Therapeutic risk factor modification has provided a significant decrease in coronary artery disease (CAD) in Western populations, however, significant risk is due to common inherited genetic variation that affects disease pathways in the vessel wall and remains poorly understood without specific therapies. To further our long-term goal of characterizing the molecular basis for this genetic risk, we have participated in genome-wide association studies (GWAS) identifying allelic variation linked to coronary artery disease (CAD) risk, and these efforts have yielded hundreds of associated loci. However, the majority of identified causal variation resides outside of protein coding exons, in regulatory regions of the genome that are poorly understood, and further efforts are required to understand the mechanisms of association and thus disease risk. Our central hypothesis is that an important subset of disease allelic variation primarily regulates long non-coding RNA (lncRNA) expression, with this effect modulating causal protein coding gene (pcGene) expression through functional genomic interactions such as chromosomal looping. Our objective here is to investigate the role these lncRNAs play in mediating expression of CAD causal pcGenes, and the mechanism by which they accomplish this function. Our rationale is that lncRNAs serve as a critical intermediary between genetic and epigenetic signaling, and that elucidating their mechanism of function is a key aspect of understanding CAD risk. To gain fundamental information regarding the mode of action of these molecules in the context of CAD, we propose to study human coronary artery smooth muscle cell (HCASMC) lncRNAs. In **Aim 1**, we will identify lncRNAs regulated in these cells by disease-related stimuli and that map to CAD GWAS loci. Co-expression network analyses will connect these lncRNAs to pcGenes, and initiate network and pathway analyses to begin to establish their biological functional associations. In **Aim 2**, we will map expression quantitative trait loci variants (eQTLs) for each of the lncRNAs, using a high-throughput allele-specific expression method that provides quantification of low abundance RNAs. Discovered lncRNA eQTLs will be investigated to determine whether they colocalize with CAD GWAS causal variation, as well as genomic molecular trait QTLs. CRISPR genome editing will be employed to validate the eQTLs, and confirm pcGene identity. In **Aim 3**, we will employ CRISPR inhibition and single cell RNA sequencing (PerturbSeq) to map the transcriptional networks regulated by the disease related lncRNAs, and also investigate their in vitro cellular effects on HCASMC. These studies will be aided by our extensive work with primary cultured HCASMC characterizing epigenome modification, chromatin accessibility, and looping, and our efforts to map CAD GWAS causal variants and genes that mediate risk in this cell type. This work is highly innovative in that it combines unique genomic datasets developed in a highly disease relevant cell type and significant since it will integrate lncRNAs, their regulatory variation, and molecular mechanisms into the etiology of CAD risk.

**SPECIFIC AIMS**

Therapeutic risk factor modification has provided a significant decrease in coronary artery disease (CAD) in Western populations; but a significant burden of risk is due to common inherited genetic variation that affects disease pathways in the vessel wall and remains poorly understood without specific therapies. To characterize the molecular basis for this genetic risk, we have participated in genome-wide association studies (GWAS) aimed at identifying allelic variation linked to CAD risk, and these efforts have yielded numerous disease loci. However, the majority of the identified causal variation resides outside of protein coding exons, in regulatory regions of the genome that are poorly understood, and further efforts are required to understand the mechanisms of association and thus disease risk. Increasingly, it is becoming apparent that long noncoding RNAs (lncRNAs), regulatory RNAs which do not encode proteins, are a likely primary target of disease-associated variation. To gain fundamental information regarding the mode of action of these lncRNAs, and to map the mechanisms by which causal variation regulates risk of complex human diseases such as CAD, we aim to identify lncRNAs regulated in human coronary artery smooth muscle cells (HCASMC) by disease-related stimuli, map the expression quantitative trait loci (eQTLs) for these genes, and determine whether they are also associated with epigenomic molecular traits such as chromosomal looping and are enriched in CAD GWAS causal variation. These studies will be aided by our work with HCASMC characterizing epigenome modification by post-translational histone marks, chromatin accessibility, and looping, and our efforts to map CAD GWAS causal variants and genes that mediate risk in this cell type. Finally, we will employ CRISPR inhibition and single cell RNA sequencing to map the transcriptional programs regulated by the disease-related lncRNAs, and investigate their cellular effects on HCASMC.

*Hypothesis 1: lncRNAs play a critical role in the regulation of HCASMC response to vascular stress, linking epigenetic signaling pathways with protein coding genes that mediate disease-related processes.*

**Specific Aim 1**. To identify and characterize HCASMC lncRNAs that respond to disease-related signaling pathways and reside in CAD loci.

The full repertoire of lncRNAs that are expressed in HCASMC will be characterized with deep RNA sequencing, with particular focus on those that are differentially expressed in response to disease relevant stimuli that modulate cellular differentiation state and pro-inflammatory profile. Such lncRNAs that map to CAD loci will be investigated with gene co-expression network approaches to identify related protein coding genes (pcGenes) and provide an initial characterization of possible cellular functions.

*Hypothesis 2: HCASMC lncRNA expression is regulated by allelic variation that is linked to genomic molecular functions and CAD risk.*

**Specific Aim 2.** To associate HCASMC lncRNA expression and function with allelic variation that mediates genomic molecular traits and risk for CAD.

The goal of this work is to map expression quantitative trait loci (eQTLs), variants that regulate lncRNA expression, in CAD-associated loci and identify which of these variants are also causally associated with disease. As lncRNAs can be non-polyadenylated and lowly-expressed, we will combine RNA-seq-based analysis with a targeted high-throughput allele-specific expression assay to detect lncRNA eQTLs. lncRNA eQTLs will also be intersected with QTLs for molecular traits, such as chromosomal looping, that we have previously mapped in HCASMC. These studies will thus characterize functional mechanisms by defining how lncRNAs are modulated to regulate protein coding gene expression, function and disease risk.

*Hypothesis 3: lncRNAs integrate with pcGenes into networks that regulate molecular genomic traits and coronary artery smooth muscle cell phenotypic features that are linked to CAD risk.*

**Specific Aim 3.** To map the downstream genes that are regulated by HCASMC lncRNAs and determine how their gene expression program alters smooth muscle cell functions.

