Circulation Research

ORIGINAL RESEARCH

Genetic Regulation of SMC Gene Expression and Splicing Predict Causal CAD Genes

Rédouane Aherrahrou, Dillon Lue, R. Noah Perry, Yonathan Tamrat Aberra, Mohammad Daud Khan, Joon Yuhl Soh, Tiit Örd, Prosanta Singha, Qianyi Yang, Huda Gilani, Ernest Diez Benavente, Doris Wong, Jameson Hinkle, Lijiang Ma, Gloria M. Sheynkman, Hester M. den Ruijter, Clint L. Miller, Johan L.M. Björkegren, Minna U. Kaikkonen, Mete Civelek

BACKGROUND: Coronary artery disease (CAD) is the leading cause of death worldwide. Recent meta-analyses of genome-wide association studies have identified over 175 loci associated with CAD. The majority of these loci are in noncoding regions and are predicted to regulate gene expression. Given that vascular smooth muscle cells (SMCs) play critical roles in the development and progression of CAD, we aimed to identify the subset of the CAD loci associated with the regulation of transcription in distinct SMC phenotypes.

METHODS: We measured gene expression in SMCs isolated from the ascending aortas of 151 heart transplant donors of various genetic ancestries in quiescent or proliferative conditions and calculated the association of their expression and splicing with ~6.3 million imputed single-nucleotide polymorphism markers across the genome.

RESULTS: We identified 4910 expression and 4412 splicing quantitative trait loci (sQTLs) representing regions of the genome associated with transcript abundance and splicing. A total of 3660 expression quantitative trait loci (eQTLs) had not been observed in the publicly available Genotype-Tissue Expression dataset. Further, 29 and 880 eQTLs were SMC-specific and sex-biased, respectively. We made these results available for public query on a user-friendly website. To identify the effector transcript(s) regulated by CAD loci, we used 4 distinct colocalization approaches. We identified 84 eQTL and 164 sQTL that colocalized with CAD loci, highlighting the importance of genetic regulation of mRNA splicing as a molecular mechanism for CAD genetic risk. Notably, 20% and 35% of the eQTLs were unique to quiescent or proliferative SMCs, respectively. One CAD locus colocalized with a sex-specific eQTL (*TERF2IP*), and another locus colocalized with SMC-specific eQTL (*ALKBH8*). The most significantly associated CAD locus, 9p21, was an sQTL for the long noncoding RNA *CDKN2B-AS1*, also known as *ANRIL*, in proliferative SMCs.

CONCLUSIONS: Collectively, our results provide evidence for the molecular mechanisms of genetic susceptibility to CAD in distinct SMC phenotypes.

GRAPHIC ABSTRACT: A graphic abstract is available for this article.

Key Words: atherosclerosis ■ coronary artery disease ■ expression and splice quantitative trait loci ■ gene expression ■ smooth muscle cell

Meet the First Author, see p 253

Coronary artery disease (CAD) is the leading cause of death worldwide.[1] Heritability estimates for CAD vary between 40% and 70%, suggesting a strong genetic contribution to disease pathology.[2] Genome-wide association studies (GWAS) have identified 175 loci associated with increased risk for CAD.[3–5] Approximately 40% of the CAD loci are associated with known risk factors, such as blood lipid levels, nitric oxide signaling, and blood pressure.[3] The remaining 60% have unknown mechanisms, but there is evidence that some of these loci function through the vessel wall where the disease develops.[6] In addition, 94% of CAD-associated genetic

Correspondence to: Rédouane Aherrahrou, PhD, University of Virginia, Center for Public Health Genomics, Old Medical School 3836, PO Box 800717, Charlottesville, VA 22908-0717, Email ra2qy@virginia.edu or Mete Civelek, PhD, University of Virginia, Center for Public Health Genomics, Old Medical School 3836, PO Box 800717, Charlottesville, VA 22908-0717, Email mete@virginia.edu

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/CIRCRESAHA.122.321586.
For Sources of Funding and Disclosures, see page 336.

Circulation Research is available at www.ahajournals.org/journal/res

ORIGINAL RESEARCH

Novelty and Significance

What Is Known?

- Genome-wide association studies (GWAS) identified 175 loci associated with coronary artery disease (CAD).
- Most of these loci are in noncoding regions, have unknown mechanisms, but are predicted to regulate gene expression.
- The most significantly associated CAD GWAS locus, 9p21, remains a mystery after many years of studies.
- Vascular smooth muscle cells (SMCs) play critical roles in the development and progression of CAD.

What New Information Does This Article Contribute?

- We identify 84 genes whose expression and 164 genes whose splicing are regulated by loci associated with increased risk for CAD.
- We discover distinct genetic architectures of gene expression in quiescent and proliferative SMC phenotypes.
- We predict that long noncoding RNA SNHG18 is a likely causal transcript for CAD and affects SMC proliferation.
- We identify the colocalization of TERF2IP with a sex-biased CAD locus.
- We show that the 9p21, the most significantly associated CAD locus, affects the splicing of the long noncoding RNA CDKN2B-AS1, also known as ANRIL, in SMCs.

CAD is the leading cause of death worldwide. Recent meta-analyses of GWAS have identified 175 loci associated with CAD. Given that vascular SMCs play critical roles in the development and progression of CAD, we hypothesized that a subset of the CAD GWAS risk loci are associated with the regulation of transcription in distinct SMC phenotypes. We identified 4910 expression quantitative trait loci (eQTL) and 4412 splicing quantitative trait loci (sQTL) that represent regions of the genome associated with transcript abundance and splicing. A total of 3660 of the eQTLs had not been observed in the publicly available Genotype-Tissue Expression dataset. We identified 84 eQTL and 164 sQTLs that colocalized with CAD loci. Notably, 20% and 35% of the eQTLs were unique to quiescent or proliferative SMCs, respectively. A CAD locus colocalized with a sex-specific eQTL, TERF2IP, and another locus colocalized with SMC-specific eQTL, ALKBH8. The most significantly associated CAD locus, 9p21, was an sQTL for the long noncoding RNA CDKN2B-AS1, also known as ANRIL, in proliferative SMCs. We created a user-friendly website (https://civeleklab.cphg.virginia.edu) for cardiovascular researchers to query our dataset. Our study provided evidence for the molecular mechanisms of genetic susceptibility to CAD in distinct SMC phenotypes.

Nonstandard Abbreviations and Acronyms

ATACseq	assay for transposase-accessible chromatin with high-throughput sequencing
CAD	coronary artery disease
eQTL	expression quantitative trait locus
GTEx	Genotype-Tissue Expression
GWAS	genome-wide association studies
LD	linkage disequilibrium
RBP	RNA-binding protein
SMC	smooth muscle cells
SNP	single-nucleotide polymorphism
sQTL	splicing quantitative trait loci

variants are in noncoding regions of the genome,[7] implying that the disease-causing loci involve regulatory mechanisms that affect the transcription of genes.[6] Therefore, gene expression studies performed in cells and tissues relevant to CAD in human populations can pinpoint the regulatory mechanisms of disease susceptibility.[8]

Expression quantitative trait loci (eQTL) and splicing quantitative trait loci (sQTL) analyses are key approaches that link genetic variants with variations in gene expression and splicing patterns, respectively.[8,9] They have enabled the prioritization of genetic variants within GWAS loci for different traits, and have shown that trait-associated genetic variants often function in a tissue- or cell type-specific manner.[10–13] Smooth muscle cells (SMCs), which make up the medial layer of arteries, play key roles in the integrity of the vessel wall, regulation of blood pressure, and initiation and development of atherosclerosis. Recent studies provided compelling evidence that SMCs can play either beneficial or detrimental roles in lesion pathogenesis depending on the nature of their phenotypic changes, for example from a quiescent to a proliferative phenotype.[14–17] In addition, SMC phenotypic switching seems to be important in explaining sex differences in atherosclerotic plaque composition.[18] Thus, identifying the genetic determinants of SMC gene expression is crucial for understanding the biological significance of CAD-associated genetic variants functioning in SMCs. This will also inform the prediction of novel drug candidates targeting the disease processes in the vessel wall.

Despite publicly available population-level gene expression datasets from various tissues and cells, including 54 tissues from the Genotype-Tissue Expression

(GTEx) project[19] and atherosclerosis-relevant tissues and cell types from the STARNET cohort (Stockholm-Tartu Atherosclerosis Reverse Networks Engineering Task),[20] aortic endothelial cells,[21,22] monocytes,[23] whole blood,[24] and coronary artery SMCs[25] as well as epigenome profiling from the Roadmap Epigenomics Project,[26] more than half of the CAD loci are still not functionally annotated. Therefore, in this study, using a multi-omics approach, we identified the genetic variants that are associated with SMC-specific gene expression derived from 151 healthy and ancestrally diverse heart transplant donors. This allowed us to identify CAD-associated loci, including the most significantly associated CAD locus, 9p21, that perturb SMC transcription, thereby leading to the prediction of candidate effector transcripts and candidate causal variants in these disease-associated loci.

METHODS

A detailed description of the methods and the experimental procedures are provided in the Supplemental Material. Please see the major resources table in the Supplemental Material.

Data Availability

The RNAseq and assay for transposase-accessible chromatin with high-throughput sequencing (ATACseq) data are available at GEO with the accession numbers GSE193817 and GSE198544, respectively. eQTL/sQTL results can be accessed at https://virginia.box.com/s/t5e1tzlaqsf85z13o4ie2f9t1i0zfypd and https://virginia.box.com/s/o81cxrj5xne3xem4au785mupikduuwbu, respectively. We have created a user-friendly website at https://civeleklab.cphg.virginia.edu to query the dataset published in this article.

RESULTS

Transcriptional Profiling of Human Aortic SMCs

We performed RNA sequencing of aortic SMCs derived from 151 ancestrally diverse healthy heart transplant donors (118 males and 33 females) to identify their transcriptional profiles. After quality control filtering, data analyses were performed on 139 and 145 samples cultured in the absence (quiescent) or presence (proliferative) of fetal bovine serum (FBS), respectively (Figure 1). There were 18 637 and 18 116 expressed genes in quiescent and proliferative SMCs, respectively (Table 1). Principal component analysis identified 2 distinct clusters of samples corresponding to the cells cultured in the 2 conditions (Figure S1A). Further, 2773 genes were differentially expressed ($P_{adj}<1\times10^{-3}$), including canonical SMC markers (*VCAM1*, *SMTN*, *ICAM1*, *TAGLN*, *CNN1*, *ACTA2*, *SPP1*) in agreement with the differences observed in the quiescent and proliferative state of SMCs (Figure S1B; Table S1). Extracellular matrix organization was among the top biological pathways identified in GO enrichment analysis of upregulated genes in proliferative versus quiescent phenotypes. In contrast, pathways that are associated with DNA replication and proliferation were repressed in the quiescent phenotype (Figure S1C).

To confirm that cultured SMCs reflect in vivo physiology, we projected their transcriptomes onto the 49 tissues profiled in version 8 of the GTEx project[28] (Figure S2). We observed that SMCs formed a distinct cluster and the closest tissue/cell type are fibroblasts, skeletal muscle, blood vessels, and heart.

cis-eQTL in SMCs

We obtained genotype information for ~6.3 million variants with at least 5% minor allele frequency in our population. Clustering of the donor genotypes with the 1000 Genomes reference population samples identified 6, 12, 64, and 69 of the individuals with East Asian, African, Admixed American, and European ancestry, respectively (Figure S3).

To identify genetic loci associated with transcript abundance, we performed association mapping with the genotypes of ~6.3 million variants and the expression levels of 18 116 and 18 637 genes in quiescent and proliferative SMCs, respectively, using tensorQTL.[29] We identified 3000 and 4188 eGenes with cis-eQTL (<1 Mb from the TSS) at false discovery rate (FDR) q-value <0.05 in the quiescent and proliferative phenotypes, respectively (Table 1). We identified that 1322 and 1899 eGenes from quiescent and proliferative SMCs, respectively, were differentially expressed between the 2 conditions ($P_{adj}<0.05$; Table S2). Next, we compared our results against the GTEx v8 eQTL dataset composed of 49 different human tissues[28] using the QTlizer R package.[30] We found that 2818 SMC cis-eQTLs (eSNP and eGene pair) were present in at least 1 GTEx tissue, whereas 3660 were unique to our dataset (Figure 2A). Most of the shared SMC eQTLs were enriched in the tissues that are anatomically rich in SMCs (Figure S4). Conditioning on the lead single-nucleotide polymorphisms (SNPs) identified 254 and 465 secondary and beyond eQTLs for quiescent and proliferative SMCs, respectively.

