REVIEW

OPEN

From genotype to phenotype in human atherosclerosis – recent findings

Lesca M. Holdt and Daniel Teupser

Purpose of review
Since 2007, genome-wide association studies (GWAS) have led to the identification of numerous loci of atherosclerotic cardiovascular disease. The majority of these loci harbor genes previously not known to be involved in atherogenesis. In this review, we summarize the recent progress in understanding the pathophysiology of genetic variants in atherosclerosis.

Recent findings
Fifty-eight loci with $P < 10^{-7}$ have been identified in GWAS for coronary heart disease and myocardial infarction. Of these, 23 loci (40%) overlap with GWAS loci of classical risk factors such as lipids, blood pressure, and diabetes mellitus, suggesting a potential causal relation. The vast majority of the remaining 35 loci (60%) are at genomic regions where the mechanism in atherogenesis is unclear. Loci most frequently found in independent GWAS were at Chr9p21.3 (*ANRIL/CDKN2B-AS1*), Chr6p24.1 (*PHACTR1*), and Chr1p13.3 (*CELSR2*, *PSRC1*, *MYBPHL*, *SORT1*). Recent work suggests that Chr9p21.3 exerts its effects through epigenetic regulation of target genes, whereas mechanisms at Chr6p24.1 remain obscure, and Chr1p13.3 affects plasma LDL cholesterol.

Summary
Novel GWAS loci indicate that our understanding of atherosclerosis is limited and implicate a role of hitherto unknown mechanisms, such as epigenetic gene regulation in atherogenesis.

Keywords
1p13.3, 6p24.1, 9p21.3, atherosclerosis, GWAS

INTRODUCTION

Findings from genome-wide association studies (GWAS) are a treasure trove for our understanding of the pathophysiology of atherosclerosis. The first GWAS in 2007 identified a locus on chromosome 9p21.3 (Chr9p21.3), which is the strongest genetic factor of atherosclerosis known today [1–4]. Since then, additional loci have been constantly added, resulting in over 50 loci. The majority is completely novel and the current challenge in the 'post GWAS era' is to identify the responsible genes and integrate them into our understanding of the pathophysiology of this frequent disease.

Here, we focus on the most robust loci identified by GWAS and review some of the approaches recently used to tease out their complex pathophysiology. These approaches include expression quantitative trait loci (eQTL) and functional studies in tissues from patients with defined genotypes, which are essential to single-out the culprit gene at loci usually containing multiple transcripts. Moreover, overlap with GWAS hits of cardiovascular risk factors and seemingly unrelated phenotypes gives hints to potentially causal relations. Finally, cell culture studies and mouse models using knockout and overexpression strategies are essential, in particular at loci involving completely novel pathophysiology. Understanding the mechanisms of these loci in atherogenesis is a prerequisite for later therapeutic targeting.

Institute of Laboratory Medicine, University Hospital Munich (LMU) and Ludwig-Maximilians-University Munich, Munich, Germany

Correspondence to Daniel Teupser, Institute of Laboratory Medicine, University Hospital Munich (LMU) and Ludwig-Maximilians-University Munich, Marchioninistr. 15, 81377 Munich, Germany. Tel: +49 89 7095 3210; fax: +49 89 7095 8888; e-mail: daniel.teupser@med.uni-muenchen.de

Curr Opin Lipidol 2013, 24:410–418

DOI:10.1097/MOL.0b013e3283654e7c

KEY POINTS

- GWAS have identified 58 loci of CHD and MI.
- Forty percent overlap with genomic loci for classical risk factors, whereas mechanisms at the remaining 60% are unclear.
- The three most frequently found loci are at Chr9p21.3, Chr6p24.1, and Chr1p13.3.
- Recent work suggests a role of epigenetic gene regulation by a noncoding RNA as a novel mechanism of atherogenesis at Chr9p21.3.
- Mechanisms at Chr6p24.1 are unclear, and Chr1p13.3 likely works through affecting plasma LDL-cholesterol.

Atherosclerosis is a disease affecting arterial blood vessels, leading to different disease phenotypes depending on the anatomical location and stage of the disease process. Most GWAS have been performed for the phenotype of coronary heart disease (CHD), which includes a broad spectrum of patients with stable and unstable coronary disease, myocardial infarction (MI) survivors and patients undergoing coronary angiography (Table 1) [3–8,9▪,10▪▪,11–17,18▪▪,19–21]. A smaller number of GWAS has specifically dealt with the phenotype MI, which overlaps with CHD because CHD almost always precedes MI. However, MI clearly involves additional mechanisms, such as thrombosis. In this review, we are not covering stroke, which requires differentiation into several subtypes of ischemic and hemorrhagic stroke with different underlying pathophysiology [22]. We are also not explicitely covering peripheral atherosclerosis and its surrogate marker ankle brachial index, where until now GWAS have only revealed the Chr9p21.3 locus with genome-wide significance in a study of more than 40 000 individuals [23].

GWAS LOCI OF CORONARY HEART DISEASE AND MYOCARDIAL INFARCTION

Searching the GWAS catalogue (www.genome.gov/gwastudies; accessed May 2013; [21]) with a stringent cutoff ($P < 10^{-7}$), we have assembled 58 loci from 18 publications for the phenotypes of CHD and MI (Table 1) reporting the best P values (including combined analyses with replication) [1,3–8,9▪,10▪▪,11–17,18▪▪,19,20]. A predominant number of these variants has been identified by the CARDIoGRAM consortium [10▪▪]. A total of 6220 single-nucleotide polymorphisms (SNPs) with $P < 0.01$ from this analysis were followed-up in the CARDIoGRAMplusC4D consortium in 63 746 coronary artery disease cases and 130 681 controls, adding 15 additional loci to the list [18▪▪]. Whereas earlier GWAS were mainly performed in cohorts of European decent, a number of novel loci were recently identified in Asian and Middle Eastern populations [13,15–17,20] (Table 1).

The Chr9p21.3 (*CDKN2B-AS1*) locus, which is the strongest genetic marker of human atherosclerosis and which is generally considered the 'gold standard' for any association study of atherosclerosis-related traits, is listed in 12 independent GWAS publications (Table 1) [1,3,4,9▪,10▪▪,11–13,16,17,18▪▪,19]. Chr9p21.3 stands out because of its relatively large effect size [odds ratio (OR) 1.3 per allele], and its allele frequency of ~50%. The second most frequently identified GWAS locus is on Chr6p24.1 (*PHACTR1*), which has been described in six publications (Table 1, OR 1.10) [9▪,10▪▪,15,17,18▪▪,19]. The third most often found locus is on Chr1p13.3 (Table 1, OR 1.11) [3,9▪,10▪▪,18▪▪,19]. Despite harboring at least four transcripts, *SORT1* is currently considered the prime candidate gene at Chr1p13.3 and was investigated in several functional studies [24–26,27▪]. The current advances in understanding the pathophysiology at each of these three major loci identified so far will be discussed later in this review.