These experiments will investigate the impact of disease regulated HCASMC-specific lncRNAs on gene expression pathways and basic cell fate phenotypes in this cell type. In cultured HCASMCs using highly-multiplexed CRISPR inhibition coupled with single-cell RNA sequencing (PerturbSeq) we aim to characterize the transcriptional impact resulting from perturbation of the CAD and molecular genomic trait related lncRNAs identified in Aims 1 & 2. A focused evaluation of lncRNA function on cellular phenotype will be investigated with in vitro assays.

These studies will integrate lncRNAs, their regulatory variation, and target pcGenes into the component of CAD risk that resides in the vascular wall, and more broadly will identify mechanisms by which this class of regulatory molecules interact with the genome to provide for exquisite control of pcGene expression. Further, given the specificity of lncRNA expression and function, these findings will facilitate the development of novel RNA-targeted disease therapies.

**SIGNIFICANCE**

***Human genetic studies identify causal molecular mechanisms of coronary artery disease -*** Atherosclerotic coronary artery disease (CAD) is the world-wide leading cause of death and is growing in developing populations and minority groups in this country.1-4 Significant CAD risk is attributable to common inherited genetic variation that affects disease pathways in the vessel wall and remains poorly understood without specific therapies. Modern human genetics can identify and establish causality for disease pathways and potential molecular therapeutic targets. Genome Wide Association Studies (GWAS) have focused on common single nucleotide polymorphisms (SNPs) to link CAD with genetic loci at a rigorous level of significance. Toward that end, we have participated in large international meta-analytical efforts aimed at mapping CAD causal loci, through the CARDIoGRAM and CARDIoGRAM+C4D consortia.5-9 These efforts have identified ~130 loci strongly associated with CAD, and another ~400 associated at a false discovery rate (FDR) of 5%. While some of the genes identified through these CAD GWAS studies appear to be related to classical risk factors, the majority are not, and must underlie cellular and molecular processes in the vessel wall. With the successful mapping of disease associated loci, efforts in the lab have turned to genomic studies aimed at identifying the causal variants and genes in these loci10, and mechanistic studies investigating the physiology and pathophysiology related to the disease genes.11-14 In particular, our recent efforts have pointed to long non-coding RNAs (lncRNAs) as potential mediators of the effects of certain GWAS loci on CAD risk.

***lncRNAs are a diverse class of regulatory RNAs that modulate chromatin epigenetic features and enhancer-promoter looping* -** lncRNAs are operationally defined as being >200 nucleotides in length and devoid of an open reading frame. Their primary function is to regulate expression of protein coding genes (pcGenes) that are involved in a wide range of cellular processes including embryonic development15, cell cycle control16, survival17, migration18, alternative splicing19, and chromatin remodeling.20, 21 lncRNA effects on pcGene expression may be mediated locally through a *cis*-acting mechanism, or may impact pcGene expression from a distance in a *trans*-acting manner, at a genome-wide level.22 While lncRNAs are more similar to pcGenes than other regulatory RNAs, they are poorly conserved in sequence across different species, far exceed pcGenes in number 23, and are expressed in a highly tissue-specific manner compared to other ncRNAs and pcGenes24, at levels commonly lower than mRNA encoding genes.24, 25

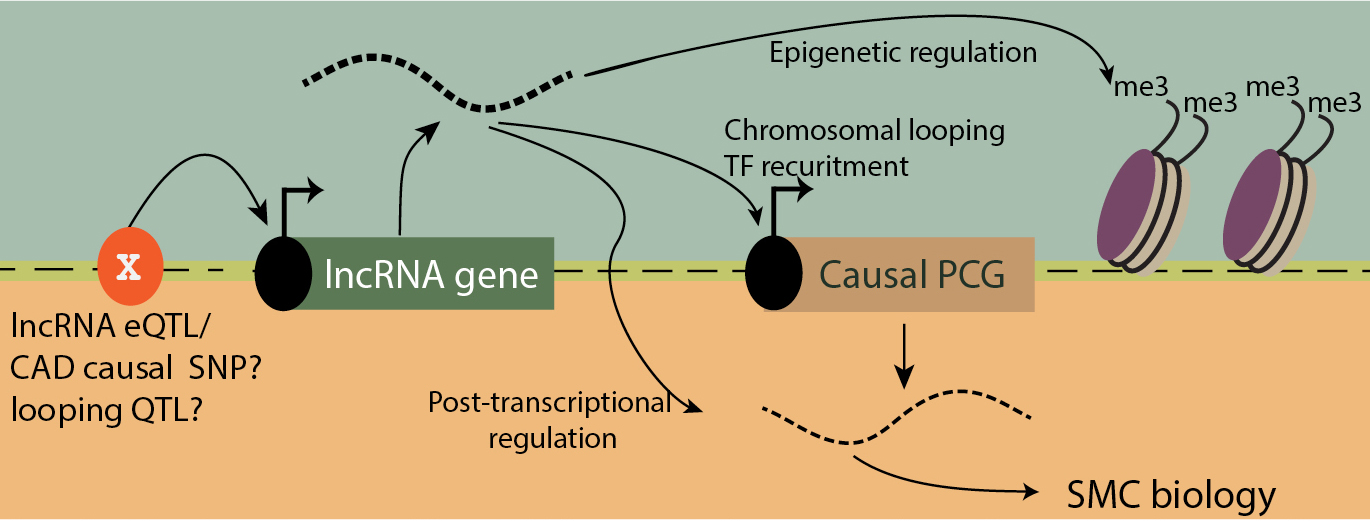
A common theme among well studied lncRNAs is that they regulate expression of nearby pcGenes by modulating epigenomic modifications and chromosomal structure. For instance, the lncRNA ANRIL recruits polycomb (PC)1 and PC2 complexes in *cis* to mediate epigenetic H3K27me3 modification of the 9p21.3 locus that has been associated with CAD26, 27, as well as numerous types of cancer and other complex human diseases.28, 29 lncRNA TARID recruits GADD45A and TET to regulate promoter methylation and expression of the CAD associated gene *TCF21* at 6q23.2.13, 14, 30 Perhaps most significant among effects at the chromatin level is the ability of lncRNAs to regulate chromosomal looping. Such looping is commonly accepted to be the primary mechanism by which enhancers physically interface with promoters to enable transcriptional activation.31-34 In this regard, an activating group of enhancer lncRNAs have been shown to interact with the mediator complex to promote looping complex formation and promote transcription.35 Further, Tan et al. have shown that genetically regulated lncRNAs are clustered at transcriptional activation domains (TADs) in conjunction with CTCF binding sites and found that the density of chromosomal contacts is significantly higher for TADs containing such lncRNAs.36 There are numerous examples of specific lncRNAs that mediate or participate in looping, including CCAT1-L that mediates long-range interaction between its locus and the distal *MYC* gene37, and the lncRNA HOTTIP which functions to organize chromatin domains to coordinate long-range gene activation in the *HOXA* locus.38 The lncRNA Firre has been shown to mediate looping between multiple loci covering ~5 Mb of the X-chromosome.39 In view of the general ability of lncRNAs to tether at their site of transcription, and to interact with various transcriptional regulatory elements, we hypothesize that such activity will be more commonly recognized as lncRNA function is investigated with newer high throughput methods such as HiChIP that allow rapid and comprehensive evaluation of chromosomal interactions.40, 41