We identified SMC-specific eQTLs using GTEx tissues as a reference[31] with METASOFT.[30] We identified 29 SMC-specific cis-eQTLs under a stringent criteria of eQTL posterior probability >0.9 for SMCs and <0.1 for all the GTEx tissues (Table S3; Figure S5). To validate these results, we queried these 29 cis-eQTLs in the STARNET dataset of 7 cardiometabolic tissues from ~600 donors.[20] We identified only 2 of the SMC cis-eQTLs (rs367077-HLA-K and rs4795548-SH3GL1P2) to be present in the STARNET dataset at FDR <0.05 (Table S4). Twelve of the 29 SMC-specific cis-eQTLs were present in both quiescent and proliferative SMCs, 9 loci were cis-eQTLs only in quiescent, and 8 loci were cis-eQTLs only in proliferative SMCs. Because many tissues in GTEx contain vascular wall cells, we examined if

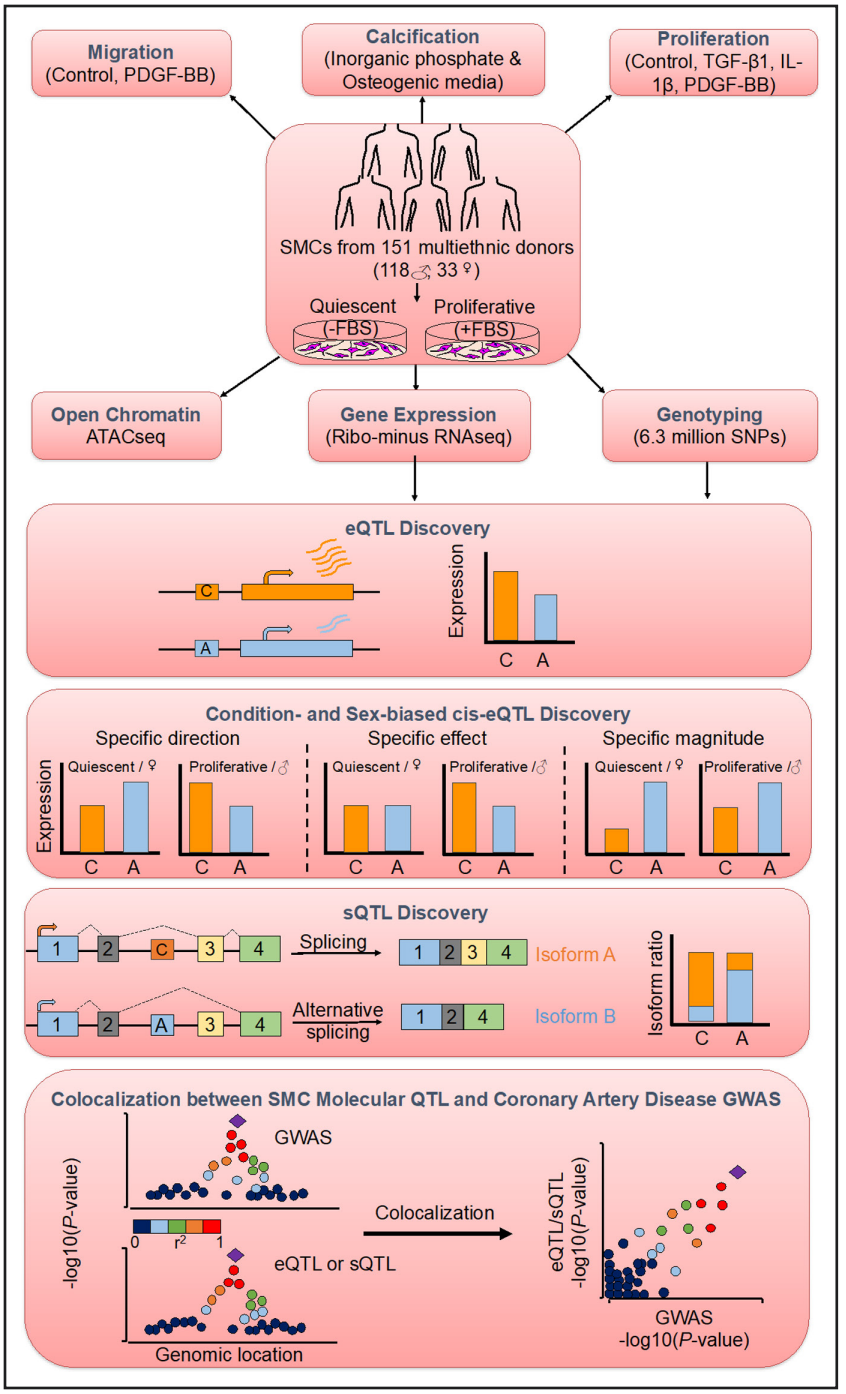

Figure 1. Study design and overview of analyses.
Aortic smooth muscle cells (SMCs) from 151 heart transplant donors of various genetic ancestries were characterized for 3 atherosclerosis-relevant phenotypes: migration, calcification, and proliferation.[27] To measure gene expression of SMCs, sequencing of ribosomal RNA-depleted total RNA isolated from SMCs cultured in the absence or presence of FBS to simulate quiescent or proliferative phenotypic state was performed. Associations of gene expression and splicing with the genotypes of ~6.3 million imputed SNPs were calculated to discover cis-eQTLs as well as condition-specific and sex-biased eQTLs and sQTLs. Colocalization between molecular QTL and coronary artery disease GWAS associations was identified using 4 different methods. CAD indicates coronary artery disease; eQTL, expression quantitative trait locus; GWAS, genome-wide association studies; SMC, smooth muscle cells; SNP, single-nucleotide polymorphism; and sQTL, splicing quantitative trait locus.

the 29 SMC eQTLs showed associations in monocytes/macrophages[23] and aortic endothelial cells.[21] None of the 29 eQTLs were present in these cells suggesting that the regulatory impact of the variants in these loci are SMC specific.

About 15% (3000) and 39% (1976) of the eGenes were unique to quiescent and proliferative cells, respectively (Figure 2A; $P_{overlap}<1\times10^{-300}$, hypergeometric test). Therefore, we determined whether the eQTL effect sizes were statistically different between the 2 phenotypic states. We compared the regression slopes of an eQTL in quiescent (β_{noFBS}) versus proliferative (β_{FBS}) phenotypes using a Z-test.[32] We identified 1248 eQTLs at FDR q-value <0.05 with varying effects between the 2 phenotypic states (Table S5). We classified the condition-specific eQTLs into 3 different categories (Figure 2B). Fifty-eight percent (527) of them showed differences in magnitude of the effect size between the 2 phenotypes, 39% (358) showed eQTL effect in only 1 phenotype, and 3% (25) showed differences in the direction of effect between the 2 phenotypes (Figure 2C). Three examples of condition-specific eQTLs are shown in Figure 2B. These results suggest that regulatory variation impacts SMC gene expression in specific contexts.

Table 1. Expression and Splicing Quantitative Trait Loci

Gene type	Gene expression		Splicing	
	No. of genes tested	No. of eGenes with at least one eQTL	No. of genes tested	No. of sGenes with at least 1 sQTL
Condition: quiescent				
Protein coding	13 341	2267	11 123	2 774
LncRNA	3183	477	959	244
Pseudogene	1484	215	291	129
Other	629	41	1	0
Total	18 637	3000	12 374	3147
Condition: proliferative				
Protein coding	13 182	3367	11 013	3180
LncRNA	2986	546	996	273
Pseudogene	1366	228	297	125
Other	582	47	1	0
Total	18 116	4188	12 307	3578

eQTL indicates expression quantitative trait locus; LncRNA, long noncoding RNAs; and sQTL, splicing quantitative trait locus.

Since sex differences in the genetic regulation of gene expression have been observed in many tissues,[33,34] we identified sex-biased cis-eQTLs in quiescent and proliferative SMCs. We identified 457 and 454 sex-biased cis-eQTLs (eigenMT value <0.05) in quiescent and proliferative SMCs, respectively (Table S6). Twenty-four percent of them (combined conditions) showed differences in magnitude of effect size between the 2 sexes, 46% showed eQTL effect in only 1 sex, and 30% showed differences in the direction of effect between the 2 sexes (Figure 2C). Three examples of sex-biased eQTLs are shown in Figure 2B.

To characterize the potential function of the eQTL signals, we evaluated the overlap of the lead variant and linkage disequilibrium (LD) proxies ($r^2 \geq 0.8$; 1000G EUR) in accessible chromatin regions of SMCs as identified by ATACseq in 5 different donors in quiescent and proliferative phenotypes. About 9.86% of the lead variants of the eQTLs were in these accessible chromatin regions. After including LD proxies of these lead variants, 1386 of 3000 eQTL signals in quiescent SMCs and 1945 of 4188 eQTL signals in proliferative SMCs overlapped accessible chromatin regions, demonstrating the potential regulatory function of these loci. We

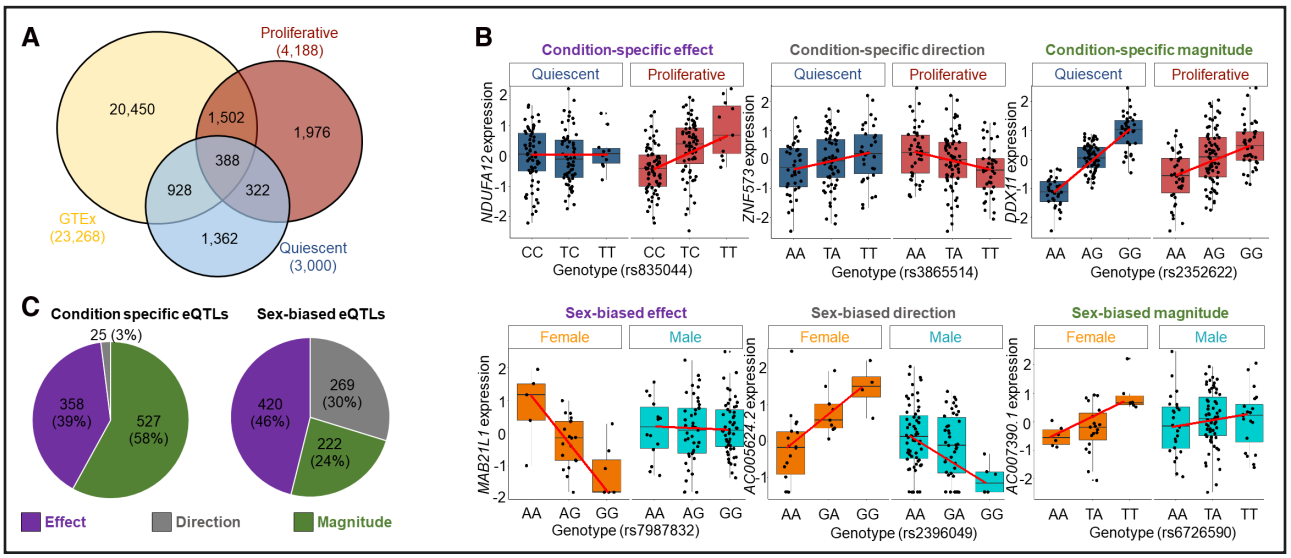

Figure 2. Identification of cis-eQTLs, condition-specific and sex-biased eQTLs in aortic smooth muscle cells.
A, Venn diagram comparing eQTL discovered in quiescent (blue) and proliferative (red) conditions vs GTEx tissues (yellow; false discovery rate [FDR] q-value <0.05). About 3660 of smooth muscle cell (SMC) cis-eQTLs (pair of SNP-gene) were absent or not significant in the GTEx dataset. A total of 1362 and 1976 of these novel cis-eQTLs were unique to quiescent and proliferative cells, respectively. **B**, Condition-specific (**top**) and sex-biased (**bottom**) eQTLs were classified into 3 different categories: condition-specific or sex-biased effect, condition-specific or sex-biased direction, and condition-specific or sex-biased magnitude. **C**, Quantification of the 3 different condition-specific (**left**) and sex-biased (**right**) eQTL groups. eQTL indicates expression quantitative trait locus; and GTEx, Genotype-Tissue Expression.

next identified transcription factor-binding sites over-represented in ATACseq peaks and overlapping eQTL SNPs in the accessible chromatin regions (Figure S6). Motif enrichment analysis showed enrichment of putative binding sites for members of the SP2 (Sp2 trancription factor), SP1 (Sp1 trancription factor), ELK4 (ETS transcription factor ELK4), GABPA (GA binding protein transcription factor subunit alpha) transcription factor (TF) families, some of which are known to play functional roles in SMCs.[35–37]

Colocalization Between eQTLs and CAD GWAS Signals

To predict the effector transcripts that are regulated by CAD loci, we performed colocalization analyses using 4 distinct approaches. Overall, we had genotypes for 169 of the 175 CAD loci in our dataset. First, to identify the eQTLs that were likely to be driven by the GWAS loci, we assessed if the GWAS and eQTL lead variants were in high LD ($r^2 \geq 0.8$) in our population. Using the LD colocalization approach only, we identified 16 and 22 eGenes (FDR q-value <0.05) in the quiescent and proliferative phenotype that showed an overlap with CAD loci, respectively. We also performed colocalization analysis using 3 additional methods: summary-data-based Mendelian randomization (SMR),[38] Colocalisation Tests of Two Genetic Traits (COLOC),[39] and eQTL and GWAS Causal Variants Identification in Associated Regions (eCAVIAR).[40] We used FDR q-value <0.05 for SMR and colocalization posterior probability >0.01 as cutoffs for eCAVIAR and posterior probability H4 (PPH4) >0.5 for COLOC. Using all 4 colocalization methods, we identified 84 eGenes that showed statistically significant colocalization. Seventeen and 30 of them showed an overlap with at least 2 of the colocalization methods in quiescent and proliferative SMCs, respectively (Figure 3; Table S7 and Figure S7). Some of the eGenes predicted by the 4 colocalization methods differed; therefore, we visualized the coincidence of the eQTL and GWAS lead SNPs by inspecting the regional colocalization plots using LocusCompare.[42] This coincidence was also supported by conditional analysis on each lead eSNP and CAD GWAS index SNP.