UNDERSTANDING FUNCTION BY CO-SEGREGATION ANALYSIS WITH OTHER TRAITS

A promising strategy for inferring function from a locus is to search for overlap with GWAS loci for other traits. We systematically screened the 58 loci in Table 1 for overlap with GWAS hits for classical risk factors of CHD or MI (lipids, blood pressure/hypertension, diabetes-related phenotypes) and added information from the CARDIoGRAMplusC4D consortium [18▪▪], which also tested for overlap with genetic variants for established risk factors (Table 1). As summarized in Fig. 1, we found that GWAS loci for CHD and MI overlap with 14 loci for lipids (24% of all risk loci), six loci for blood pressure/hypertension (10%), one locus for diabetes mellitus (2%), and two loci with at least two risk factors (4%). Thirty-five (60%) loci did not co-segregate with loci of classical risk factors but out of these, six overlapped with loci from seemingly unrelated GWAS (Table 1; Supplementary material).

Another approach to get insights into function is to investigate the effects of the genotype on mRNA expression of genes at GWAS loci and to map eQTLs. This might be particularly helpful to identify the culprit gene(s) at loci harboring many

Table 1. Summary of 58 GWAS loci for CHD and MI with $P < 1 \times 10^{-7}$ as of May 2013

Region	Genes	SNP	FRA	p-value	OR	Reference	WTCCC[4]	Samani NJ [3]	Tregouet DA[5]	Erdmann J[6]	Erdmann J[7]	Reilly MP [8]	C4D [9]	Schunkert H [10]	Wild PS[11]	Slavin TP[12]	Takeuchi F[13]	Davies RW [14]	Hager J [15]	Lu X [16]	Lee JY [17]	Cardio.+C4D [18]	Helgadottir A[1]	Kathiresan S[19]	Aoki A[20]	# citations CHD	# citations CHD+M	Lipids	Blood pressure	Diabetes	Other	None
Cohorts							E	E	E	E	E	E	EA	E	E	E	J	E	M	C	K	E	E	E	J							
# of loci							1	4	1	1	1	2	8	24	2	4	2	2	1	7	3	45	1	9	1			16	8	2	15	33
1p13.3	*SORT1, PSRC1, CELSR2*	rs599839 -A	0.78	3×10^{-10}	1.11	[10, 19]		■					■	■								■		■		4	5	■				
1p32.2	*PPAP2B*	rs17114036 -A	0.91	4×10^{-19}	1.17	[10]								■								■				2	2					■
1p32.3	*PCSK9*	rs11206510 -T	0.82	9×10^{-8}	1.08	[10, 19]								■								■		■		2	3	■				
1q21.3	*IL6R*	rs4845625 -T	0.47	4×10^{-10}	1.06	[18]																■				1	1					■
1q41	*MIA3*	rs17465637 -C	0.74	1×10^{-8}	1.14	[10, 19]								■								■		■		2	3					■
2p21	*ABCG5/8*	rs6544713 -T	0.30	$2x\ 10^{-9}$	1.06	[18]																■				1	1	■			A	
2p24.1	*TTC32, WDR35*	rs2123536 -T	0.39	7×10^{-11}	1.12	[16]														■						1	1					
2p24.1	*ABOB*	rs515135 -G	0.83	3×10^{-10}	1.07	[18]																■				1	1	■				
2q22.3	*ZEB2-ACO74093.1*	rs22526 41 -G	0.46	5×10^{-8}	1.06	[18]																■				1	1					■
2q31.1	*VAMP5/8-GGCX*	rs1561198 -A	0.45	1×10^{-10}	1.06	[18]																■				1	1					■
2q33.2	*WDR12, NBEAL1*	rs6725887 -C	0.15	1×10^{-9}	1.14	[10, 19]								■								■		■		2	3					■
3q22.3	*MRAS*	rs2306374 -C	0.18	3×10^{-8}	1.12	[10]				■				■								■				3	3					■
3q26 .1	*Intergenic*	TSH1		4×10^{-10}	NR	[12]										■										1	1					
4p16.2	*STK32B*	TSH2		2×10^{-11}	NR	[12]										■										1	1					
4q31.23	*EDNRA*	rs1878406 -T	0.15	3×10^{-8}	1.10	[18]																■				1	1				B	■
4q32.1	*GUCY1A3*	rs1842896 -T	0.76	1×10^{-11}	1.14	[16, 18]														■		■				2	2		■			
5p15.33	*IRX1, ADAMTS16*	rs11748327 -?	NR	5×10^{-13}	1.25	[20]																		■		0	1					
5q31.1	*SLC22A4 -SLC22A5*	rs273909 -A	0.14	1×10^{-9}	1.07	[18]																■				1	1					■
6p21.1	*VEGFA -C6orf223*	rs6905288 -T	NR	7×10^{-8}	1.23	[14]												■								1	1			■	C	
6p21.2	*KCNK5*	rs10947789 -T	0.76	1×10^{-8}	1.07	[18]																■				1	1					■
6p21.31	*ANKS1A*	rs17609940 -G	0.75	1×10^{-8}	1.07	[10]								■								■				2	2	■			D	
6p21.32	*C6orf10, BTNL2*	rs9268402 -G	0.59	3×10^{-15}	1.16	[16]														■						1	1				E	
6p21.33	*HCG27, HLA -C*	rs3869109 -G	NR	1×10^{-9}	1.14	[14]												■								1	1					
6p24.1	*PHACTR1*	rs12526453 -C	0.67	1×10^{-9}	1.1	[10, 19]							■	■					■	■		■		■		5	6		■		F	■
6q23.2	*TCF21*	rs12190287 -C	0.62	1×10^{-12}	1.08	[10]								■								■				2	2					■
6q25.1	*MTHFD1L*	rs6922269 -A	0.25	3×10^{-8}	1.23	[3]		■																		1	1					
6q25.3	*SLC22A3 -LPAL2-LPA*	rs3798220 -C	0.02	3×10^{-11}	1.51	[10]			■					■								■				3	3	■			G	
6q26	*PLG*	rs4252120 -T	0.73	5×10^{-10}	1.07	[18]																■				1	1				H	■
7p21.1	*HDAC9*	rs2023938 -G	0.10	5×10^{-8}	1.08	[18]																■				1	1					■
7q22.3	*BCAP29*	rs10953541 -C	0.8	3×10^{-8}	1.08	[9]							■									■				2	2					■
7q32.2	*ZC3HC1*	rs11556924 -C	0.62	9×10^{-18}	1.09	[10]								■								■				2	2		■			
8q24.13	*TRIB*	rs2954029 -A	0.55	5×10^{-9}	1.06	[18]																■				1	1	■				
9p21.3	*CDKN2B-AS1*	rs4977574 -G	0.46	1×10^{-22}	1.29	[10, 19]	■	■					■	■	■	■				■	■	■		■	■	10	12					■
9q34.2	*ABO*	rs579459 -C	0.21	4×10^{-14}	1.10	[10]						■		■								■				3	3	■			I	
8p21.3	*LPL*	rs264-G	0.86	3×10^{-9}	1.11	[18]																■				1	1	■				
10p11.23	*KIAA1462*	rs2505083 -C	0.38	4×10^{-8}	1.07	[9]					■		■									■				3	3					■
10q11.21	*CXCL12*	rs1746048 -C	0.87	3×10^{-10}	1.09	[10, 19]		■						■								■		■		3	4					■
10q23.31	*LIPA*	rs1412444 -T	0.42	3×10^{-13}	1.09	[9]							■		■							■				3	3					■
10q24.32	*CYP17A1,CNNM2, NT5C2*	rs12413409 -G	0.89	1×10^{-9}	1.12	[10]							■	■								■				2	2		■			
11q22.3	*PDGFD*	rs974819 -T	0.32	2×10^{-9}	1.07	[9]							■									■				2	2					■
11q23.3	*ZNF259 -APOA5-APOA1*	rs964184 -G	0.13	1×10^{-17}	1.13	[10]								■								■				2	2	■				
12q21.33	*ATP2B1*	rs7136259 -T	0.39	6×10^{-10}	1.11	[16]														■						1	1					
12q23.3	*HSP90B1*	TSH3		3×10^{-9}	NR	[12]										■										1	1					
12q24.11	*MYL2*	rs3782889 -C	0.21	4×10^{-14}	1.26	[17]															■					1	1					
12q24.12	*ACAD10, ALDH2*	rs11066015 -A	NR	5×10^{-11}	1.41	[17]											■					■				2	2				J	
12q24.12	*SH2B3*	rs3184504 -T	0.40	5×10^{-11}	1.07	[18]																■				1	1	■	■	■	K	
12q24.13	*C12orf51*	rs11066280 -A	0.17	2×10^{-11}	1.19	[16]														■						1	1				L	
13q12.3	*FLT1*	rs9319428 -A	0.32	7×10^{-11}	1.06	[18]																■				1	1					■
13q34	*COL4A1, COL4A2*	rs4773144 -G	0.44	4×10^{-9}	1.07	[10]								■								■				2	2				M	■
14q32.2	*HHIPL1*	rs2895811 -C	0.43	1×10^{-10}	1.07	[10]								■								■				2	2					■
15q25.1	*ADAMTS7*	rs3825807 -A	0.57	1×10^{-12}	1.08	[10]						■	■	■								■				4	4					■
15q26.1	*FURIN-FES*	rs17514846 -A	0.44	9×10^{-11}	1.07	[18]																■				1	1		■			
17p11.2	*RAI1- PEMT- RASD1*	rs12936587 -G	0.56	4×10^{-10}	1.07	[10]								■								■				2	2				N	■
17p13.3	*SMG6, SRR*	rs1231206 -A	0.37	9×10^{-10}	1.07	[10]								■								■				2	2					■
17q21.32	*UBE2Z, GIP, ATP5G1, SNF8*	rs46522 -T	0.53	2×10^{-8}	1.06	[10]								■								■				2	2				O	■
19p13.2	*SMARCA4, LDLR*	rs1122608 -G	0.77	1×10^{-8}	1.14	[10, 19]								■								■		■		2	3	■				
19q13.32	*ApoE-ApoC1*	rs2075650 -G	0.14	6×10^{-11}	1.11	[18]								■												1	1	■				
21q22.11	*MRPS6, SLC5A3, KCNE2*	rs9982601 -T	0.15	4×10^{-10}	1.18	[10, 19]								■								■		■		2	3					■