***lncRNA transcription is regulated by epigenetic and genetic mechanisms* -** lncRNAs show spatial- and temporal-specific patterns of expression, indicating that lncRNA transcription is highly regulated.42 They have similar but more compact promoter regions than pcGenes, marked by H3K4me3 histone modifications36, 43, and characterized by open chromatin configuration in cells and tissues where the lncRNAs are expressed. Additionally, lncRNAs are regulated by enhancer regions, similar to mRNA genes, and these appear to be in close proximity to the structural gene and to be marked by open chromatin and enhancer marks such as H3K4me1 and H3K27Ac.36 Also, similar to mRNA genes, variation in lncRNA gene expression results from common SNPs that modulate the binding of transcription factors and other chromatin regulating factors that determine enhancer-promoter interactions. Such regulatory variation is commonly referred to as expression quantitative trait loci (eQTL), which may function in *cis* or *trans*.44, 45 eQTL analysis, associating differences in gene expression with genomic variation, revealed that lncRNAs have more *cis*-eQTLs than do equally expressed pcGenes with the same exon number46, but that lncRNA eQTLs are often missed because of low RNA sequencing depth.36 lncRNA *cis*-eQTLs are located closer to transcription start sites and their effect sizes are greater than *cis*-eQTLs found for pcGenes.46 The GTEx project found that ~50% of all annotated lncRNAs have a *cis-*eQTL, indicating a widespread effect of genetic variation on lncRNA expression.47

***lncRNA regulatory SNPs alter CAD risk by modulating expression of disease causal pcGenes*** –GWAS over the past decade with ever increasing numbers of study participants have identified hundreds of CAD locus associations.48 Despite this very significant advance in knowledge regarding the genetic risk of CAD, there have been very few studies detailing the mechanisms by which causal variants mediate disease epigenomic signaling and causal genes mediate vessel wall disease pathways. This is due to the fact that only 7% of disease-associated SNPs are exonic, while the remaining 93% are located in gene regulatory or intergenic regions.49 Indeed, extensive pcGene exon sequencing has not been able to identify novel significant CAD associating variants.50 Poor annotation of the regulatory genome, and lack of understanding regarding transcriptional regulatory networks are the fundamental reasons for lack of progress in the field.

Interestingly, many CAD SNPs that map to non-coding intervals have been associated with lncRNAs **(Fig. 1)**, and through study of eQTL, histone modification, and chromatin accessibility data, along with in vitro functional studies, a number of causal lncRNA genes have been identified. This is true for the 9p21.3 CAD associated locus, with the identification of the lncRNA *ANRIL* gene as casual.28, 51-55 Our lab identified *CDKN2B* as a key pcGene target of ANRIL, and established a role for this gene in the smooth muscle cell (SMC) response to vascular stress.56, 57 In another highly replicated CAD locus at 6q23.2 disease variants are eQTLs for the antisense lncRNA TARID, which regulates expression of *TCF21*. Again, we have established a functional role for *TCF21* in the regulation of SMC phenotypic response to disease stimuli.11, 13, 14 In our pilot studies, an average ~1.1 lncRNAs were found in each CAD locus (see **Aim 1**), and recent CAD meta-analyses have found numerous lncRNA eQTLs primarily associated with disease, including for example *TEX41* in the *ZEB2* locus.58 The involvement of lncRNAs in CAD genetic risk is substantiated by the role of lncRNAs in a number of other complex diseases including numerous forms of cancer where lncRNAs have been identified as the target of variant-related expression differences.37, 59-62