Next, we assessed the LD between the sex-biased eQTL and CAD GWAS SNPs and identified *TERF2IP* whose cis-eQTL colocalized with the 16q23.1 CAD locus (Figure 4). The sex-biased eQTL for *TERF2IP* was present only in proliferative SMCs. The CAD risk allele of the eSNP rs12929673, T, was associated with lower expression of *TERF2IP* in females and higher expression in males (Figure 4A). The same locus was also associated with 2 SMC phenotypes relevant to atherosclerosis in a sex-stratified manner.[27] The CAD risk allele of the eSNP rs12929673, T, was associated with reduced proliferation response to IL-1β stimulation and lower calcification in females compared to males which showed the opposite effect (Figure 4B). The CAD GWAS signal at this locus was stronger in males ($P_{rs12929673,male}=2\times10^{-6}$; $\beta_{rs12929673,male}=0.28$) than in females ($P_{rs12929673,female}=1\times10^{-2}$; $\beta_{rs2929673,female}=0.15$) in the UK Biobank cohort[43] (Figure 4C), suggesting a protective role for lower *TERF2IP* expression against atherosclerosis.

To identify SMC-specific genetic regulation of CAD risk, we asked if the SMC-specific eQTLs (Table S3) colocalized with CAD GWAS loci. We found that the cis-eQTL for *ALKBH8* colocalized with the 11q22.3 CAD locus (Figure 5A). This SMC-specific eQTL is also colocalized with systolic and diastolic blood pressure (Figure 5B and 5C), suggesting a role for this gene, which encodes a methyltransferase, in regulating blood pressure and atherosclerosis risk. The risk allele, G, of the eSNP rs7926602 is associated with lower expression of *ALKBH8* in quiescent and proliferative SMCs (Figure 5D and 5E).

Functional Annotation of CAD GWAS-Colocalized SMC eQTLs

The 84 colocalized eQTLs contain 3811 SNPs that are in high LD in our study population ($r^2 \geq 0.8$). We overlapped these SNPs with accessible chromatin regions identified in ATACseq experiments performed in quiescent and proliferative SMCs from 5 donors. We determined that 128 SNPs in 30 eQTL loci in quiescent SMCs and 140 SNPs in 37 eQTL loci in proliferative SMCs were in accessible chromatin peaks (Table S8). We predict these SNPs to have a regulatory impact on eQTL gene expression and potentially be causal for SMC gene expression and CAD risk.

We previously characterized the SMC donors for 12 atherosclerosis-relevant phenotypes.[27] First, we assessed the association of the eSNPs in the CAD loci with these phenotypes. Second, we assessed the correlation between the phenotypes and the colocalized eQTL gene expression. The risk allele of the eSNP rs7195958 at the *DHODH* locus was associated with increased SMC proliferation compared to the non-risk allele. The risk allele was also associated with decreased expression of *DHODH* in quiescent SMCs. As would be predicted by these association results, there was a significant positive correlation between *DHODH* expression and SMC proliferation, suggesting that the 16q22 CAD locus regulates SMC proliferation by perturbing *DHODH* expression (Figure S8A). Similarly, the risk allele of the eSNP rs12817989 at the *FGD6* locus had lower association with SMC proliferation compared to the non-risk allele. The risk allele was associated with higher expression of *FGD6* in proliferative SMCs. As would be predicted by these association results, there was a significant negative correlation between *FGD6* expression and SMC proliferation, suggesting that the 12q22 CAD locus regulates SMC proliferation by perturbing *FGD6* expression (Figure S8B).

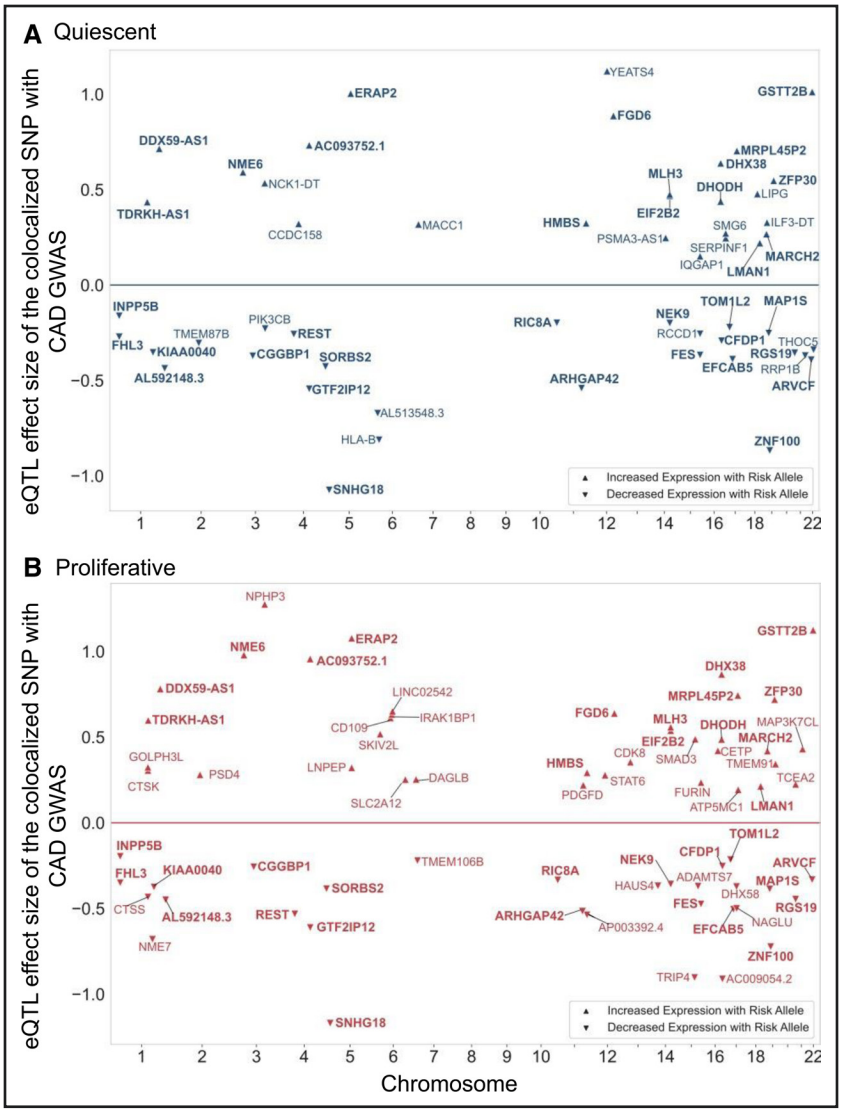

Figure 3. Summary of the smooth muscle cell (SMC) eQTL and CAD GWAS colocalization.
Plots show the colocalization of CAD GWAS and eQTL signals using a combination of 4 different methods in quiescent (**A**) and proliferative (**B**) SMCs. The *x* axis shows the chromosomal position of the colocalized SNP. The *y* axis shows the effect size and direction of the eQTL with respect to the risk allele from the coronary artery GWAS.[4] The effect size value is proportional to gene expression residual after PEER correction.[41] Risk alleles that are associated with increased gene expression level are shown with up-triangle and risk alleles that are associated with decreased gene expression level are shown with down-triangle. Bolded gene symbols are common between the 2 phenotypic states. CAD indicates coronary artery disease; eQTL, expression quantitative trait locus; GWAS, genome-wide association studies; and SNP, single-nucleotide polymorphism.

Finally, human single-cell RNAseq analysis from coronary atherosclerotic plaques confirmed the expression of most of the eQTL genes in SMCs (Figure S9).[17] We were able to assess the expression of 76 of the 84 eQTL genes in the coronary artery scRNA-Seq dataset and found that 50 were higher expressed in SMCs, pericytes, and fibroblasts compared to endothelial cells, monocytes, macrophages, and other immune cells. For example, we found that the expression of the long noncoding RNA SNHG18 was regulated by the 5p15 CAD locus in both the quiescent and proliferative SMCs. The risk allele was associated with decreased expression of SNHG18 (Figure 6A). Of the 16 potentially causal SNPs in high LD in this locus, 6 of them were in accessible chromatin regions in both SMC phenotypic states (Figure 6B and Table S8). Nine SNPs also displayed an allelic effect in a massively parallel reporter assay performed in SMCs exposed to cholesterol to induce phenotypic switching to resemble modulated SMCs found in atherosclerotic plaques[44] (Figure 6C). SNPs rs1651285, rs1706987, and rs1398337 were in both accessible regions and showed allelic effects in massively parallel reporter assay, identifying them as the potential causal variants at this locus. The expression of *SNHG18* was highly enriched in SMC, fibroblast/fibromyocyte, and pericyte clusters in a human carotid artery single-cell RNAseq dataset[17] (Figure 6D). Further, SNHG18 expression was negatively correlated with PDGF-BB-induced proliferation (Figure 6E). When we silenced the SNHG18 expression in immortalized human coronary aortic SMCs (Figure S10), we observed increased proliferation (Figure 6F).

