnic cohort, Front. Cardiovasc. Med. 5 (2018) 19.