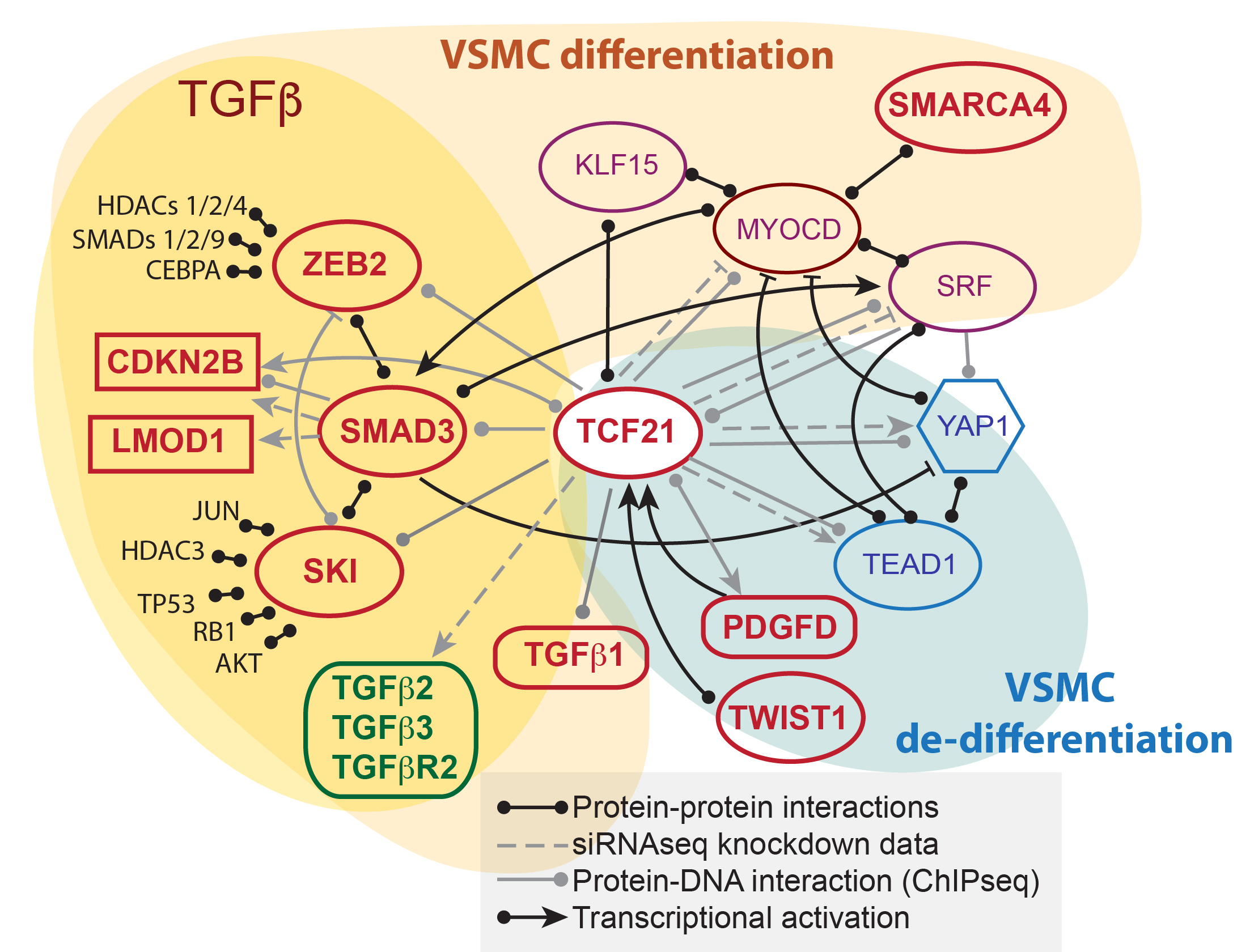
It is thus our hypothesis that lncRNAs are an important intermediate between causal CAD associated allelic variation and disease related modulation of causal pcGene gene expression. This hypothesis is based on the observations that: *i)* there are more eQTLs regulating expression of individual lncRNAs compared to individual pcGenes46, so more opportunity for disruption of regulatory circuits, *ii)* there are at least 5-times more lncRNAs compared to pcGenes, further contributing to the total number of lncRNA regulatory variants that may be disrupted, and resulting in greater genomic exposure to spontaneous variation (LNCipedia), *iii)* the primary focus of lncRNAs is to mediate cell type and environmental influences on pcGene expression, processes directly relevant to disease etiology22, 42, 63, *iv)* lncRNAs show greater cell-specificity than pcGenes, as do complex human diseases10, 24, 64-67, and *v)* they appear to be an integral component of chromatin architecture and more highly regulated by chromosomal elements.68, 69



**Fig. 1. Mechanisms of lncRNA regulation and function.** Allelic variation (X) regulates lncRNA expression, which regulates pcGene expression, and through various functions of lncRNAs affects CAD risk and chromosomal features such as looping. Abbreviations: eQTL, expression quantitative trait locus (variant); PCG, protein coding gene; me3, histone 3 lysine 27 trimethylation (repressive histone modification); TF, transcription factor.

***Smooth muscle cell phenotype as a mechanism of genetic CAD risk* -** The risk of CAD events is inversely correlated to the number of SMC in atherosclerotic plaque, and it is speculated that this cell type stabilizes the lesion to protect against plaque rupture and myocardial infarction.70-74 SMC contribute over 50% of cells in atherosclerotic plaque75 and express a majority of CAD-associated genes.11 Also, SMC etiology has been identified in Mendelian forms of early onset CAD through the mapping of causal gene mutations to SMC-specific genes (e.g., *ACTA2, LRP6*) and these mutations shown to affect fundamental SMC phenotypic properties.76-79 Further, there are numerous genes in CAD loci that have been convincingly identified as causal that are known to be involved in SMC disease processes5, 9, 58, 80-82, e.g., *LMOD1*83, *EDNRA*84, *PDGFD*85, *KLF4*75, 82, and *REST*86. Importantly, SMC can undergo a process of epigenetic reprogramming, termed “phenotypic modulation,” that is characterized by downregulation of lineage marker expression, proliferation, and migration to the luminal surface of the atheroma where these cells contribute to the plaque stabilizing fibrous cap.72, 75, 87 This process is also increasingly linked to the transdifferentiation of vascular SMC to a macrophage like phenotype75, 88-90, suggesting complex and competing SMC state changes that might contribute both to cellular events that promote an inflammatory milieu but also inhibit plaque rupture. Our studies have shown that CAD causal gene *TCF21* promotes phenotypic modulation11, and that CAD GWAS gene *SMAD3* opposes this process by promoting a differentiated state in HCASMC **(Fig. 2)**.91

To further support the hypothesis that HCASMC contribute to CAD risk, we conducted a number of genomic studies. Linkage disequilibrium score regression92 indicated that HCASMC, along with coronary artery tissues (GTEx), make a significant contribution to CAD risk. Also, we estimated the overlap between CAD variants and open chromatin in HCASMC and ENCODE cell types, and found that HCASMC had highly significant overlap, along with endothelial cells.10, 93 Further, investigation of the downstream transcriptional network regulated by TCF21 revealed an enrichment for binding in other CAD loci, implicating additional HCASMC genes in the disease process.12 Thus a significant portion of CAD risk appears to reside in HCASMC where it is associated with the phenotypic changes in this cell type related to vascular stresses that promote disease risk **(Fig. 2)**.