Loci are from the GWAS catalogue (www.genome.gov/gwastudies; accessed May 2013; [21]). Overlap with loci for risk factors of atherosclerosis was added. Black – loci for CHD from GWAS catalogue, dark grey – loci for MI from GWAS catalogue; grey – loci for risk factors of atherosclerosis (lipids, blood pressure/hypertension, diabetes) and other phenotypes from GWAS catalogue, light blue – loci for risk factors and other phenotypes reported in [18▪▪], dark blue – loci for risk factors and other phenotypes reported both in GWAS catalogue and in [18▪▪]. Notes A–O with information of co-segregating GWAS hits ($P < 10^{-5}$) can be found in the Supplementary material. C, Chinese cohort; CHD, coronary heart disease; E, European cohort; EA, European and South-Asian cohorts; FRA, frequency of risk allele; GWAS, genome-wide association studies; J, Japanese cohort; K, Korean cohort; M, Middle Eastern cohort; MI, myocardial infarction; NR, not reported. Loci with two-SNP haplotypes (TSH) [12]: TSH1 – rs11924705, rs6789378 (risk alleles C,A); TSH2 – rs7697839, rs7673097 (risk alleles G,G); TSH3 – rs1165668, rs1165669 (risk alleles G,C). Loci for MI [19]: Chr1p13.3 – rs646776-T, FRA 0.81, *P* value 8×10^{-12}, OR 1.19; Chr1p32.3 – FRA 0.81, *P* value 1×10^{-8}, OR 1.15; Chr1q41 – FRA 0.72, *P* value 1×10^{-9}, OR 1.14; Chr9p21.3 – FRA 0.56, *P* value 3×10^{-44}, OR 1.29; Chr10q11.21 – FRA 0.84, *P* value 7×10^{-9}, OR 1.17; Chr19p13.2 – FRA 0.75, *P* value 2×10^{-9}, OR 1.15; Chr21q22.11 – FRA 0.13, *P* value 6×10^{-11}, OR 1.2.

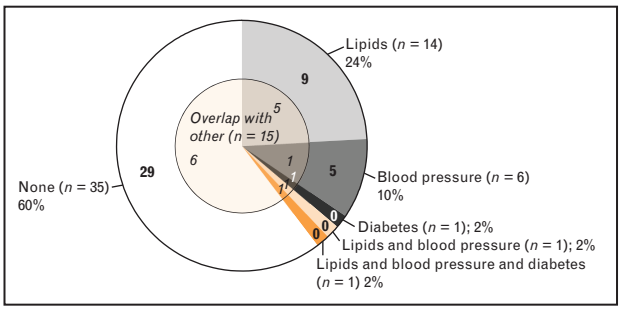

FIGURE 1. Overlap between atherosclerosis loci and loci for common risk factors. Out of 58 loci for coronary heart disease (CHD) and myocardial infarction (MI), 24% overlapped with lipid loci (LDL cholesterol, HDL cholesterol, total cholesterol, triglycerides), 10% with blood pressure, 2% with diabetes-related traits, 2% with lipids and diabetes-related traits, and 2% with all three risk factors. Sixty percent ($n = 35$) of CHD and MI loci did not overlap with loci for common risk factors suggesting novel pathophysiology. The inner circle shows additional overlap with genome-wide association studies hits for other nonrisk factor-associated traits (26% of all loci).

genes. Testing *cis*-regulation of a genetic variant at a genome-wide level requires large cohorts where transcriptome-wide mRNA expression has been assayed in each individual and where genome-wide SNP data are also available. Folkersen *et al.* [28] have systematically tested lead SNPs from GWAS of CHD and MI and found evidence for *cis*-regulation at five loci in different vascular tissues and liver samples (Chr1p13.3: *SORT1, PSRC1, CELSR2;* Chr2q33.2: *NBEAL1*; Chr3q22.3: *MRAS;* Chr6q25.1: *MTHFD1L*; Chr21q22.11: *SLC5A3)*. Wild *et al.* [11] performed a comparable analysis using mRNA expression data from monocytes of 1494 individuals from a population-based study [29] and found three eQTLs (Chr1p13.3: *PSRC1*; Chr2q33.2: *WDR12*; Chr10q23.31: *LIPA*). Results at Chr2q33.2 are particularly interesting since expression analysis in different tissues apparently led to different findings. A similar approach was taken by the C4D consortium, which systematically tested for eQTLs at newly identified loci [9▪]. A current limitation of this very promising approach is the limited availability of large cohorts with tissue collections for transcriptome-wide expression analysis.