To determine the transcriptional profile of *SNHG18* downregulation, we performed transcriptome analysis using RNAseq in human coronary artery SMCs and found 778 genes that were differentially expressed (P_{adj}<0.05; Table S9). Of these, 375 genes showed a $\log_2$(fold-change) above 0.2, and 403 genes were downregulated with $\log_2$(fold-change)<−0.2. Upregulation of *MKI67*, a cellular marker for proliferation, in response to *SNHG18* knockdown confirmed our proliferation results. Furthermore, proliferation/cell cycle, cell migration, cell motility, and blood vessel development/angiogenesis were among

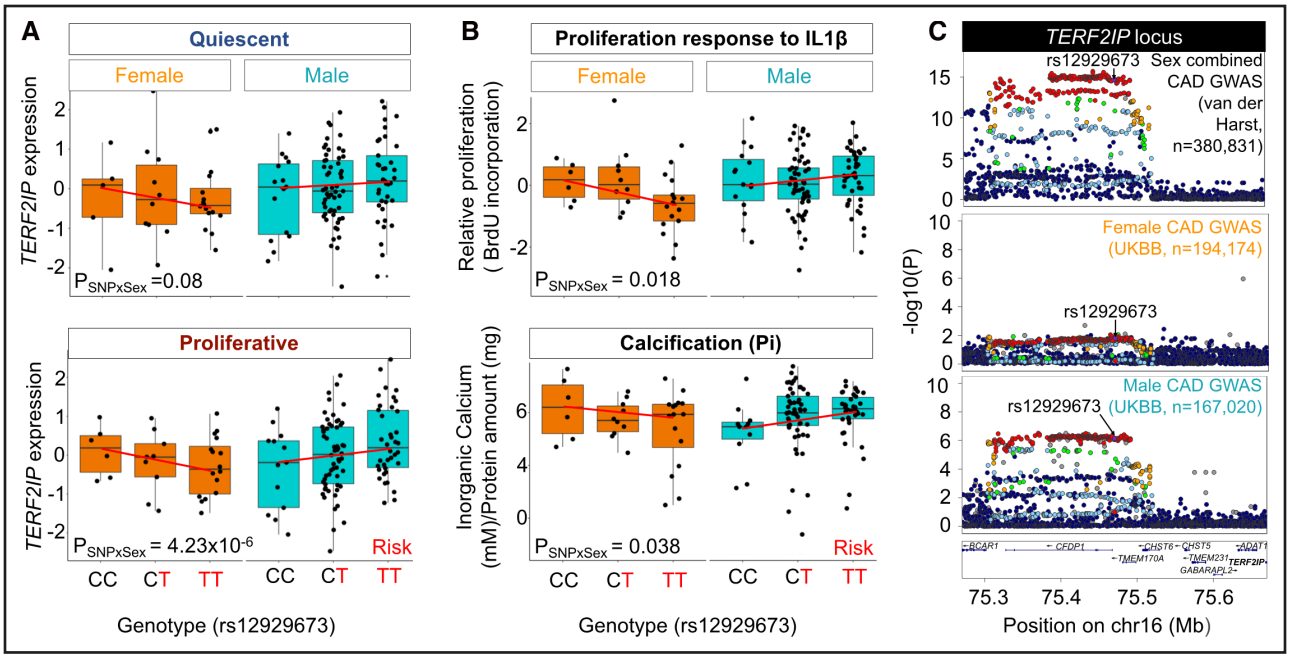

Figure 4. Colocalization between TERF2IP sex-biased expression quantitative trait locus (eQTL) and CAD GWAS.
Colocalization based on linkage disequilibrium between sex-biased eQTL SNPs and CAD GWAS identified *TERF2IP* as a colocalized gene. **A**, Genotype-gene expression plots for *TERF2IP* for the colocalized SNP (rs12929673) in quiescent and proliferative smooth muscle cells (SMCs) in males and females. **B**, Sex-biased association of rs12929673 with SMC proliferation in response to IL-1β and calcification in response to high inorganic phosphate based on data described in our previous publication.[27] **C**, LocusZoom[42] plots showing the association signal for sex-combined (Harst and Verweij[4]) and sex-stratified (UK Biobank[43]) GWAS for coronary artery disease at the *TERF2IP* locus. CAD indicates coronary artery disease; GWAS, genome-wide association studies; and SNP, single-nucleotide polymorphism.

the top biological pathways identified in the GO enrichment analysis of downregulated genes (Table S10), while pathways that are associated with actin cytoskeleton organization, cell morphogenesis, and regulation of cellular component biogenesis were identified in GO enrichment analysis of upregulated genes (Table S11).

In contrast, none of the differentially expressed genes are cis-regulated suggesting that the effects of *SNHG18* are in *trans* to affect other genes. Next, we assessed *SNHG18* mRNA expression in human aortic SMCs from 3 donors using RNA-scope in situ hybridization and immunofluorescence. Our results confirmed that *SNHG18* mRNA localized to the nucleus and cytoplasm and was co-expressed with SMC marker *ACTA2* mRNA in all the cells we imaged. In agreement with our results, Zhen et al[45] demonstrated, using in situ hybridization assays, that *SNHG18* is localized to both the nucleus and cytoplasm in different cancer cells and tissues.

Collectively, these lines of evidence point to 3 variants in the 5p15 locus as potential causal SNPs regulating the expression of *SNHG18* and SMC proliferation, thereby affecting the CAD risk in this locus.

sQTL in SMCs

Previous studies showed that RNA splicing explains a large proportion of heritable risk for complex diseases.[9] To identify genetic loci associated with mRNA splicing, we quantified RNA splicing with LeafCutter[46] and performed association mapping with tensorQTL.[29] We identified 3147 and 3578 sGenes with cis-sQTL (<200 kb from splice sites) in the quiescent and proliferative phenotypes, respectively, (FDR q-value <0.05; Table 1). About 1919 and 1700 sGenes from quiescent and proliferative SMC, respectively, were differentially expressed (Table S12). Similar to the eQTL findings, the majority of the sGenes were shared (51%) between the 2 conditions ($P_{overlap}$<1×10^{-300}, hypergeometric test). Nineteen percent (834) and 29% (1265) of the sGnes were unique to quiescent and proliferative conditions, respectively (Figure S11A). Conditioning on the lead SNPs identified 144 and 120 secondary and beyond sQTLs for quiescent and proliferative conditions, respectively. We also determined the overlap of genes with cis-eQTL or cis-sQTL. In quiescent SMCs, only 20% (1008) of the 5139 genes with a cis-eQTL or cis-sQTL were genetically regulated both at the mRNA splicing or expression levels ($P_{overlap}$=4.7×10^{-136}, hypergeometric test). Similarly, in proliferative SMCs, only 24% (1522) of the 6244 with a cis-eQTL or cis-sQTL were genetically regulated both at the mRNA splicing or expression levels ($P_{overlap}$=9.1×10^{-189}, hypergeometric test; Figure S11B). This suggests that genetic regulation of mRNA abundance and splicing are largely independent in SMCs, in agreement with studies in other tissues.[9,47]

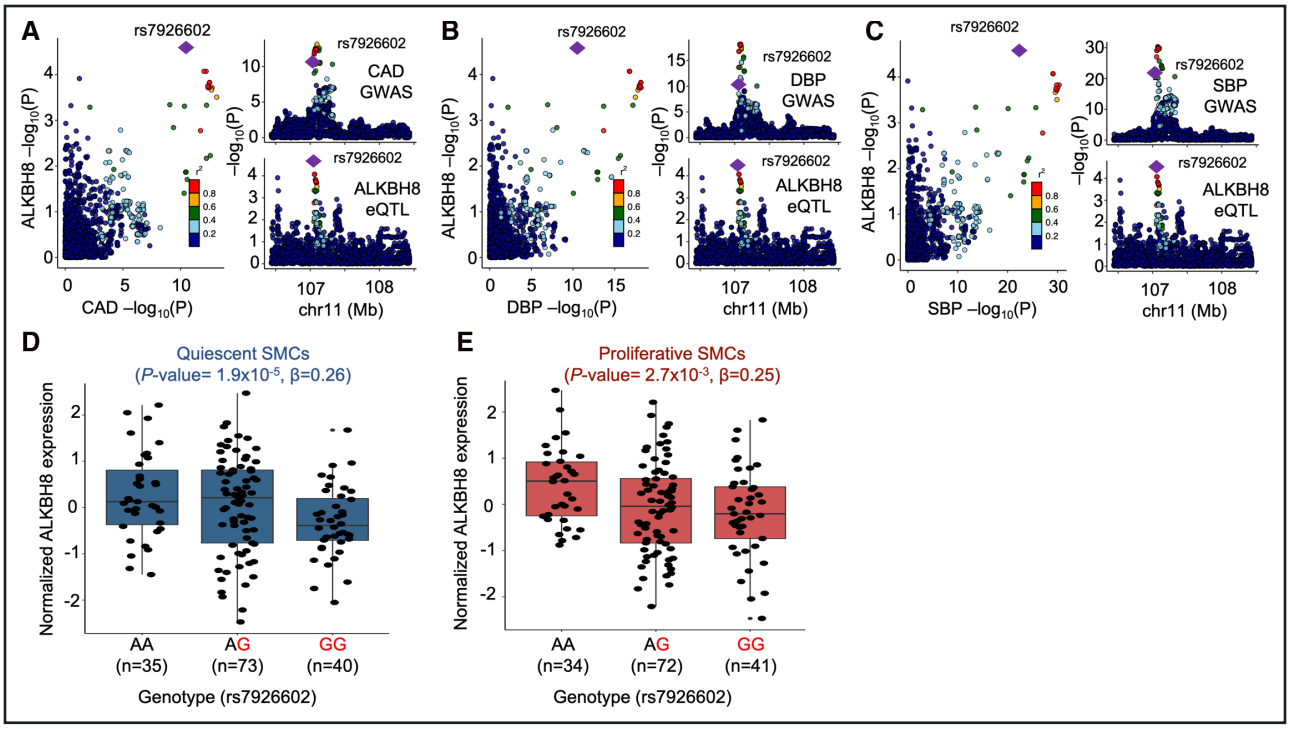

Figure 5. Colocalization between SMC-specific eQTL and vascular disease GWAS.
cis-eQTL for *ALKBH8* expression colocalized with the 11q22.3 coronary artery disease (CAD; **A**), diastolic blood pressure (DBP; **B**), and systolic blood pressure (SBP; **C**) GWAS locus. The risk allele (G) of SNP rs7926602 is associated with lower *ALKBH8* expression in quiescent (**D**) and proliferative (**E**) SMCs. CAD indicates coronary artery disease; eQTL, expression quantitative trait locus; GWAS, genome-wide association studies; SMC, smooth muscle cells; and SNP, single-nucleotide polymorphism.

Colocalization Between sQTLs and CAD GWAS Signals

To identify genes whose alternative splicing is associated with genetic risk for CAD, we performed colocalization analyses of splicing QTLs and CAD loci using 4 distinct approaches similar to the eQTL analysis described above. The intronic excision levels, as measured by LeafCutter,[46] of 100 and 120 sGenes in the quiescent and proliferative phenotypes were significantly associated with CAD loci, respectively. Colocalization of cis-sQTLs with 44 and 60 genes with CAD loci was unique to quiescent or proliferative SMCs, respectively (Figure 7A; Table S13). Significantly more CAD genes were colocalized with sQTLs (164) than eQTLs (84). We examined whether the identified SMC sQTLs that colocalized with CAD could impact the ability of RNA-binding proteins (RBPs) playing a role in splicing events. We used RBP-Var,[48] which provides extensive annotation for the functional variants involved in posttranscriptional interactions. RBP-Var includes collection of RBP-binding SNPs, which may disrupt the binding of RBPs, derived from crosslinking immunoprecipitation sequencing data sets for 60 RBPs and motif matching sites for 153 RBPs. When considering only 303 sQTL lead variants in 164 loci, 122 (40%) overlapped the binding site of an RBP (Table S14).

We identified 11 genes whose expression and alternative splicing was associated with the same CAD loci (Table S15). These results point to the significant role of the genetic regulation of mRNA splicing as a molecular mechanism for CAD genetic risk.

We observed that the 9p21 locus, which has been the most significantly associated CAD locus in many populations, contains an sQTL for *CDKN2B-AS1*, also known as ANRIL (Figure 7A through 7C). We detected the expression of 25 of the 28 *CDKN2B-AS1* transcripts in SMCs (Figure 7D; Figure S12). Our results showed that the most significantly differentially excised intron at 9p21 (chr9:22064018-22096372) was found in *CDKN2B-AS1*. The frequency of this splicing event found in SMC proliferative phenotype (P=2.7×10^{-4}; β=−0.28) was colocalized with the genotype of the rs10217586 SNP in the CAD locus (P=1.9×10^{-122}; β=−0.16). Previous studies conducted in aortic endothelial cells,[21,22] monocytes,[23] whole blood,[24] coronary artery SMCs,[25] and umbilical artery SMCs[49] did not identify an eQTL or sQTL for *CDKN2B-AS1* in this locus, suggesting that the genetic variants in the 9p21 locus act in aortic SMCs through splicing.