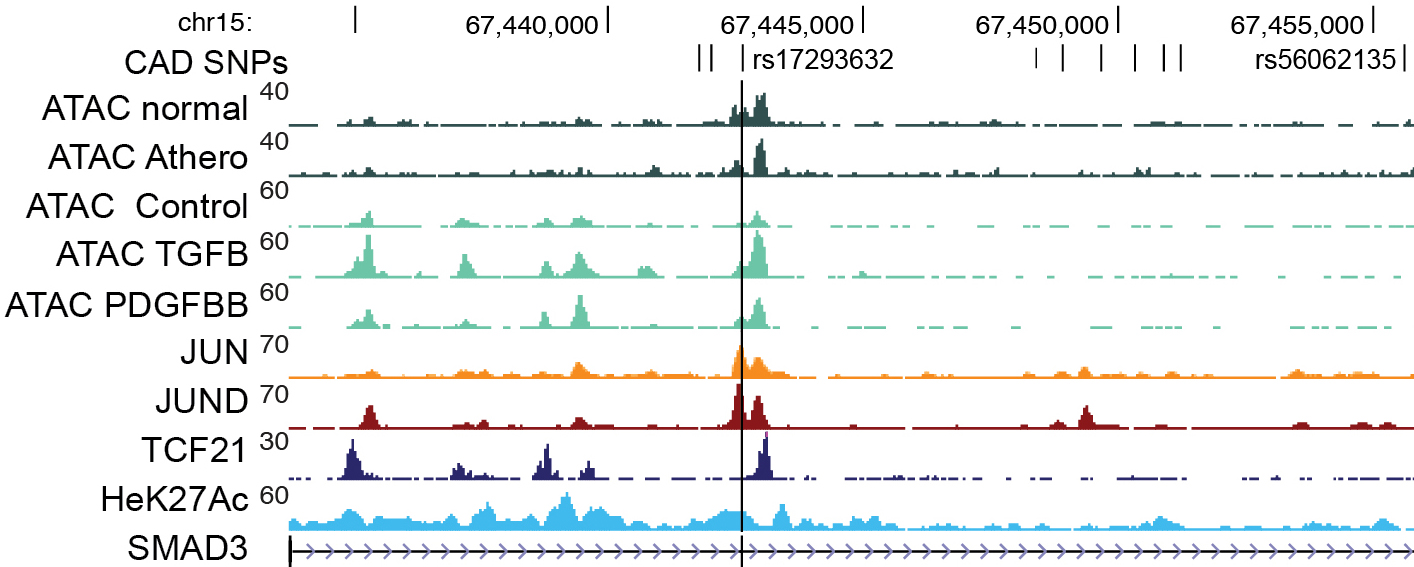


**Fig. 2. HCASMC transcriptional CAD network based on studies in this laboratory.** GWAS genes identified in CAD associated loci are shown in bold red, transcription factors are shown as ellipses. Yellow shading indicates CAD associated genes in the TGFβ pathway, beige shading more broadly identifies pro-differentiation factors. Blue shading indicates growth factor and transcription pathways that promote de-differentiation.