Chr9P21.3 (*ANRIL*): ROLE OF A LONG NONCODING RNA (ncRNA) IN ATHEROGENESIS

Chr9p21.3 is the most replicated locus of human atherosclerosis (reviewed in [30,31]). The locus lacks associations with common cardiovascular risk

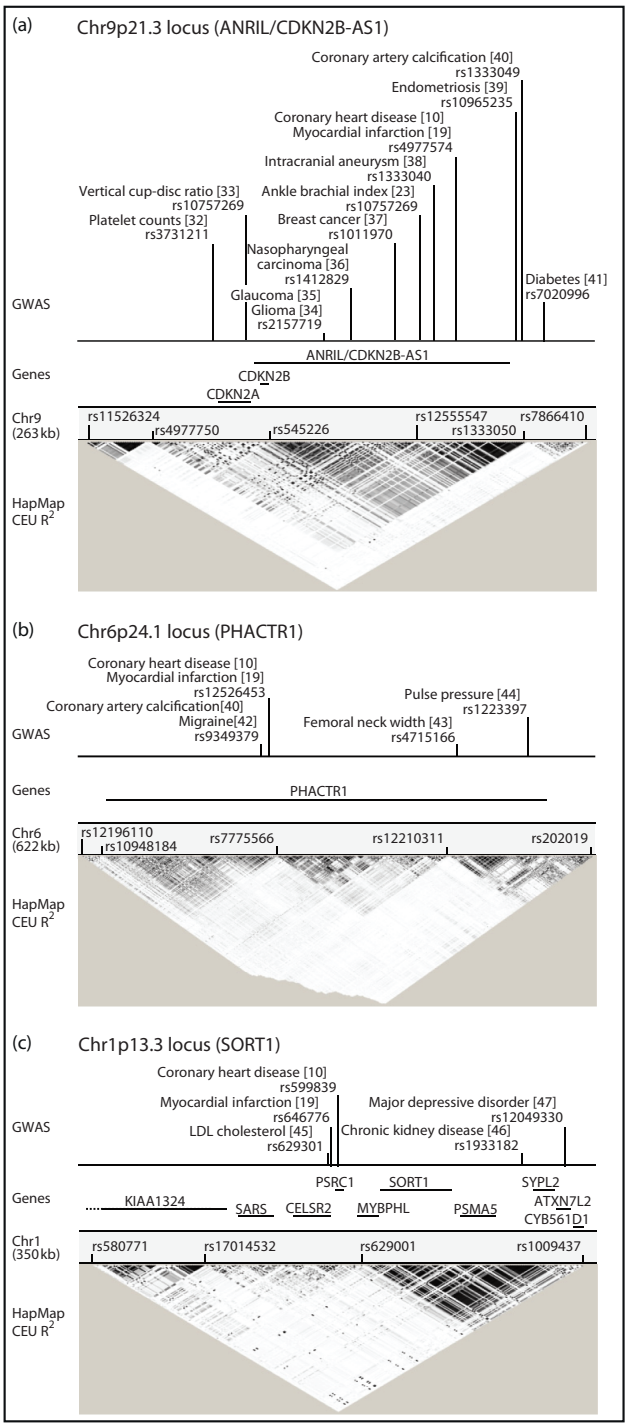

FIGURE 2. Haplotype analysis (HapMap CEU) and annotated genes at the three most frequently identified loci for coronary heart disease (CHD) and myocardial infarction (MI). Single-nucleotide polymorphisms with strongest signals of the respective phenotype and corresponding references are given. (a) Chr9p21.3 CHD and MI locus and adjacent hits for cancer, diabetes, and other traits. (b) Chr6p24.1 CHD and MI locus overlapping with migraine. Significance of pulse pressure and femoral neck width loci is unclear. (c) Chr1p13.3 CHD and MI locus co-segregating with genome-wide association studies (GWAS) hits for lipids.

factors suggesting that it exerts its effect through an alternative mechanism. The core risk haplotype spans approximately 50 kb [10▪▪,19,22,32–47] (Fig. 2a) and does not contain protein-coding genes but the 3′end of the long ncRNA *antisense noncoding RNA in the INK4 locus* (*ANRIL*). The synonyms *CDKN2B antisense RNA 1* (*CDKN2B-AS1*) and *CDKN2BAS* are used for *ANRIL* and refer to its antisense orientation to *cyclin-dependent kinase inhibitor 2B* (*CDKN2B*), which is located proximal to the core CHD region. Together with *CDKN2A*, which is located further proximal of *ANRIL*, this region depicts a GWAS hotspot for different tumor entities [30,34–37] and other phenotypes [32,33], which is in line with loss of function of these genes in many human cancers (Fig. 2a) [48]. In an adjacent haplotype block, an independent locus for diabetes was identified [41]. Despite their expression in human plaques [49], several lines of evidence argue against a role of *CDKN2A* and *CDKN2B* as major Chr9p21.3 effector genes. First, SNPs within these genes are not in linkage disequilibrium with the lead CHD SNPs (Fig. 2a). Second, *cis*-regulation of these genes is lacking in the majority of human studies (reviewed in [30]). Third, mouse models speak against a causal role of *CDKN2B* in atherogenesis [50▪,51] and yielded conflicting results for *CDKN2A* [50▪,52–54].

In contrast, there is growing evidence for a role of *ANRIL* in modulating atherosclerosis susceptibility at Chr9p21.3. *ANRIL* expression is tightly regulated by the Chr9p21.3 genotype [55–58,59▪, 60–62] (for review see [30]). In addition, a positive correlation of *ANRIL* expression with atherosclerosis severity has been described [58]. Transcription of *ANRIL* is complex and more than 20 linear and several circular isoforms are known today [55,57,59▪]. As a mechanism for differential expression, Harismendy *et al.* [63] proposed that *ANRIL* expression in Chr9p21.3 risk allele carriers was induced by disruption of an inhibitory STAT1-binding site. Functional studies in mammalian cells revealed that *ANRIL* knock-down led to decreased proliferation [64–67]. Recent work has extended these findings, showing that *ANRIL* overexpression not only led to accelerated proliferation but also increased adhesion and decreased apoptosis [59▪]. These are key mechanisms of atherogenesis and the direction of effects would be in line with the pro-atherogenic role of *ANRIL* suggested from expression studies (Fig. 3) [59▪].