DISCUSSION

GWAS have successfully identified 175 loci associated with CAD risk; however, the genes and mechanisms responsible for many of these loci remain unknown. Majority of the variants are in noncoding regions making

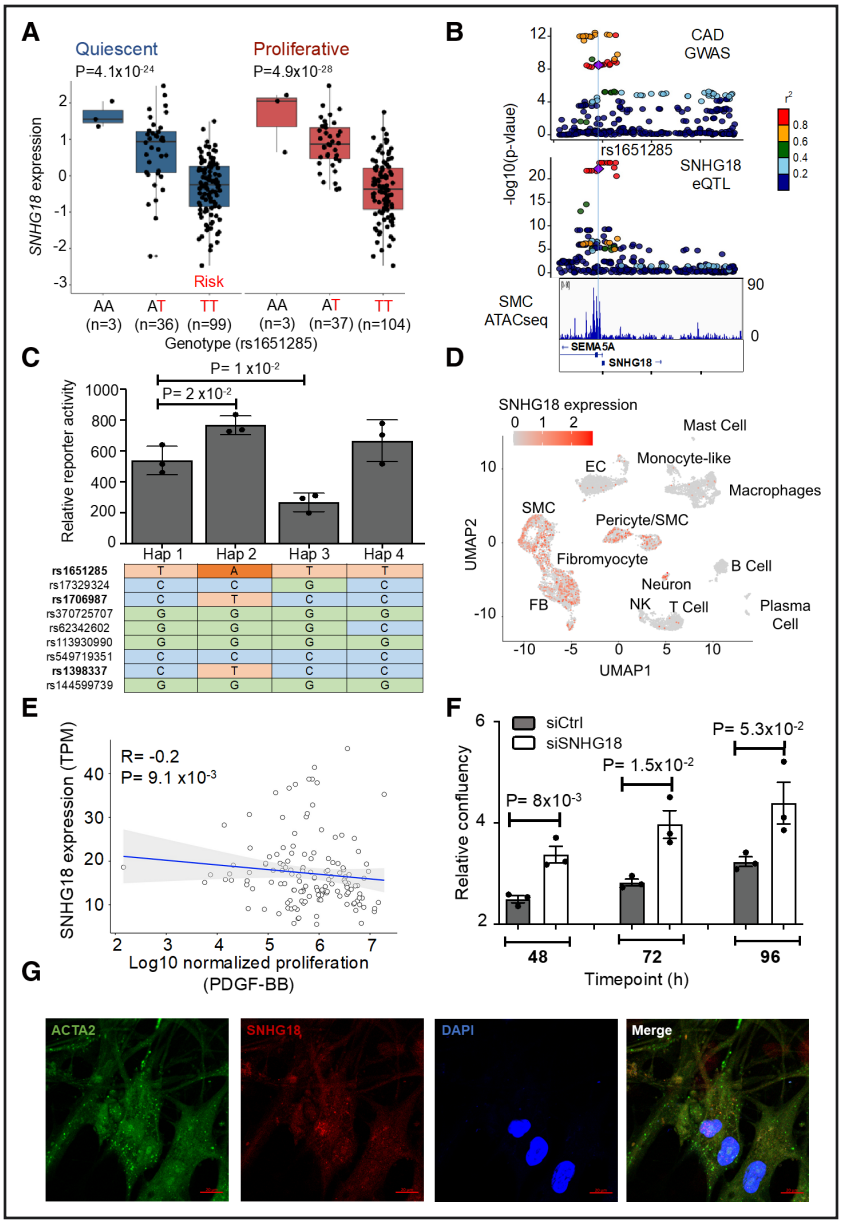

Figure 6. Colocalization between *SNHG18* eQTL and 5p15.31 CAD GWAS locus.

The risk allele (T) of SNP rs1651285 is associated with lower *SNHG18* expression in quiescent and proliferative SMCs (**A**) LocusZoom[42] plots of the CAD GWAS and SMC cis-eQTL (**B**) in the 5p15.31 locus. SNPs rs1651285, rs1706987, and rs1398337 are located in an accessible chromatin region identified by assay for transposase-accessible chromatin with high-throughput sequencing (ATACseq) in SMCs (lower panel). **C**, Bar plot summarizing the CAD haplotypes in the chr5p15/SNHG18 locus (Hap 1, 2, 3, 4) that demonstrated significant allele-specific enhancer activity in massively parallel reporter assays performed in cholesterol-loaded SMCs (n=3). Significant differences (2-sided *t* test) are shown between Hap2/Hap3 vs Hap1 (reference haplotype). A different color represents each nucleotide. Adenine (A) is indicated by the dark orange color; Cytosine (C) is indicated by the blue color; Guanine (G) is indicated by the green color; and the orange light indicates Thymine (T). **D**, Uniform manifold approximation and projection plot of single-cell RNA-sequencing data from human coronary atherosclerotic plaques.[37] **E**, Negative correlation of SNHG18 expression with platelet-derived growth factor BB (PDGF-BB)-induced proliferation in SMCs using Spearman correlation. **F**, Downregulation of SNHG18 in SMCs increased proliferation (n=3). **G**, Representative RNAScope Double in situ hybridization images (40X) of human aortic SMCs stained for ACTA2, SNHG18, and DAPI. The representative images were chosen since they captured the critical cell structures and evident staining. Scale bars = 20 μm. CAD indicates coronary artery disease; eQTL, expression quantitative trait locus; GWAS, genome-wide association studies; Hap, haplotype; and SMC, smooth muscle cells.

the task of identifying causal variants and genes difficult. Systems genetics aims to address this challenge by associating genetic variants with molecular phenotypes to comprehensively uncover the relationship between genotype and phenotype. Colocalized molecular QTL signals enable identification of reasonable candidate genes in disease-relevant cells and tissues. Therefore, we conducted, as far as we know, the largest transcriptome and whole-genome analyses using human aortic SMCs derived from a multiethnic population. We cultured these critical vascular cell types to CAD in 2 different media formulations to recapitulate the atherosclerosis-relevant quiescent and proliferative state of SMCs. PCA of the transcriptome confirmed the distinctiveness of the 2 conditions and the comparison with publicly available datasets in GTEx revealed regulatory patterns specific to human SMCs.

Intersection of our SMC eQTL data with the GTEx dataset showed that more than half of SMC eQTLs were not evident in GTEx tissues, indicating genetic regulation of gene expression unique to SMCs. Another study also found about half of the eQTLs from aortic endothelial cells of up to 157 donors were absent in the GTEx dataset.[22] This is possibly because most GTEx eQTLs have been performed in heterogeneous tissue samples containing various cell types and the genetic effects that are functioning only in rare cell types within a sampled tissue may not be detected. Indeed, most of the SMC eQTLs shared with GTEx samples were in the tissues that are rich in SMCs. Differences in RNA sequencing methods may have also contributed to the differences between GTEx and our study. Cell-type-specific eQTL analysis in disease-relevant tissues will lead to the identification of novel and more precise disease associations that can

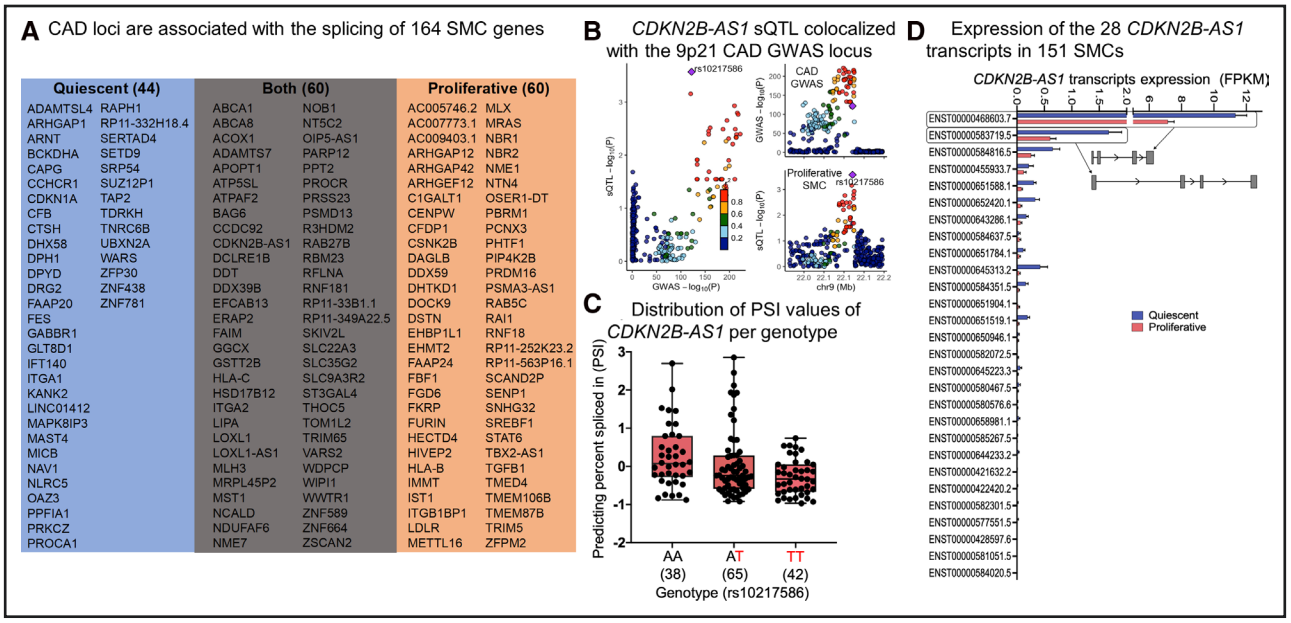

Figure 7. Summary of the SMC sQTL and CAD GWAS colocalization and the 9p21 locus.
A, Shows the colocalization summary of CAD GWAS and sQTLs signals using 4 different methods in SMCs. **B**, *CDKN2B-AS1 cis*-sQTL signal colocalized with the 9p21 CAD GWAS locus. **C**, Distribution of the PSI values for CDKN2B-AS1 intron based on the rs10217586 genotype. **D**, Average expression of the 28 *CDKN2B-AS1* transcripts in 151 quiescent and proliferative SMCs. The SEM is shown. The isoform structures of the 2 most abundant transcripts are shown. Detailed isoform structures for all transcripts are in Figure S14. CAD indicates coronary artery disease; GWAS, genome-wide association studies; PSI, percent spliced in index; SMC, smooth muscle cells; and sQTL, splicing quantitative trait locus.

help elucidate the molecular mechanisms by which the genetic variants affect the disease.

Overlapping TF binding sites with eQTL SNPs identified enrichment of putative binding sites for members of the SP2, SP1, ELK4, and GABPA TF families. While SP2 has an unknown role in SMCs, we predicted that the eQTL SNPs would impact SP2 binding to DNA in both the quiescent and proliferative phenotypes, suggesting an important role for this transcription factor in the regulation of gene expression by genetic variants in SMCs. Chromatin immunoprecipitation followed by sequencing data for the promising TFs, SP2, SP1, ELK4, and GABPA, have yet to be generated in SMCs to support our ATACseq analyses. We also predicted the transcription factors KLF5 (KLF transcription factor 5), E2F1 (E2F transcription factor 1), and CTCF (CCCTC-binding factor) to be important for atherosclerosis as their binding sites are enriched in the proliferative SMCs. E2F1 is not only known to regulate cell proliferation[50] but it is also upregulated in proliferative SMCs compared to quiescent SMCs suggesting it may be contributing to the phenotypic transition of SMCs from a quiescent to a proliferative state.

Several colocalization approaches have been developed in recent years.[51] They are sensitive to the parameters, such as thresholds applied to the prior probabilities, and differences in haplotype structures of the populations from which GWAS and molecular QTL data are derived. When the lead variants for the GWAS and eQTL studies are the same or in high LD in both populations, colocalization is straightforward.[52] Since the donors in our study population had various genetic ancestries and the CAD GWAS participants were of European ancestry, we used 4 different colocalization methods that may help to account for the differences in the LD structure. While multiple eQTLs and sQTLs had evidence for colocalization with CAD loci with 2 or more approaches, only 7 genes (*AL513548.3*, *EIF2B2*, *FES*, *FURIN*, *MAP3K7CL*, *SMAD3*, and *EST*) had evidence from all 4 approaches. Further development of colocalization methods is needed to increase the confidence in molecular QTL studies for identifying candidate genes for GWAS loci. For example, analytical approaches by modeling varying LD patterns across datasets with multiethnic populations will most likely enhance colocalization discoveries. Notably, most existing molecular QTL and GWAS studies are limited by the cost of the phenotyping, genotyping, and sequencing power and focus on people of European ancestry. In addition, the combination of the improved colocalization methods with data aggregation, for example, a meta-analysis of molecular QTLs (sQTL, eQTL, pQTL, etc) and other relevant molecular phenotype studies, such as data from single-cell RNAseq, methylation, chromatin accessibility, and histone modification, could significantly enhance both the sensitivity and specificity of the colocalization findings.