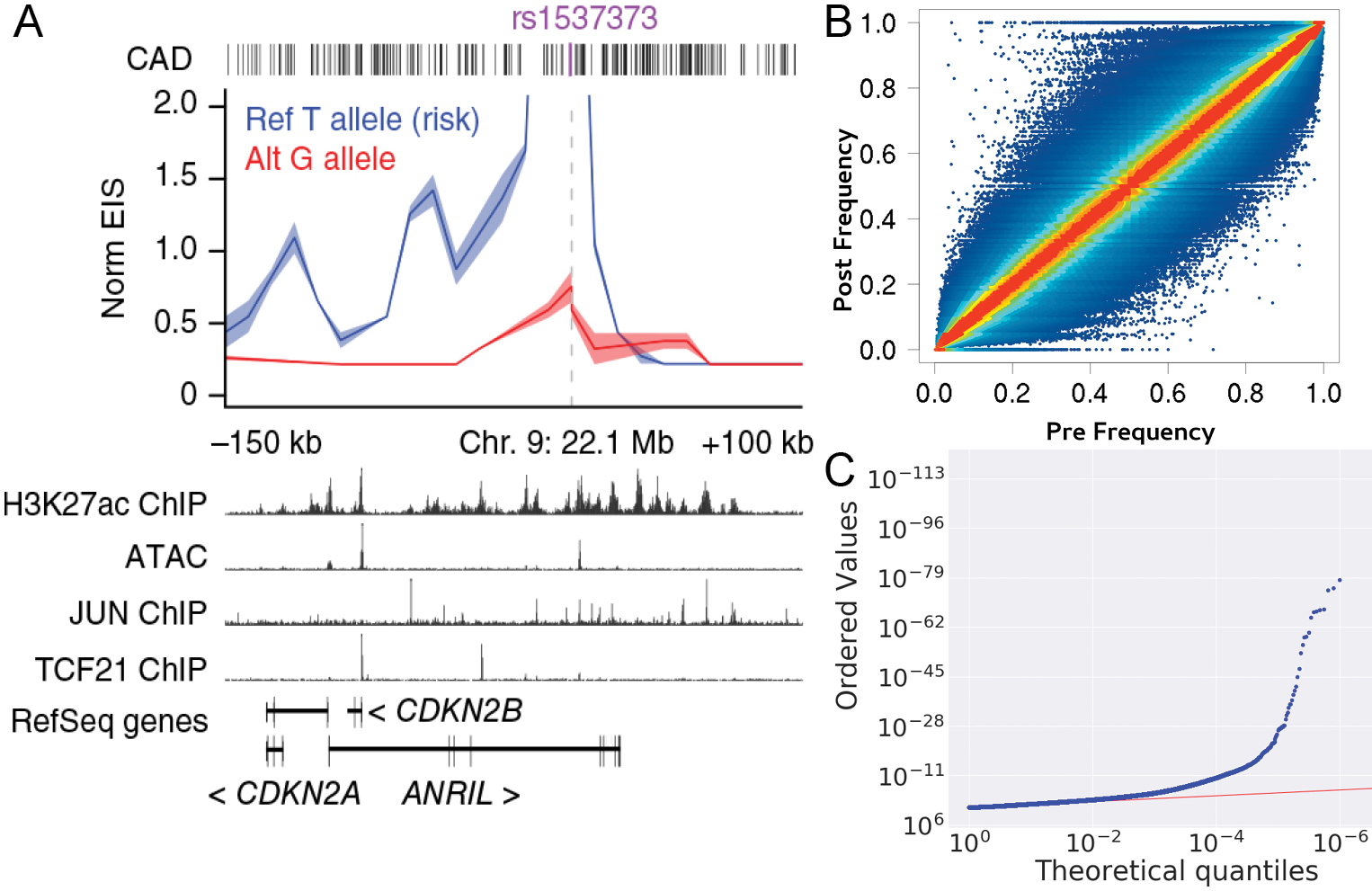
***Extensive genetic and genomic HCASMC datasets support studies of lncRNAs in CAD*** *–* A primary goal of the laboratory is to investigate the fundamental role that is played by lncRNAs in mediating connections between the epigenome, causal variation and causal pcGene regulation in HCASMC. However, such studies are impossible without comprehensive mapping of eQTLs, chromatin accessibility, epigenetic state, and the identity of causal variation in associated CAD loci. Thus, we have collected and performed whole genome sequencing in 60 HCASMC lines, called variants with GATK, and imputed these data to 1000 Genomes. The algorithm RASQUAL94, that integrates eQTL and allele-specific expression, was employed to identify eQTLs that regulate pcGene expression in this cell type.93. eCAVIAR95 and SMR96 identified colocalization of GWAS SNPs and eQTLs informing high confidence colocalization of 20-30 causal variants and genes in CAD loci. Further, we have conducted ATAC-seq97 studies of open chromatin in basal and cytokine/growth factor stimulated cells as well as in diseased and normal coronary tissue **(Fig. 3)**. ChIPseq studies in HCASMC have mapped H3K4me1, H3K4me3, and H3K27Ac histone marks, identifying active promoter and enhancer regions.10 Also, we were the first lab to use the newly developed method of HiChIP to identify chromosomal loops that juxtapose causal gene promoters with enhancers containing CAD causal variation **(Fig. 4A)**.40 We have also mapped CAD-associated transcription factor binding sites to characterize the downstream networks regulated by these factors, including *TCF21*, *SMAD3*, and the aryl hydrocarbon receptor (*AHR*), all of which regulate coronary SMC phenotype**.**10, 12, 91, 98 Ongoing efforts are mapping additional regulatory variation that controls genomic molecular traits such as transcription factor binding, chromatin accessibility, etc., employing the pooling approach developed in the lab of Hunter Fraser (Consultant, see letter of support).99 Such efforts to map looping QTLs (clQTLs) have identified ~50,000 high confidence clQTLs in HCASMC **(Fig. 4B, C)**. Our expectation is that these molecular genomic trait QTLs will be significantly enriched for lncRNA eQTLs and CAD causal variants, suggesting that chromosomal regulatory variation regulates disease risk by modulating lncRNA and pcGene expression, through local control of chromatin accessibility and architecture.

**INNOVATION**

Biological innovation- There have been no previous comprehensive studies mapping lncRNA regulatory variation and investigating the relationship between the target lncRNAs, their related pcGenes, and the mechanisms by which these two types of genes interact to mediate disease risk. Our proposed studies are unique in that we have available extensive genetic and genomic information that will allow much better identification of lncRNAs, their eQTLs, and their integration into risk mechanisms. Our focus on the smooth muscle cell as a key component of CAD is also quite unique, and these studies in conjunction with our previous work will bring focus to their contribution to disease pathophysiology. Technical innovation- We will use advanced genomics techniques, including the application of microfluidics multiplex PCR sequencing to derive allele-specific expression information from lowly-abundant or non-polyadenylated lncRNA genes, colocalization algorithms to identify the most interesting lncRNA eQTLs, innovative network analysis of lncRNA-mRNA associations, and cutting edge CRISPRi single cell RNA sequencing (PerturbSeq) to identify downstream networks and pathways that are regulated by lncRNAs in this disease cell type in response to CAD GWAS pathway activation. Overall, our primary innovation is the paradigm developed here to establish large-scale lncRNA expression, epigenetic, and genomic data in a highly relevant cell type, map regulatory variation for the lncRNAs and integrate their function into a causal disease framework. This paradigm should be widely useful for the study of lncRNAs in other complex diseases.



**Fig. 3. HCASMC epigenome and ATACseq open chromatin mapping data from this lab identifies CAD causal variant rs17293632 in the SMAD3 locus.** Shown are ATAC-seq open chromatin tracks in coronary medial tissue (ATAC normal, ATAC Athero) and stimulated HCASMC, transcription factor binding ChIPseq for TCF21, JUN, and JUND, and active enhancer histone modification H3K27Ac ChIPseq.



**Fig 4. Chromatin looping mapped in HCASMC with HiChIP.**

***A)*** Promoter-enhancer interaction signal (EIS) at 9p21 was significantly different for the two alleles at rs1537373. HCASMC ATACseq and ChIPseq data are shown. ***B)*** Pooled HiChIP identification of variants that alter chromosomal looping. Pre-HiChIP reference allele frequencies are graphed against Post HiChIP frequencies, with blue dots off the diagonal representing differences in looping in relation to allelic variation. ***C)*** Pooled HiChIP QQ plot indicates high significance of HCASMC clQTLs.

**APPROACH**

**Specific Aim 1**. To identify and characterize HCASMC lncRNAs that respond to disease related signaling pathways and reside in CAD loci.