But how does *ANRIL* exert these effects at the molecular level? *ANRIL* belongs to the group of large noncoding RNAs which have been shown to regulate gene expression through RNA–RNA, RNA–DNA, or RNA–protein interactions [68–70]. For

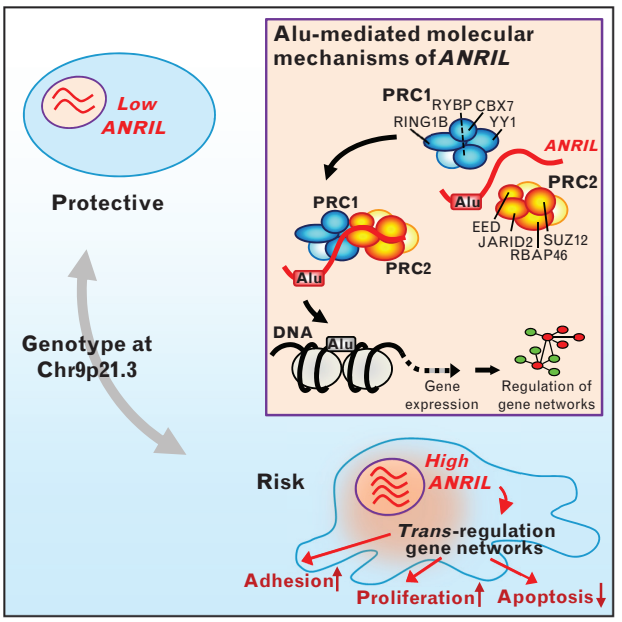

FIGURE 3. Model of *ANRIL/CDKN2B-AS1* function at Chr9p21 according to [59▪]. The atherosclerosis risk allele leads to up-regulation of the long ncRNA *ANRIL*. Increased *ANRIL* expression modulates networks of genes *in-trans*, leading to pro-atherogenic cell properties (increased cell adhesion, increased proliferation, decreased apoptosis). On the molecular level, *ANRIL* may act as a scaffold, guiding epigenetic modifier proteins of Polycomb repressive complexes 1 and 2 (PRC1, PRC2) and potentially others to chromatin. These functions depend on Alu motifs, which mark the promoters of *ANRIL* target genes and are mirrored in *ANRIL* RNA, suggesting an Alu-mediated RNA-DNA interaction as effector mechanism.

ANRIL, binding to epigenetic silencer Polycomb repressive complexes 1 and 2 (PRC1 and PRC2) [59▪,66,67] and to PRC-associated activating proteins RYBP and YY1 [71,72] has been demonstrated (Fig. 3) [59▪]. In accordance, modulation of *ANRIL* expression led to the epigenetic regulation of target genes expression in *cis* [66,67] and in *trans* [59▪,64,73]. We have recently shown that *trans*-regulation was dependent on an Alu-DEIN motif [74,75], which marked the promoters of *ANRIL* target genes and was mirrored in *ANRIL* RNA transcripts (Fig. 3). The functional relevance of Alu motifs in *ANRIL* was confirmed by deletion and mutagenesis, reversing *trans*-regulation and restoring normal cellular functions [59▪]. Recent work by Jeck *et al.* has also demonstrated that Alu motifs are preferably incorporated in noncoding RNA lariats, which might represent inactive isoforms and were also shown to exist for *ANRIL* [55,76]. Whether integration of Alu motifs in ncRNA lariats leads to silencing of the effector sequences remains to be determined.

In summary, the robust association of *ANRIL* with the risk genotype, its correlation with atherosclerosis severity, and functional data strongly support *ANRIL* as Chr9p21.3 effector gene. Recent work has not only broadened our understanding of *ANRIL's* function but also suggested a novel molecular mechanism for long ncRNA-mediated *trans*-regulation.

Chr6P24.1 (*PHACTR1*): FREQUENTLY REPLICATED BUT POORLY UNDERSTOOD

Chr6p24.1 is the second most often identified GWAS hit for CHD and MI. The locus was found in European, Asian, and Middle Eastern populations and therefore appears to be relevant across ethnicities [9▪,10▪▪,15,16,18▪▪,19]. Chr6p24.1 is also associated with coronary calcification [40]. Until now, virtually nothing is known about the mechanism of Chr6p24.1 in atherogenesis. The region contains a single gene, *protein phosphatase and actin regulator 1* (*PHACTR1*), spanning a very large genomic distance of ~500 kb, and extending over three haplotype blocks (Fig. 2b). Lead SNPs for CHD and MI are in the proximal haplotype block and the same SNPs were independently identified in a GWAS for migraine [42]. Intriguingly, alleles conferring migraine susceptibility were also associated with risk for CHD suggesting a common pathophysiology. The distal haplotype block of *PHACTR1* also contains hits in the GWAS catalogue (www.genome. gov/gwastudies; accessed May 2013; [21]), originating from a 100k GWAS for femoral neck width in females of the Framingham Heart Study [43] and a linkage study for pulse pressure in 63 Chinese sib-pairs [44] (Fig. 2b). However, these findings have not been firmly replicated and their significance is still unclear. In addition, these SNPs are ~300-kb apart and seemingly unrelated to the lead atherosclerosis SNPs, speaking against a causal relation.

PHACTR1 is highest expressed in human heart and brain [77] and is a member of a family of proteins that bind actin and interact with protein phosphatase 1 (PP1) [78]. PP1 is an ubiquitous enzyme, regulating essential cellular processes such as cell cycle progression, protein synthesis, muscle contraction, carbohydrate metabolism, transcription, and neuronal signaling (reviewed in [79]). For *PHACTR1*, a role in cell migration, motility and invasiveness of breast cancer, and melanoma tumor cells was described [80,81]. Moreover, *PHACTR1* is expressed in endothelial cells and involved in regulation of endothelial tubulogenesis and apoptosis [82,83]. In summary, even though *PHACTR1* is an obvious candidate gene at Chr6p24.1, current data on its function is scarce and its mechanism in atherogenesis is still unclear.