In addition, while performing the colocalization analyses using molecular QTL results from a multiethnic population could identify common variants that have disease relevance for all ethnic groups, they could also miss some genetic-ancestry-specific colocalizations due to differing haplotype structures between the populations where the

molecular QTL studies and CAD GWAS were performed. Therefore, more studies, like the most recent CAD GWAS performed in a multiethnic population, are needed.[53]

A previous study predicted the colocalization of 5 genes with CAD loci using eQTL data from coronary artery SMCs of 52 donors cultured only under proliferative conditions.[25] Our study of 151 donors whose SMCs were cultured in 2 conditions significantly expands these previous findings. Only 2 of the 5 predicted causal genes in the coronary SMC study, *FES* and *SMAD3*, were replicated in our study. One of the 5 genes, *TCF21*, which encodes a transcription factor that inhibits SMC differentiation,[17,54] is expressed in coronary but not aortic SMCs; therefore, we were not able to test its association with genetic variants. While the other 2 genes, *SIPA1* and *PDGFRA*, are expressed in aortic SMCs, we did not detect their colocalization with CAD loci. A recent study that was published[49] while this article was being revised identified 42 257 eGenes in umbilical cord aortic SMC from 1486 donors. They identified 85 genes colocalized with CAD GWAS loci, 16 of which overlapped with our study (*AL513548.3 [MIA3-AS1], ARHGAP42, CETP, EIF2B2, FES, FGD6, LINC02542, MAP1S, REST, MLH3, NEK9, SKIV2L, SMAD3, SMG6, SNHG18, TDRKH-AS1*). The discrepancies between our study and these 2 studies may be related to differences in methods, such as the *P*-value thresholds for declaring a cis-eQTL or its colocalization significant in addition to the aortic bed where the SMCs were isolated from. By performing our studies in a large number of donors with deep RNA sequencing, we provide a significant number of the predicted causal CAD genes playing a role in SMCs. The differences between ascending aorta, coronary artery, and umbilical cord aorta SMC eQTLs point to differences in transcriptional regulation among the vascular beds. Thus, our aortic SMC design provides more generalizable knowledge for aortic diseases. Larger numbers of donors or meta-analysis of the SMC eQTL studies should lead to the identification of more causal genes associated with CAD.

Nearby genes to the index SNP in a locus are usually used as signposts for annotation. Previous CAD GWAS identified 366 nearby genes as potentially causal.[4] We found that only 26 and 33 of these nearby genes matched eQTL and sQTL, respectively. On the contrary, 173 genes were distinct from the initial locus annotations. A total of 58 were derived from eQTL and 127 were derived from sQTL colocalizations. For example, rs11810571 in the 1q21.3 locus is located near *TDRKH* and *RP11-98D18.9*. However, our results showed significant colocalizations with the expression level of *GOLPH3L*, *CTSK*, and *CTSS* located ~1 MB from the variant. For 8 loci, we identified multiple genes associated with the risk variants. For example, 15q21 risk locus was associated with *FES* and *FURIN* genes in proliferative state, while in quiescent SMCs it was associated with the expression of *FES*, *RCCD1*, and *IQGAP1*. Finally, previous GWAS identified missense mutations in 20 genes.[4] For 2 of the genes, *LIPA* and *TRIM5*, we also observed an eQTL effect. For 3 of the genes, *ADAMTS7*, *DAGLB, DHX58*, we also observed both eQTL and sQTL effects. These examples demonstrate the complexity of the molecular mechanisms by which CAD loci affect disease risk.

Long noncoding RNAs (lncRNAs) are typically >200 nucleotides in length and do not contain a functional open reading frame. They can be encoded within protein coding genes or can be encoded in the intergenic regions from the sense or antisense DNA. They are expressed at much lower levels relative to their protein coding counterparts.[55] By performing library preparation with ribosomal RNA depletion, as opposed to polyA selection, and deep sequencing, we were able to assess the expression of ~3000 lncRNAs. Recent studies have shown that lncRNA plays an essential role in SMC biology and CAD.[56,57] We identified that CAD loci were associated with the expression of 12 lncRNAs and with the splicing of 15 lncRNAs. One of the colocalized lncRNA was small nucleolar RNA host gene 18 (*SNHG18*), which was regulated by the variants in the 5p15 CAD locus in both the quiescent and proliferative SMCs. While the role of SNHG18 in SMC biology and CAD have not been studied, it was observed to be upregulated in glioma and regulate the progression of epithelial-mesenchymal transition and cytoskeleton remodeling of glioma cells.[58]

Identifying tissue- and cell-specific mechanisms of GWAS loci has been challenging with few notable exceptions.[59–61] We found that SMC-specific eQTL for *ALKBH8* colocalized with the 11q22.3 CAD locus, with the risk allele leading to lower *ALKBH8* expression. The CAD risk allele is also associated with higher blood pressure,[43] suggesting a role for this gene, which encodes a tRNA methyltransferase, in regulating the vascular tone. Embryonic fibroblasts isolated from *Alkbh8*-deficient mice were shown to have increased levels of intracellular reactive oxygen species, lipid peroxidation products, and a transcript expression signature indicative of oxidative stress compared to fibroblasts isolated from wild-type littermates.[62] ALKBH8 agonists have been proposed for treating myocardial infarction injury due to its function on the modulation of autophagy and oxidative stress.[63]

Significant differences between sexes in the underlying pathology of atherosclerosis and its gene regulation have been described by us and others.[18,64] We had 118 male and 33 female donors in our population. The sex ratio of the donors is similar to the reported cases over a 20-year period in heart transplant donor registries, where 31.3% of the transplanted hearts are from deceased women.[65] Despite the imbalance in the numbers of males and females, which may have affected the statistical power of the interaction test, we were able to identify ~1000 sex-biased eQTLs. One example of sex-biased eQTLs that colocalized with CAD GWAS signal was the *TERF2IP* gene. Lower *TERF2IP* leads to telomere elongation,[66] which is associated with decreased CAD risk.[67] Risk allele

at this locus was associated with higher *TERF2IP* expression in males and lower expression in females. The same locus had an association with CAD risk only in males, suggesting that the lower *TERF2IP* expression in females may be playing a protective role against atherosclerosis. These results suggest that considering sex as a biological variable in cardiovascular research is essential to enhance our understanding of sex differences and to inform the development of sex-specific preventions and interventions in multiethnic populations.

Identifying genes whose expression is influenced by colocalizing cis-eQTL is just the first step in dissecting SMC and CAD GWAS loci. Discovering the functions of the predicted causal genes in SMC biology and CAD risk is also needed. We had previously shown that CAD loci are associated with atherosclerosis-relevant cellular phenotypes in the same donors.[27] We combined the 2 datasets to predict that *DHODH* and *FGD6* are candidate causal genes that regulate SMC proliferation. *DHODH* encodes dihydroorotate dehydrogenase, which catalyzes the fourth enzymatic step in de novo pyrimidine biosynthesis. Its role in SMCs is not known but pyrimidine nucleotides are involved in the energetics of smooth muscle contracture.[68] A missense variant in *FGD6* has been shown to increase the risk of polypoidal choroidal vasculopathy, which primarily affects the vascular layer of blood vessels in the choroid.[69]

The 175 CAD loci contain >6000 SNPs and identifying which of these alter transcriptional activity in SMCs is a necessary step to dissect the molecular mechanism of the loci. We overlaid the genomic locations of eQTL SNPs with accessible chromatin regions to predict that 194 SNPs may alter gene expression, thereby significantly reducing the number of predicted causal variants in CAD loci. Our studies show that integrating across multiple scales, from genotype to cellular phenotypes, allows us to focus on a few plausible hypotheses to test in subsequent in vitro and in vivo studies to identify the molecular and cellular genetic mechanisms of CAD loci.

The landscape of CAD-relevant RNA splicing events are mostly unknown. We observed that significantly more CAD loci were associated with splicing than expression, suggesting that the majority of the genetic risk for CAD acts through regulating transcript splicing in SMCs rather than transcript abundance. This observation is in agreement with a previous study that showed that sQTLs are more likely to be enriched for Alzheimer disease GWAS SNPs than eQTLs.[47] Further, we observed that sQTLs that colocalized with CAD loci were associated with a distinct group of genes than eQTLs, indicating that our results can explain additional factors of the genetic architecture of CAD.

Of significant note, we identified the colocalization of *CDKN2B-AS1* (*ANRIL*) sQTL with the 9p21 locus. This locus has been under intense scrutiny because it is the most significantly associated CAD locus and has been replicated in many populations with diverse ancestries.[70] Since there is no association with traditional risk factors such as dyslipidemia, diabetes, age, and sex, previous studies focused on identifying the effects at the vessel wall. eQTL studies in endothelial cells did not identify an impact of the genetic variants in this locus on gene expression[22]; however, they have been shown to regulate adhesion, contraction, and proliferation in SMCs derived from induced pluripotent stem cells.[71] Aortic SMCs isolated from mice with a knock-out of the homologous region showed excessive proliferation and diminished senescence.[72] When these mice were bred to an atheroprone background, they developed larger atherosclerotic plaques with no changes in blood pressure, lipid levels, body weight, or fasting glucose.[73] Primary SMCs were prone to dedifferentiation and had accelerated calcification, reflective of the susceptibility mechanisms of the humans carrying the risk allele. This region contains 5 tightly clustered genes, which partly overlap. *CDKN2B-AS1* overlaps in antisense the full length of the *CDKN2B* gene body while sharing a bidirectional promoter with *CDKN2A*. There are 28 linear and multiple circular isoforms of *ANRIL*. We detected the expression of 25 of the 28 linear isoforms in SMCs. Previous studies showed associations of linear *ANRIL* isoforms, as well as *CDKN2A* and *CDKN2B* with the variants in the 9p21 locus in whole blood, peripheral blood monocytes, peripheral blood T lymphocytes, lymphoblastoid cells lines, vascular tissues such as carotid atherosclerotic plaque samples, aorta, mammary artery, as well as subcutaneous or omental adipose tissue.[70,74] Our study shows these variants affect linear ANRIL splicing in SMCs; however, the associations of these variants with circular forms of ANRIL remain to be determined. The mechanism by which *ANRIL* isoforms affect SMC functions such as proliferation, migration, and calcification also needs to be explored.

Collectively, our results predicted candidate causal genes playing a role in SMCs that modulate the genetic risk for CAD. Some of the loci act differentially in quiescent and proliferative SMC phenotypes emulating different stages of atherosclerosis. They also have distinct effects in males and females, and some are SMC-specific. Taken together, our results provide evidence for the complexity of the molecular mechanisms of CAD loci. We expect that our findings will provide a rich catalog of molecular QTLs to the cardiovascular community and candidates for future preclinical development.

ARTICLE INFORMATION

Received June 24, 2022; accepted December 20, 2022.

Affiliations

Center for Public Health Genomics, University of Virginia, Charlottesville (R.A., R.N.P., Y.T.A., M.D.K., J.Y.S., D.W., J.H., G.M.S., C.L.M., M.C.). Department of Biomedical Engineering, University of Virginia, Charlottesville (D.L., R.N.P., Y.T.A., J.Y.S., M.C.). A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio, Finland (T.Ö., P.S., Q.Y., H.G., M.U.K.). Laboratory of Experimental Cardiology, University Medical Center Utrecht, Utrecht University, The Netherlands

(E.D.B., H.M.d.R.). Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, NY (L.M., J.L.M.B.). Icahn Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, NY (L.M., J.L.M.B.). Cancer Center, University of Virginia, Charlottesville (G.M.S.). Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville (G.M.S.). Integrated Cardio Metabolic Centre, Department of Medicine, Karolinska Institutet, Karolinska Universitetssjukhuset, Huddinge, Sweden (J.L.M.B.).

Sources of Funding

This work was supported by an American Heart Association Postdoctoral Fellowship 18POST33990046 (to R. Aherrahrou), Transformational Project Award 19TPA34910021 (to M. Civelek), National Institutes of Health Grants: R21HL135230 (to M. Civelek); R01HL148239 (to C. Miller); F31HL156463 (to D. Wong), Academy of Finland (Grant Nos 287478 and 319324 to M. Kaikkonen), European Research Council Horizon 2020 Research and Innovation Programme (Grant No 802825 to M. Kaikkonen), the Finnish Foundation for Cardiovascular Research (to M. Kaikkonen), and Transatlantic Network of Excellence Awards (12CVD02, 18CVD02) from Foundation Leducq (to M. Civelek, J. Björkegren, H. Ruijter, C. Miller) and its Junior Investigator Award (to R. Aherrahrou).

Disclosures

J. Björkegren is a shareholder in Clinical Gene Network AB that has an invested interest in STARNET (Stockholm-Tartu Atherosclerosis Reverse Networks Engineering Task). The other authors report no conflicts.