**Rationale:** To integrate lncRNAs into modern biology and disease risk mechanisms, we need to investigate the response of lncRNAs to disease signaling pathways, in the context of cell specificity and biological setting, and these data need to be integrated with GWAS and genomic data that allows mapping of causal variation and causal genes for complex human diseases such as CAD. We hypothesize that many HCASMC lncRNAs are missing from current public annotation databases due to lncRNA tissue specificity in a cell type not well studied, inadequate sequencing depth, and lack of experiments under physiological contexts. Also, while there have been a number of well executed comprehensive studies cataloging the repertoire of lncRNAs in other cardiometabolic cells and tissues100-102 and a small number of lncRNAs have been investigated in cardiovascular cells,103-106 their regulatory framework has not been studied and incorporated into a highly characterized disease cell-specific genetic and epigenomic landscape that is critical for disease mechanistic investigation. We have shown that HCASMC harbor a significant amount of CAD risk10, 12, 93 and have mapped epigenomic traits, chromatin accessibility, chromatin looping, and CAD causal variants and genes in these cells. Further, we have studied lncRNAs at a mechanistic level in a few prominent CAD loci, 9p21.3 (ANRIL*/CDKN2B*)56, 57 and 6q23.2 (TARID*/TCF21*).10-14. We now propose to investigate in detail the contribution of lncRNAs to CAD mechanisms on a broader scale. To advance the field and serve as a starting point for studies in subsequent Aims, we propose to do an exhaustive de novo characterization of lncRNAs in HCASMC, by activation of growth factor, differentiation, and inflammatory pathways (PDGFD, TGFβ1, TNFα, ET1, IL-6) that are directly linked to CAD GWAS causal genes, in genetically diverse HCASMC lines **(Fig. 5)**. Ribosome depleted cellular RNA will be subjected to deep (2.5e8 reads/sample) strand-specific HiSeq 4000 sequencing, with de novo assembly methods to identify novel lncRNAs. These differentially regulated and novel lncRNAs will be intersected with a list of 503 CAD (non-lipid) loci, including all of the replicated genome-wide significant alleles, as well as those that have been identified in multiple meta-analyses to be associated with an FDR of 5% to generate a list of CAD lncRNAs. Hypothesis free WGCNA/MEGENA network construction using the large RNA sequencing datasets will integrate these lncRNAs into coexpression networks that will begin to associate HCASMC lncRNAs with pcGene partners and inform on functional molecular interactions. This is a required first step to investigate the role of lncRNAs in CAD and determine the mechanisms by which they regulate disease causal pcGene expression.

**Preliminary data:** *Identification of HCASMC lncRNAs in CAD loci:* To begin to investigate how HCASMC lncRNAs respond to cell state changes, we have identified lncRNAs in two of our most interesting RNAseq datasets. In one experiment, HCASMC were stimulated in vitro with media containing fetal bovine serum, and these cells compared to those that underwent 48 hrs of serum starvation. Stimulation with serum is well known to promote de-differentiation107, and we have established CAD causality for this process and shown that *TCF21* also promotes de-differentiation of HCASMC.11 The second dataset is thus one comparing HCASMC with and without TCF21 gain of function (GOF). For the serum stimulation experiment, polyA+ RNA was isolated and underwent RNAseq analysis, while for the TCF21 GOF study RNA isolation employed ribosome depletion. For both studies, ~3.0e7 reads were mapped against the LNCipedia and GENCODE databases to identify known lncRNAs. The objective of this study was to identify lncRNAs that might be specifically involved in the phenotypic modulation by intersecting these two datasets and focusing on the shared effect on HCASMC.



We identified 6,230 (serum) and 3,199 (*TCF21* GOF) lncRNAs that were expressed with > 50 average read counts. We identified 261 (serum) and 195 (*TCF21* GOF), 456 total differentially expressed (DE) lncRNAs at adjusted p <0.01 using DESeq2. Fifteen percent of lncRNAs were in common between the two experiments and were found to be associated with pcGenes related to “chromatin remodeling,” cell matrix,” and “cell cycle”, terms that reflect phenotypic modulation of HCASMC. Interestingly, isoforms of the lncRNA TARID showed bimodal regulation, with one short isoform (TARID:14-577 bp) being downregulated by serum (24.4 vs. 1.6 normalized count values corrected, *P*adj 0.043), and a long isoform (TARID:2-4,894 bp) being activated with TCF21 GOF (93.1 vs. 6458 normalized count values, *P*adj 2.1e-239). This interesting finding suggests that *TCF21* and *TARID* counter-regulate each other to modulate CAD risk. In total, 31 of 503 CAD GWAS loci (lead SNP plus LD SNPs with r2>0.8, ±100,000 bp) were found to harbor 34 DE lncRNAs.

**Experimental details:** *Cell culture and RNA sequencing -* Primary cultured HCASMC11, 13 will be serum starved for 48 hrs and then stimulated for 6 hrs with 5 ng/ml human recombinant TGFβ1 (R&D Systems), 4 hrs with 50 ng/ml human recombinant PDGFD (R&D Systems), 6 hrs with 10 ng/ml human recombinant TNFα (Sigma), 10 nM/L synthetic ET1 (Sigma), or human recombinant 100 ng/ ml IL-6 (R&D Systems), or sterile water **(Fig. 5)**. All conditions will be evaluated in triplicate, in HCASMC lines derived from one Caucasian and one African-America subject. Cells will be collected at the indicated time points and RNA will be isolated by Qiagen RNeasy Mini Kit. RNA quantity and quality will be characterized with Agilent Bioanalyzer and RNA samples that pass the QC metrics will be transferred to Genewiz where 36 strand-specific RNA library preparations (6 conditions x 3 triplicates x 2 lines) with rRNA depletion will be generated for sequencing on an Illumina 4000 machine at a depth of 2.5e8 reads per library.