Chr1p13.3 (*PSRC1/CELSR2/MYBPHL/SORT1*): LIPIDS AND CORONARY HEART DISEASE

The Chr1p13.3 locus has been discovered in the first surge of GWAS for CHD even before it was also identified as one of the top GWAS hits for plasma LDL cholesterol concentrations [84–86]. Genetic variation at the locus is associated with reduced plasma LDL-cholesterol and reduced risk of coronary artery disease [10▪▪,25,45] suggesting that Chr1p13.3 exerts its effect on atherosclerosis by modulating LDL-cholesterol levels. The lead SNPs of CHD and LDL-cholesterol are located in a haplotype block encoding three genes, *cadherin EGF LAG seven-pass G-type receptor 2* (*CELSR2*), *proline/serine-rich coiled-coil 1* (*PSRC1*), and *myosin binding protein H-like* (*MYBPHL*) (Fig. 2c). Wild *et al.* found differential expression of *PSRC1* in monocytes at the locus [11]. The majority of functional work, however, has focused on *sortilin 1* (*SORT1*), which is located in a haplotype block distal of *PSRC1*, *CELSR2*, and *MYBPHL* (Fig. 2c) containing GWAS hits for major depressive disorder [46] and chronic kidney disease [47]. Schadt *et al.* [87] and Folkersen *et al.* [28] found that mRNA expression of *CELSR2*, *PSRC1*, and *SORT1* were all strongly associated with Chr1p13.3 in liver. Although *SORT1* was highly expressed in many tissues, genotype-dependent differential regulation was only seen in liver [28]. Musunuru *et al.* [26] identified a SNP in linkage disequilibrium with the lead SNP, creating a C/EBP transcription factor binding site in the 3' UTR of *CELSR2* and altering expression of *SORT1*. These data suggested that *SORT1* expression might be affected by *cis*-regulation through the neighboring haplotype block [26].

SORT1 is a member of the VSP10P receptor family of sorting receptors, which have been intensively studied in neuroscience and direct proteins through secretory and endocytic pathways of the cell (for review see [88,89]). In 2010, three independent groups published first mechanistic work on the role of *SORT1* in LDL-metabolism with in part paradoxical results: The first study overexpressed *SORT1* in HEK293 cells, resulting in increased uptake of LDL and LDL-receptor-related protein [25]. A second article used viral overexpression in mouse liver, demonstrating that increased *SORT1* decreased plasma LDL-cholesterol and VLDL levels by reducing hepatic VLDL secretion [26]. Inverse results were seen after *SORT1* knock-down [26]. Both studies were well in line with the observation that increased expression of *SORT1* mRNA in human liver was correlated with decreased LDL-cholesterol [26], even though the proposed mechanisms would be either through increased LDL uptake [25] or reduced

VLDL secretion [26]. Results of a third article, published virtually at the same time, were seemingly at odds with the two previous articles. Using mice on the *Ldlr*$^{-/-}$ background, these authors demonstrated that complete *Sort1* deficiency ameliorated hypercholesterolemia and atherosclerosis [24]. Additional studies on the subcellular level indicated that SORT1 interacts with apoB100 in the Golgi apparatus, thereby facilitating formation and hepatic export of apolipoprotein B containing lipoproteins [24].

Recent work [27▪] has reconciled the divergent hypotheses on the function of *SORT1* in lipoprotein metabolism. These authors proposed a model in which hepatic SORT1 binds intracellular apoB100 containing particles in the Golgi as well as extracellular LDL at the plasma membrane and traffics them to lysosomal degradation. They suggested a hyperbolic relationship in which complete lack as well as increased *SORT1* would both lead to a reduction in apoB and VLDL secretion, whereas intermediate *SORT1* expression would increase secretion [27▪]. Although common variants in *SORT1* have subtle effects on LDL-cholesterol, a recent publication provided data speaking against a role of *SORT1* missense mutations in autosomal dominant hypercholesterolemia [90].

Until now, the majority of work on the molecular mechanism at Chr1p13.3 has clearly focused on *SORT1*. Very little is known about the functions of *PSRC1*, *CELSR2*, and *MYBPHL*, which are closer to the lead Chr1p13.3 SNPs. More work is clearly warranted to establish or firmly exclude a role of these genes in lipid metabolism and atherogenesis.

CONCLUSION

Current GWAS have added additional loci to the 'genomic landscape' of CHD and MI bringing the total count to 58 at a significance cutoff of $P < 10^{-7}$. Recent advances in functional characterization of some loci promise the discovery of hitherto unknown pathways influencing atherosclerosis risk. One such example is the most replicated locus on Chr9p21.3, which might influence atherogenesis through epigenetic chromatin modification by the long ncRNA *ANRIL*. Nevertheless, our current understanding of potential causal variants and mechanisms at most GWAS loci of atherosclerotic cardiovascular disease is very limited. Although some of these loci co-segregate with known risk factors suggesting a potential causal relation, the majority is still 'terra incognita'. This is exemplified by the second most frequently found locus on Chr6p24.1, where virtually nothing is known about its function in atherogenesis. Owing to their small effect size, the utility of genetic variants for diagnostic purposes is limited. The major promise of identified GWAS loci therefore lies in understanding their function in atherogenesis as a prerequisite for later therapeutic targeting.

Acknowledgements

None.

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:
- ▪ of special interest
- ▪▪ of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (p. 456).

1. Helgadottir A, Thorleifsson G, Manolescu A, *et al.* A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007; 316:1491–1493.
2. McPherson R, Pertsemlidis A, Kavaslar N, *et al.* A common allele on chromosome 9 associated with coronary heart disease. Science 2007; 316:1488–1491.
3. Samani NJ, Erdmann J, Hall AS, *et al.* Genome-wide association analysis of coronary artery disease. N Engl J Med 2007; 357:443–453.
4. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447:661–678.
5. Tregouet DA, Konig IR, Erdmann J, *et al.* Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. Nat Genet 2009; 41:283–285.
6. Erdmann J, Grosshennig A, Braund PS, *et al.* New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat Genet 2009; 41:280–282.
7. Erdmann J, Willenborg C, Nahrstaedt J, *et al.* Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. Eur Heart J 2011; 32:158–168.
8. Reilly MP, Li M, He J, *et al.* Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. Lancet 2011; 377:383–392.
9. ▪ Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat Genet 2011; 43:339–344.

This is a large GWAS of coronary artery disease in Europeans and South Asians by the C4D consortium, adding five new loci of coronary artery disease and systematically testing for eQTLs at these loci.

10. ▪▪ Schunkert H, Konig IR, Kathiresan S, *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011; 43:333–338.

This is the currently largest meta-analysis of 14 GWAS of coronary artery disease by the CARDIoGRAM consortium, comprising 22 233 cases and 64 762 controls of European descent with follow-up in additional 56 682 individuals. The study confirmed 10 previously reported loci and added 13 new loci.

11. Wild PS, Zeller T, Schillert A, *et al.* A genome-wide association study identifies LIPA as a susceptibility gene for coronary artery disease. Circ Cardiovasc Genet 2011; 4:403–412.
12. Slavin TP, Feng T, Schnell A, *et al.* Two-marker association tests yield new disease associations for coronary artery disease and hypertension. Hum Genet 2011; 130:725–733.
13. Takeuchi F, Yokota M, Yamamoto K, *et al.* Genome-wide association study of coronary artery disease in the Japanese. Eur J Hum Genet 2012; 20:333–340.
14. Davies RW, Wells GA, Stewart AF, *et al.* A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. Circ Cardiovasc Genet 2012; 5:217–225.
15. Hager J, Kamatani Y, Cazier JB, *et al.* Genome-wide association study in a Lebanese cohort confirms PHACTR1 as a major determinant of coronary artery stenosis. PLoS One 2012; 7:e38663.
16. Lu X, Wang L, Chen S, *et al.* Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. Nat Genet 2012; 44:890–894.