Supplemental Material

Expanded Materials and Methods
Figures S1–S14
Tables S1–S15
Data Set
References 28–32,38–42,44,46,75–102

REFERENCES

1. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, Boehme AK, Buxton AE, Carson AP, Commodore-Mensah Y, et al. Heart disease and stroke statistics—2022 update: a report from the American Heart Association. *Circulation*. 2022;145:e153–e639. doi: 10.1161/CIR.0000000000001052
2. McPherson R, Tybjaerg-Hansen A. Genetics of coronary artery disease. *Circ Res*. 2016;118:564–578. doi: 10.1161/circresaha.115.306566
3. Erdmann J, Kessler T, Munoz Venegas L, Schunkert H. A decade of genome-wide association studies for coronary artery disease: the challenges ahead. *Cardiovasc Res*. 2018;114:1241–1257. doi: 10.1093/cvr/cvy084
4. van der Harst P, Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease novelty and significance. *Circ Res*. 2018;122:433–443. doi: 10.1161/CIRCRESAHA.117.312086.
5. Ishigaki K, Akiyama M, Kanai M, Takahashi A, Kawakami E, Sugishita H, Sakaue S, Matoba N, Low SK, Okada Y, et al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat Genet*. 2020;52:669–679. doi: 10.1038/s41588-020-0640-3
6. Khera AV, Kathiresan S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat Rev Genet*. 2017;18:331–344. doi: 10.1038/nrg.2016.160
7. Braenne I, Civelek M, Vilne B, Di Narzo A, Johnson AD, Zhao Y, Reiz B, Codoni V, Webb TR, Asl HF, et al. Prediction of causal candidate genes in coronary artery disease loci. *Arter Thromb Vasc Biol*. 2015;35:2207–2217. doi: 10.1161/ATVBAHA.115.306108
8. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. *Nat Rev Genet*. 2015;16:197–212. doi: 10.1038/nrg3891
9. Li YI, van de Geijn B, Raj A, Knowles DA, Petti AA, Golan D, Gilad Y, Pritchard JK. RNA splicing is a primary link between genetic variation and disease. *Science*. 2016;352:600–604. doi: 10.1126/science.aad9417
10. Brown CD, Mangravite LM, Engelhardt BE. Integrative modeling of eQTLs and cis-regulatory elements suggests mechanisms underlying cell type specificity of eQTLs. *PLoS Genet*. 2013;9:e1003649. doi: 10.1371/journal.pgen.1003649
11. Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, Ellis P, Langford C, Vannberg FO, Knight JC. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet*. 2012;44:502–510. doi: 10.1038/ng.2205
12. Fu J, Wolfs MG, Deelen P, Westra HJ, Fehrmann RSN, Te Meerman GJ, Buurman WA, Rensen SSM, Groen HJM, Weersma RK, et al. Unraveling the regulatory mechanisms underlying tissue-dependent genetic variation of gene expression. *PLoS Genet*. 2012;8:e1002431. doi: 10.1371/journal.pgen.1002431
13. Hormozdiari F, Gazal S, van de Geijn B, Finucane HK, Ju CJ, Loh P-R, Schoech A, Reshef Y, Liu X, O'Connor L, et al. Leveraging molecular quantitative trait loci to understand the genetic architecture of diseases and complex traits. *Nat Genet*. 2018;50:1041–1047. doi: 10.1038/s41588-018-0148-2
14. Cherepanova OA, Gomez D, Shankman LS, Swiatlowska P, Williams J, Sarmento OF, Alencar GF, Hess DL, Bevard MH, Greene ES, et al. Activation of the pluripotency factor OCT4 in smooth muscle cells is atheroprotective. *Nat Med*. 2016;22:657–665. doi: 10.1038/nm.4109
15. Shankman LS, Gomez D, Cherepanova OA, Salmon M, Alencar GF, Haskins RM, Swiatlowska P, Newman AAC, Greene ES, Straub AC, et al. KLF4-dependent phenotypic modulation of smooth muscle cells has a key role in atherosclerotic plaque pathogenesis. *Nat Med*. 2015;21:628–637. doi: 10.1038/nm.3866
16. Pan H, Xue C, Auerbach Benjamin J, Fan J, Bashore AC, Cui J, Yang DY, Trignano SB, Liu W, Shi J, et al. Single-cell genomics reveals a novel cell state during smooth muscle cell phenotypic switching and potential therapeutic targets for atherosclerosis in mouse and human. *Circulation*. 2020;142:2060–2075. doi: 10.1161/CIRCULATIONAHA.120.048378
17. Wirka RC, Wagh D, Paik DT, Pjanic M, Nguyen T, Miller CL, Kundu R, Nagao M, Coller J, Koyano TK, et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. *Nat Med*. 2019;25:1280–1289. doi: 10.1038/s41591-019-0512-5
18. Hartman RJG, Owsiany K, Ma L, Koplev S, Hao K, Slenders L, Civelek M, Mokry M, Kovacic JC, Pasterkamp G, et al. Sex-stratified gene regulatory networks reveal female key driver genes of atherosclerosis involved in smooth muscle cell phenotype switching. *Circulation*. 2021;143:713–726. doi: 10.1161/CIRCULATIONAHA.120.051231
19. Gamazon ER, Segrè AV, van de Bunt M, Wen X, Xi HS, Hormozdiari F, Ongen H, Konkashbaev A, Derks EM, Aguet F, et al. Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. *Nat Genet*. 2018;50:956–967. doi: 10.1038/s41588-018-0154-4
20. Franzen O, Ermel R, Cohain A, Akers NK, Di Narzo A, Talukdar HA, Foroughi-Asl H, Giambartolomei C, Fullard JF, Sukhavasi K, et al. Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. *Science*. 2016;353:827–830. doi: 10.1126/science.aad6970
21. Erbilgin A, Civelek M, Romanoski CE, Pan C, Hagopian R, Berliner JA, Lusis AJ. Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. *J Lipid Res*. 2013;54:1894–1905. doi: 10.1194/jlr.m037085
22. Stolze LK, Conklin AC, Whalen MB, López Rodríguez M, Õunap K, Selvarajan I, Toropainen A, Örd T, Li J, Eshghi A, et al. Systems genetics in human endothelial cells identifies non-coding variants modifying enhancers, expression, and complex disease traits. *Am J Hum Genet*. 2020;106:748–763. doi: 10.1016/j.ajhg.2020.04.008
23. Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, Maouche S, Germain M, Lackner K, Rossmann H, et al. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS One*. 2010;5:e10693. doi: 10.1371/journal.pone.0010693
24. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*. 2013;45:1238–1243. doi: 10.1038/ng.2756
25. Liu B, Pjanic M, Wang T, Nguyen T, Gloudemans M, Rao A, Castano VG, Nurnberg S, Rader DJ, Elwyn S, et al. Genetic regulatory mechanisms of smooth muscle cells map to coronary artery disease risk loci. *Am J Hum Genet*. 2018;103:377–388. doi: 10.1016/j.ajhg.2018.08.001
26. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, et al; Roadmap Epigenomics C. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518:317–330. doi: 10.1038/nature14248
27. Aherrahrou R, Guo L, Nagraj VP, Aguhob A, Hinkle J, Chen L, Yuhl Soh J, Lue D, Alencar GF, Boltjes A, et al. Genetic regulation of atherosclerosis-relevant phenotypes in human vascular smooth muscle cells. *Circ Res*. 2020;127:1552–1565. doi: 10.1161/circresaha.120.317415

28. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*. 2020;369:1318–1330. doi: 10.1126/science.aaz1776
29. Taylor-Weiner A, Aguet F, Haradhvala NJ, Gosai S, Anand S, Kim J, Ardlie K, Van Allen EM, Getz G. Scaling computational genomics to millions of individuals with GPUs. *Genome Biol*. 2019;20:228. doi: 10.1186/s13059-019-1836-7
30. Munz M, Wohlers I, Simon E, Reinberger T, Busch H, Schaefer AS, Erdmann J. Qtlizer: comprehensive QTL annotation of GWAS results. *Sci Rep*. 2020;10:20417. doi: 10.1038/s41598-020-75770-7
31. Kumar S, Ambrosini G, Bucher P. SNP2TFBS – a database of regulatory SNPs affecting predicted transcription factor binding site affinity. *Nucleic Acids Res*. 2017;45:D139–D144. doi: 10.1093/nar/gkw1064
32. Kim-Hellmuth S, Bechheim M, Pütz B, Mohammadi P, Nédélec Y, Giangreco N, Becker J, Kaiser V, Fricker N, Beier E, et al. Genetic regulatory effects modified by immune activation contribute to autoimmune disease associations. *Nat Commun*. 2017;8:266. doi: 10.1038/s41467-017-00366-1
33. Oliva M, Muñoz-Aguirre M, Kim-Hellmuth S, Wucher V, Gewirtz ADH, Cotter DJ, Parsana P, Kasela S, Balliu B, Viñuela A, et al. The impact of sex on gene expression across human tissues. *Science*. 2020;369:eaba3066. doi: 10.1126/science.aba3066
34. Anderson WD, Soh JY, Innis SE, Dimanche A, Ma L, Langefeld CD, Comeau ME, Das SK, Schadt EE, Björkegren JL, et al. Sex differences in human adipose tissue gene expression and genetic regulation involve adipogenesis. *Genome Res*. 2020;30:1379–1392. doi: 10.1101/gr.264614.120
35. Crook MF, Olive M, Xue HH, Langenickel TH, Boehm M, Leonard WJ, Nabel EG. GA-binding protein regulates KIS gene expression, cell migration, and cell cycle progression. *FASEB J*. 2008;22:225–235. doi: 10.1096/fj.07-8573com
36. Kimura TE, Duggirala A, Hindmarch CCT, Hewer RC, Cui MZ, Newby AC, Bond M. Inhibition of Egr1 expression underlies the anti-mitogenic effects of cAMP in vascular smooth muscle cells. *J Mol Cell Cardiol*. 2014;72:9–19. doi: 10.1016/j.yjmcc.2014.02.001
37. Zhang X, Li R, Qin X, Wang L, Xiao J, Song Y, Sheng X, Guo M, Ji X. Sp1 pays an important role in vascular calcification both in vivo and in vitro. *J Am Heart Assoc*. 2018;7:e007555. doi: 10.1161/JAHA.117.007555
38. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. 2016;48:481–487. doi: 10.1038/ng.3538
39. Wallace C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet*. 2020;16:e1008720. doi: 10.1371/journal.pgen.1008720
40. Hormozdiari F, van de Bunt M, Segrè AV, Li X, Joo JWJ, Bilow M, Sul JH, Sankararaman S, Pasaniuc B, Eskin E. Colocalization of GWAS and eQTL signals detects target genes. *Am J Hum Genet*. 2016;99:1245–1260. doi: 10.1016/j.ajhg.2016.10.003
41. Stegle O, Parts L, Piipari M, Winn J, Durbin R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat Protoc*. 2012;7:500–507. doi: 10.1038/nprot.2011.457
42. Liu B, Gloudemans MJ, Rao AS, Ingelsson E, Montgomery SB. Abundant associations with gene expression complicate GWAS follow-up. *Nat Genet*. 2019;51:768–769. doi: 10.1038/s41588-019-0404-0
43. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562:203–209. doi: 10.1038/s41586-018-0579-z
44. Örd T, Õunap K, Stolze LK, Aherrahrou R, Nurminen V, Toropainen A, Selvarajan I, Lönnberg T, Aavik E, Ylä-Herttuala S, et al. Single-cell epigenomics and functional fine-mapping of atherosclerosis GWAS loci. *Circ Res*. 2021;129:240–258. doi: 10.1161/circresaha.121.318971
45. Zheng R, Yao Q, Ren C, Liu Y, Yang H, Xie G, Du S, Yang K, Yuan Y. Upregulation of long noncoding RNA small nucleolar RNA host gene 18 promotes radioresistance of glioma by repressing semaphorin 5A. *Int J Radiat Oncol*. 2016;96:877–887. doi: 10.1016/j.ijrobp.2016.07.036
46. Li YI, Knowles DA, Humphrey J, Barbeira AN, Dickinson SP, Im HK, Pritchard JK. Annotation-free quantification of RNA splicing using LeafCutter. *Nat Genet*. 2018;50:151–158. doi: 10.1038/s41588-017-0004-9
47. Raj T, Li YI, Wong G, Humphrey J, Wang M, Ramdhani S, Wang Y-C, Ng B, Gupta I, Haroutunian V, et al. Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility. *Nat Genet*. 2018;50:1584–1592. doi: 10.1038/s41588-018-0238-1
48. Mao F, Xiao L, Li X, Liang J, Teng H, Cai W, Sun ZS. RBP-Var: a database of functional variants involved in regulation mediated by RNA-binding proteins. *Nucleic Acids Res*. 2016;44:D154–D163. doi: 10.1093/nar/gkv1308
49. Solomon CU, McVey DG, Andreadi C, Gong P, Turner L, Stanczyk PJ, Khemiri S, Chamberlain JC, Yang W, Webb TR, et al. Effects of coronary artery disease–associated variants on vascular smooth muscle cells. *Circulation*. 2022;146:917–929. doi: 10.1161/circulationaha.121.058389
50. Dubrez L. Regulation of E2F1 transcription factor by ubiquitin conjugation. *Int J Mol Sci* . 2017;18:2188. doi: 10.3390/ijms18102188
51. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, Ermel R, Ruusalepp A, Quertermous T, Hao K, et al. Opportunities and challenges for transcriptome-wide association studies. *Nat Genet*. 2019;51:592–599. doi: 10.1038/s41588-019-0385-z
52. Wu Y, Broadaway KA, Raulerson CK, Scott LJ, Pan C, Ko A, He A, Tilford C, Fuchsberger C, Locke AE, et al. Colocalization of GWAS and eQTL signals at loci with multiple signals identifies additional candidate genes for body fat distribution. *Hum Mol Genet*. 2019;28:4161–4172. doi: 10.1093/hmg/ddz263
53. Tcheandjieu C, Zhu X, Hilliard AT, Clarke SL, Napolioni V, Ma S, Lee KM, Fang H, Chen F, Lu Y, et al. Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. *Nat Med*. 2022;28:1679–1692. doi: 10.1038/s41591-022-01891-3
54. Nagao M, Lyu Q, Zhao Q, Wirka RC, Bagga J, Nguyen T, Cheng P, Kim JB, Pjanic M, Miano JM, et al. Coronary disease-associated gene TCF21 inhibits smooth muscle cell differentiation by blocking the myocardin-serum response factor pathway. *Circ Res*. 2020;126:517–529. doi: 10.1161/circresaha.119.315968
55. Turner AW, Wong D, Khan MD, Dreisbach CN, Palmore M, Miller CL. Multi-omics approaches to study long non-coding RNA function in atherosclerosis. *Front Cardiovasc Med*. 2019;6:9. doi: 10.3389/fcvm.2019.00009
56. Fasolo F, Jin H, Winski G, Chernogubova E, Pauli J, Winter H, Li DY, Glukha N, Bauer S, Metschl S, et al. Long noncoding RNA MIAT controls advanced atherosclerotic lesion formation and plaque destabilization. *Circulation*. 2021;144:1567–1583. doi: 10.1161/circulationaha.120.052023
57. Jaé N, Dimmeler S. Noncoding RNAs in vascular diseases. *Circ Res*. 2020;126:1127–1145. doi: 10.1161/circresaha.119.315938
58. Zheng R, Yao Q, Li X, Xu B. Long noncoding ribonucleic acid SNHG18 promotes glioma cell motility via disruption of α-enolase nucleocytoplasmic transport. *Front Genet*. 2019;10:1140. doi: 10.3389/fgene.2019.01140
59. Small KS, Todorčević M, Civelek M, El-Sayed Moustafa JS, Wang X, Simon MM, Fernandez-Tajes J, Mahajan A, Horikoshi M, Hugill A, et al. Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. *Nat Genet*. 2018;50:572–580. doi: 10.1038/s41588-018-0088-x
60. Krause MD, Huang R-T, Wu D, Shentu TP, Harrison DL, Whalen MB, Stolze LK, Di Rienzo A, Moskowitz IP, Civelek M, et al. Genetic variant at coronary artery disease and ischemic stroke locus 1p32.2 regulates endothelial responses to hemodynamics. *Proc Natl Acad Sci U S A*. 2018;115:E11349–E11358. doi: 10.1073/pnas.1810568115
61. Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, Li X, Li H, Kuperwasser N, Ruda VM, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*. 2010;466:714–719. doi: 10.1038/nature09266
62. Endres L, Begley U, Clark R, Gu C, Dziergowska A, Małkiewicz A, Melendez JA, Dedon PC, Begley TJ. Alkbh8 regulates selenocysteine-protein expression to protect against reactive oxygen species damage. *PLoS One*. 2015;10:e0131335. doi: 10.1371/journal.pone.0131335
63. Xiao M-Z, Liu J-M, Xian C-L, Chen K-Y, Liu Z-Q, Cheng Y-Y. Therapeutic potential of ALKB homologs for cardiovascular disease. *Biomed Pharmacother*. 2020;131:110645. doi: 10.1016/j.biopha.2020.110645
64. Yahagi K, Davis HR, Arbustini E, Virmani R. Sex differences in coronary artery disease: Pathological observations. *Atherosclerosis*. 2015;239:260–267. doi: 10.1016/j.atherosclerosis.2015.01.017
65. Khush KK, Kubo JT, Desai M. Influence of donor and recipient sex mismatch on heart transplant outcomes: analysis of the International Society for Heart and Lung Transplantation Registry. *J Heart Lung Transplant*. 2012;31:459–466. doi: 10.1016/j.healun.2012.02.005
66. Cai Y, Kandula V, Kosuru R, Ye X, Irwin MG, Xia Z. Decoding telomere protein Rap1: its telomeric and nontelomeric functions and potential implications in diabetic cardiomyopathy. *Cell Cycle Georget Tex*. 2017;16:1765–1773. doi: 10.1080/15384101.2017.1371886
67. Codd V, Nelson CP, Albrecht E, Mangino M, Deelen J, Buxton JL, Hottenga JJ, Fischer K, Esko T, Surakka I, et al. Identification of seven loci affecting mean telomere length and their association with disease. *Nat Genet*. 2013;45:422–7, 427e1-2. doi: 10.1038/ng.2528

68. James SG, Appleby GJ, Miller KA, Steen JT, Colquhoun EQ, Clark MG. Purine and pyrimidine nucleotide metabolism of vascular smooth muscle cells in culture. *Gen Pharmacol Vasc Syst*. 1996;27:837–844. doi: 10.1016/0306-3623(95)02087-x
69. Huang L, Zhang H, Cheng C-Y, Wen F, Tam POS, Zhao P, Chen H, Li Z, Chen L, Tai Z, et al. A missense variant in FGD6 confers increased risk of polypoidal choroidal vasculopathy. *Nat Genet*. 2016;48:640–647. doi: 10.1038/ng.3546
70. Holdt LM, Teupser D. Long noncoding RNA ANRIL: Lnc-ing genetic variation at the chromosome 9p21 locus to molecular mechanisms of atherosclerosis. *Front Cardiovasc Med*. 2018;5:145. doi: 10.3389/fcvm.2018.00145
71. Lo Sardo V, Chubukov P, Ferguson W, Kumar A, Teng EL, Duran M, Zhang L, Cost G, Engler AJ, Urnov F, et al. Unveiling the role of the most impactful cardiovascular risk locus through haplotype editing. *Cell*. 2018;175:1796–1810.e20. doi: 10.1016/j.cell.2018.11.014
72. Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, Blow MJ, Cohen JC, Rubin EM, Pennacchio LA. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature*. 2010;464:409–412. doi: 10.1038/nature08801
73. Kojima Y, Ye J, Nanda V, Wang Y, Flores AM, Jarr K-U, Tsantilas P, Guo L, Finn AV, Virmani R, et al. Knockout of the murine ortholog to the human 9p21 coronary artery disease locus leads to smooth muscle cell proliferation, vascular calcification, and advanced atherosclerosis. *Circulation*. 2020;141:1274–1276. doi: 10.1161/circulationaha.119.043413
74. Civelek M, Wu Y, Pan C, Raulerson CK, Ko A, He A, Tilford C, Saleem NK, Stančáková A, Scott LJ, et al. Genetic regulation of adipose gene expression and cardio-metabolic traits. *Am J Hum Genet*. 2017;100:428–443. doi: 10.1016/j.ajhg.2017.01.027
75. Navab M, Imes SS, Hama SY, Hough GP, Ross LA, Bork RW, Valente AJ, Berliner JA, Drinkwater DC, Laks H. Monocyte transmigration induced by modification of low density lipoprotein in cocultures of human aortic wall cells is due to induction of monocyte chemotactic protein 1 synthesis and is abolished by high density lipoprotein. *J Clin Invest*. 1991;88:2039–2046. doi: 10.1172/jci115532
76. Zimmermann O, Zwaka TP, Marx N, Torzewski M, Bucher A, Guilliard P, Hannekum A, Hombach V, Torzewski J. Serum starvation and growth factor receptor expression in vascular smooth muscle cells. *J Vasc Res*. 2006;43:157–165. doi: 10.1159/000090945
77. Han M, Wen J-K, Zheng B, Cheng Y, Zhang C. Serum deprivation results in redifferentiation of human umbilical vascular smooth muscle cells. *Am J Physiol Cell Physiol*. 2006;291:C50–C58.
78. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA; 1000Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56–65. doi: 10.1038/nature11632
79. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48:1284–1287. doi: 10.1038/ng.3656
80. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinforma Oxf Engl*. 2010;26:2867–2873. doi: 10.1093/bioinformatics/btq559
81. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21. doi: 10.1093/bioinformatics/bts635
82. DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire M-D, Williams C, Reich M, Winckler W, Getz GR. RNA-seq metrics for quality control and process optimization. *Bioinformatics*. 2012;28:1530–1532.
83. Leung A, Stapleton K, Natarajan R. Functional long non-coding RNAs in vascular smooth muscle cells [Internet]. In: Morris KV, eds. *Long Non-coding RNAs in Human Disease*. Cham: Springer International Publishing; 2016:127–141. Accessed February 16, 2020. https://doi.org/10.1007/82_2015_441.
84. Lee S, Lee S, Ouellette S, Park W-Y, Lee EA, Park PJ. NGSCheckMate: software for validating sample identity in next-generation sequencing studies within and across data types. *Nucleic Acids Res*. 2017;45:e103–e103. doi: 10.1093/nar/gkx193
85. Jun G, Flickinger M, Hetrick KN, Romm JM, Doheny KF, Abecasis GR, Boehnke M, Kang HM. Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am J Hum Genet*. 2012;91:839–848. doi: 10.1016/j.ajhg.2012.09.004
86. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. doi: 10.1186/s13059-014-0550-8
87. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinforma Oxf Engl*. 2012;28:882–883. doi: 10.1093/bioinformatics/bts034
88. Tarazona S, Furió-Tarí P, Turrà D, Pietro AD, Nueda MJ, Ferrer A, Conesa A. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Res*. 2015;43:e140–e140. doi: 10.1093/nar/gkv711
89. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11:R25. doi: 10.1186/gb-2010-11-3-r25
90. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A*. 2003;100:9440–9445. doi: 10.1073/pnas.1530509100
91. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010;26:2336–2337. doi: 10.1093/bioinformatics/btq419
92. Davis JR, Fresard L, Knowles DA, Pala M, Bustamante CD, Battle A, Montgomery SB. An efficient multiple-testing adjustment for eQTL studies that accounts for linkage disequilibrium between variants. *Am J Hum Genet*. 2016;98:216–224. doi: 10.1016/j.ajhg.2015.11.021
93. Han B, Eskin E. Interpreting meta-analyses of genome-wide association studies. *PLoS Genet*. 2012;8:e1002555. doi: 10.1371/journal.pgen.1002555
94. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods*. 2017;14:959–962. doi: 10.1038/nmeth.4396
95. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–359. doi: 10.1038/nmeth.1923
96. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078–2079. doi: 10.1093/bioinformatics/btp352
97. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9:R137. doi: 10.1186/gb-2008-9-9-r137
98. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841–842. doi: 10.1093/bioinformatics/btq033
99. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, Garcia MU, Di Tommaso P, Nahnsen S. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol*. 2020;38:276–278. doi: 10.1038/s41587-020-0439-x
100. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*. 2020;36:2628–2629. doi: 10.1093/bioinformatics/btz931
101. Barry DM, Liu X-T, Liu B, Liu XY, Gao F, Zeng X, Liu J, Yang Q, Wilhelm S, Yin J, et al. Exploration of sensory and spinal neurons expressing gastrin-releasing peptide in itch and pain related behaviors. *Nat Commun*. 2020;11:1397. doi: 10.1038/s41467-020-15230-y
102. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9:671–675. doi: 10.1038/nmeth.2089