17. Lee JY, Lee BS, Shin DJ, *et al.* A genome-wide association study of a coronary artery disease risk variant. J Hum Genet 2013; 58:120–126.

18. CARDIoGRAMplusC4D Consortium. Large-scale association analysis ■■ identifies new risk loci for coronary artery disease. Nat Genet 2013; 45:25–33.

CARDIoGRAMplusC4D consortium follow-up of the top 6226 SNPs with $P < 0.01$ from CARDIoGRAM in 63 746 coronary artery disease cases and 130 681 controls, adding 15 additional loci to the list. The study systematically tested for overlap with risk factors for coronary artery disease.

19. Kathiresan S, Voight BF, Purcell S, *et al.* Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 2009; 41:334–341.

20. Aoki A, Ozaki K, Sato H, *et al.* SNPs on chromosome 5p15.3 associated with myocardial infarction in Japanese population. J Hum Genet 2011; 56:47–51.

21. Hindorff LA, Sethupathy P, Junkins HA, *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 2009; 106:9362–9367.

22. Lanktree MB, Dichgans M, Hegele RA. Advances in genomic analysis of stroke: what have we learned and where are we headed? Stroke 2010; 41:825–832.

23. Murabito JM, White CC, Kavousi M, *et al.* Association between chromosome 9p21 variants and the ankle-brachial index identified by a meta-analysis of 21 genome-wide association studies. Circ Cardiovasc Genet 2012; 5:100–112.

24. Kjolby M, Andersen OM, Breiderhoff T, *et al.* Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. Cell Metab 2010; 12:213–223.

25. Linsel-Nitschke P, Heeren J, Aherrahrou Z, *et al.* Genetic variation at chromosome 1p13.3 affects sortilin mRNA expression, cellular LDL-uptake and serum LDL levels which translates to the risk of coronary artery disease. Atherosclerosis 2010; 208:183–189.

26. Musunuru K, Strong A, Frank-Kamenetsky M, *et al.* From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature 2010; 466:714–719.

27. Strong A, Ding Q, Edmondson AC, *et al.* Hepatic sortilin regulates both ■ apolipoprotein B secretion and LDL catabolism. J Clin Invest 2012; 122:2807–2816.

An elegant study reconciling paradoxical hypotheses about *SORT1* in modulating plasma LDL-cholesterol.

28. Folkersen L, van't Hooft F, Chernogubova E, *et al.* Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. Circ Cardiovasc Genet 2010; 3:365–373.

29. Zeller T, Wild P, Szymczak S, *et al.* Genetics and beyond–the transcriptome of human monocytes and disease susceptibility. PLoS One 2010; 5:e10693.

30. Holdt LM, Teupser D. Recent studies of the human chromosome 9p21 locus, which is associated with atherosclerosis in human populations. Arterioscler Thromb Vasc Biol 2012; 32:196–206.

31. Roberts R, Stewart AF. 9p21 and the genetic revolution for coronary artery disease. Clin Chem 2012; 58:104–112.

32. Gieger C, Radhakrishnan A, Cvejic A, *et al.* New gene functions in megakaryopoiesis and platelet formation. Nature 2011; 480:201–208.

33. Ramdas WD, van Koolwijk LM, Ikram MK, *et al.* A genome-wide association study of optic disc parameters. PLoS Genet 2010; 6:e1000978.

34. Sanson M, Hosking FJ, Shete S, *et al.* Chromosome 7p11.2 (EGFR) variation influences glioma risk. Hum Mol Genet 2011; 20:2897–2904.

35. Wiggs JL, Yaspan BL, Hauser MA, *et al.* Common variants at 9p21 and 8q22 are associated with increased susceptibility to optic nerve degeneration in glaucoma. PLoS Genet 2012; 8:e1002654.

36. Bei JX, Li Y, Jia WH, *et al.* A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. Nat Genet 2010; 42:599–603.

37. Turnbull C, Ahmed S, Morrison J, *et al.* Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet 2010; 42:504–507.

38. Yasuno K, Bilguvar K, Bijlenga P, *et al.* Genome-wide association study of intracranial aneurysm identifies three new risk loci. Nat Genet 2010; 42:420–425.

39. Uno S, Zembutsu H, Hirasawa A, *et al.* A genome-wide association study identifies genetic variants in the CDKN2BAS locus associated with endometriosis in Japanese. Nat Genet 2010; 42:707–710.

40. O'Donnell CJ, Kavousi M, Smith AV, *et al.* Genome-wide association study for coronary artery calcification with follow-up in myocardial infarction. Circulation 2011; 124:2855–2864.

41. Zeggini E, Scott LJ, Saxena R, *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008; 40:638–645.

42. Freilinger T, Anttila V, de Vries B, *et al.* Genome-wide association analysis identifies susceptibility loci for migraine without aura. Nat Genet 2012; 44:777–782.

43. Kiel DP, Demissie S, Dupuis J, *et al.* Genome-wide association with bone mass and geometry in the Framingham Heart Study. BMC Med Genet 2007; 8 (Suppl 1):S1.

44. Zhang D, Pang Z, Li S, *et al.* Genome-wide linkage and association scans for pulse pressure in Chinese twins. Hypertens Res 2012; 35:1051–1057.

45. Teslovich TM, Musunuru K, Smith AV, *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010; 466:707–713.

46. Kottgen A, Pattaro C, Boger CA, *et al.* New loci associated with kidney function and chronic kidney disease. Nat Genet 2010; 42:376–384.

47. Shi J, Potash JB, Knowles JA, *et al.* Genome-wide association study of recurrent early-onset major depressive disorder. Mol Psychiatry 2011; 16:193–201.

48. Kim WY, Sharpless NE. The regulation of INK4/ARF in cancer and aging. Cell 2006; 127:265–275.

49. Holdt LM, Sass K, Gabel G, *et al.* Expression of Chr9p21 genes CDKN2B (p15(INK4b)), CDKN2A (p16(INK4a), p14(ARF)) and MTAP in human atherosclerotic plaque. Atherosclerosis 2011; 214:264–270.

50. Kim JB, Deluna A, Mungrue IN, *et al.* Effect of 9p21.3 coronary artery disease ■ locus neighboring genes on atherosclerosis in mice. Circulation 2012; 126:1896–1906.

This is a systematic analysis of knockout models of all relevant protein-coding genes at the Chr9p21 locus of atherosclerosis.

51. Leeper NJ, Raiesdana A, Kojima Y, *et al.* Loss of CDKN2B promotes p53-dependent smooth muscle cell apoptosis and aneurysm formation. Arterioscler Thromb Vasc Biol 2013; 33:e1–e10.

52. Fuster JJ, Molina-Sanchez P, Jovani D, *et al.* Increased gene dosage of the Ink4/Arf locus does not attenuate atherosclerosis development in hypercholesterolaemic mice. Atherosclerosis 2012; 221:98–105.

53. Kuo CL, Murphy AJ, Sayers S, *et al.* Cdkn2a is an atherosclerosis modifier locus that regulates monocyte/macrophage proliferation. Arterioscler Thromb Vasc Biol 2011; 31:2483–2492.

54. Wouters K, Cudejko C, Gijbels MJ, *et al.* Bone marrow p16INK4a-deficiency does not modulate obesity, glucose homeostasis or atherosclerosis development. PLoS One 2012; 7:e32440.

55. Burd CE, Jeck WR, Liu Y, *et al.* Expression of linear and novel circular forms of an INK4/ARF-associated noncoding RNA correlates with atherosclerosis risk. PLoS Genet 2010; 6:e1001233.

56. Cunnington MS, Santibanez Koref M, Mayosi BM, *et al.* Chromosome 9p21 SNPs associated with multiple disease phenotypes correlate with ANRIL expression. PLoS Genet 2010; 6:e1000899.

57. Folkersen L, Kyriakou T, Goel A, *et al.* Relationship between CAD risk genotype in the chromosome 9p21 locus and gene expression. Identification of eight new ANRIL splice variants. PLoS One 2009; 4:e7677.

58. Holdt LM, Beutner F, Scholz M, *et al.* ANRIL expression is associated with atherosclerosis risk at chromosome 9p21. Arterioscler Thromb Vasc Biol 2010; 30:620–627.

59. Holdt LM, Hoffmann S, Sass K, *et al.* Alu elements in ANRIL non-coding RNA ■ at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks. PLoS Genet 2013; 9:e1003588.

This is a comprehensive study of molecular mechanism of *ANRIL* ncRNA at Chr9p21 in *trans*-regulation and pro-atherogenic cell functions.

60. Jarinova O, Stewart AF, Roberts R, *et al.* Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. Arterioscler Thromb Vasc Biol 2009; 29:1671–1677.

61. Johnson AD, Hwang SJ, Voorman A, *et al.* Resequencing and clinical associations of the 9p21.3 region: a comprehensive investigation in the Framingham heart study. Circulation 2013; 127:799–810.

62. Liu Y, Sanoff HK, Cho H, *et al.* INK4/ARF transcript expression is associated with chromosome 9p21 variants linked to atherosclerosis. PLoS One 2009; 4:e5027.

63. Harismendy O, Notani D, Song X, *et al.* 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. Nature 2011; 470:264–268.

64. Congrains A, Kamide K, Katsuya T, *et al.* CVD-associated noncoding RNA, ANRIL, modulates expression of atherogenic pathways in VSMC. Biochem Biophys Res Commun 2012; 419:612–616.

65. Congrains A, Kamide K, Oguro R, *et al.* Genetic variants at the 9p21 locus contribute to atherosclerosis through modulation of ANRIL and CDKN2A/B. Atherosclerosis 2012; 220:449–455.

66. Kotake Y, Nakagawa T, Kitagawa K, *et al.* Long noncoding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. Oncogene 2011; 30:1956–1962.

67. Yap KL, Li S, Munoz-Cabello AM, *et al.* Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol Cell 2010; 38:662–674.

68. Hung T, Chang HY. Long noncoding RNA in genome regulation: prospects and mechanisms. RNA Biol 2010; 7:582–585.

69. Lee JT. Epigenetic regulation by long noncoding RNAs. Science 2012; 338:1435–1439.

70. Wapinski O, Chang HY. Long noncoding RNAs and human disease (vol 21, pg 354, 2011). Trends in Cell Biology 2011; 21:561–1561.

71. Gonzalez I, Busturia A. High levels of dRYBP induce apoptosis in Drosophila imaginal cells through the activation of reaper and the requirement of trithorax, dredd and dFADD. Cell Res 2009; 19:747–757.

72. Gregoire S, Karra R, Passer D, *et al.* Essential and unexpected role of YY1 to promote mesodermal cardiac differentiation. Circ Res 2013; 112:900–910.

73. Sato K, Nakagawa H, Tajima A, *et al.* ANRIL is implicated in the regulation of nucleus and potential transcriptional target of E2F1. Oncol Rep 2010; 24:701–707.

74. Deininger PL, Jolly DJ, Rubin CM, *et al.* Base sequence studies of 300 nucleotide renatured repeated human DNA clones. J Mol Biol 1981; 151:17–33.

75. Weibrecht I, Gavrilovic M, Lindbom L, *et al.* Visualising individual sequence-specific protein-DNA interactions in situ. N Biotechnol 2012; 29:589–598.

76. Jeck WR, Sorrentino JA, Wang K, *et al.* Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA 2013; 19:141–157.

77. Nagase T, Kikuno R, Hattori A, *et al.* Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. DNA Res 2000; 7:347–355.

78. Allen PB, Greenfield AT, Svenningsson P, *et al.* Phactrs 1-4: A family of protein phosphatase 1 and actin regulatory proteins. Proc Natl Acad Sci U S A 2004; 101:7187–7192.

79. Peti W, Nairn AC, Page R. Structural basis for protein phosphatase 1 regulation and specificity. FEBS J 2013; 280:596–611.

80. Fils-Aime N, Dai M, Guo J, *et al.* MicroRNA-584 and the protein phosphatase and actin regulator 1 (PHACTR1), a new signaling route through which transforming growth factor-beta mediates the migration and actin dynamics of breast cancer cells. J Biol Chem 2013; 288:11807–11823.

81. Wiezlak M, Diring J, Abella J, *et al.* G-actin regulates the shuttling and PP1 binding of the RPEL protein Phactr1 to control actomyosin assembly. J Cell Sci 2012; 125:5860–5872.

82. Allain B, Jarray R, Borriello L, *et al.* Neuropilin-1 regulates a new VEGF-induced gene, Phactr-1, which controls tubulogenesis and modulates lamellipodial dynamics in human endothelial cells. Cell Signal 2012; 24:214–223.

83. Jarray R, Allain B, Borriello L, *et al.* Depletion of the novel protein PHACTR-1 from human endothelial cells abolishes tube formation and induces cell death receptor apoptosis. Biochimie 2011; 93:1668–1675.

84. Kathiresan S, Melander O, Guiducci C, *et al.* Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat Genet 2008; 40:189–197.

85. Sandhu MS, Waterworth DM, Debenham SL, *et al.* LDL-cholesterol concentrations: a genome-wide association study. Lancet 2008; 371:483–491.

86. Willer CJ, Sanna S, Jackson AU, *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet 2008; 40:161–169.

87. Schadt EE, Molony C, Chudin E, *et al.* Mapping the genetic architecture of gene expression in human liver. PLoS Biol 2008; 6:e107.

88. Willnow TE, Kjolby M, Nykjaer A. Sortilins: new players in lipoprotein metabolism. Curr Opin Lipidol 2011; 22:79–85.

89. Willnow TE, Petersen CM, Nykjaer A. VPS10P-domain receptors - regulators of neuronal viability and function. Nat Rev Neurosci 2008; 9:899–909.

90. Tveten K, Strom TB, Cameron J, *et al.* Mutations in the SORT1 gene are unlikely to cause autosomal dominant hypercholesterolemia. Atherosclerosis 2012; 225:370–375.