*Identification of known and novel lncRNAs* –FastQC will confirm sufficient quality of the sequencing reads. Known lncRNAs will be identified by mapping with the STAR second pass algorithm onto the hg38 human genome and reads counted using the collections of lncRNAs from GENCODE v27 and LNCipedia 5.0. For detection of novel HCASMC lncRNAs43, 101, 108-110, BAM files will be merged across samples into a unique dataset. De novo RNA calls will be performed with a complete merged set of sequence reads across all experimental conditions in HCASMC, and separately for each condition in order to determine condition specific transcripts. De novo assembly will be performed using Trinity, a method for the reconstruction of transcriptomes that constructs multiple de Bruijn graphs for each cluster of sequence contigs to depict the transcriptional complexity of a single locus, processing them separately to generate full-length splicing isoforms and separate paralogous genes. Other tools will be employed in order to obtain a comprehensive and robust set of RNA isoforms, such as Trans-ABySS, a de novo short-read transcriptome assembler that addresses variation in local read densities. Putative novel lncRNAs will be overlapped with GENCODE v27 non-lncRNAs, RefSeq (NM annotation), and UCSC coding gene annotations, and further filtered using sophisticated sequence analysis algorithms PhyloCSF111 and iSeeRNA.112

*Identification of DE lncRNA genes in CAD loci* – The expression counts for lncRNAs identified above will be analyzed using the count based method DESeq2 at FDR 5% to determine significant DE lncRNAs. To map DE genes to CAD loci we will use a comprehensive collection of CAD GWAS SNPs5, 58 to derive a set of lead plus LD SNPs for CAD (using an LD r2 ≥ 0.8 in 1000 Genomes -ph3) and fetch the closest DE lncRNAs using Bedtools closest algorithm in a window of ±100kb **(Fig. 7)**. CAD SNPs mapped to lipid traits will be excluded as per the Global Lipids Consortium.113, 114 The outcome will be a set of potential CAD-related lncRNAs.

*Coexpression network analyses to assess global relationships among lncRNAs and between lncRNAs and pcGenes* - We will construct coexpression networks comprised of both lncRNAs and pcGenes using two global network modeling approaches to organize highly coregulated genes into subgroups, termed modules. The first method is the well-established Weighted Gene Co-expression Network Analysis (WGCNA)115, which can capture broad network modules of co-regulated genes. We will also use the Multiscale Embedded Gene Co-expression Network Analysis (MEGENA)116 to identify more compact and coherent network modules. Both methods use Spearman correlation coefficient and hierarchical clustering to identify modules containing coregulated molecules (lncRNAs and pcGenes in this case). WGCNA is based on agglomerative hierarchical clustering, uses a distance measure derived from Topological Overlap Matrix that considers pair-wise correlations as well as correlations of shared neighbors, and requires that each gene only belongs to one module. In contrast, MEGENA is based on divisive clustering, shortest path distance, and allows a gene to be in multiple modules at different scales. Our recent application of both approaches indicates that these methods are complementary and allow us to comprehensively capture the potential organization patterns among genes at different scales.117 For each WGCNA and MEGENA module, we will search for the presence of CAD lncRNAs and examine the within-module relationships between each CAD lncRNAs with other lncRNAs and pcGenes. The pcGenes within the same modules as the CAD lncRNAs will be annotated with biological pathways to assess the functional implications of the CAD lncRNAs.

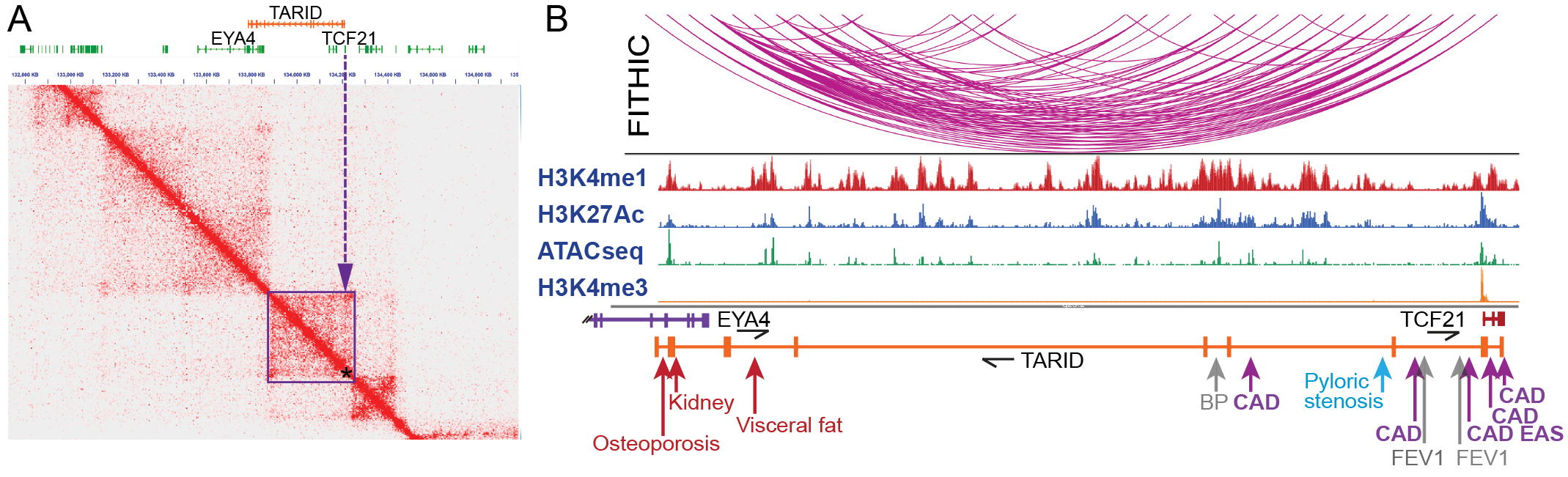
**Expected results:** Given that we are studying 3-times more stimuli with 10-times greater sequencing depth than previously used in our pilot study, we estimate that we will identify ~5-times more total regulated lncRNAs (~2250 DE lncRNAs), as well as those that colocalize with CAD GWAS loci (≥150 lncRNAs). With correlation and coexpression network analyses of this extensive RNA dataset, we expect to be able to connect lncRNAs to pcGenes that are also relevant to the vascular wall disease process. Having identity of the related pcGenes will allow the construction of new networks and pathways. In HCASMC, TGFβ1 promotes while PDGF inhibits SMC differentiation and promotes phenotypic modulation. Thus, we expect to find lncRNAs related to SMC lineage markers such as *ACTA2*, *TAGLN*, *MYH11* that will be downregulated by PDGFD and proliferation and migratory lncRNAs and pcGenes upregulated, while *TGFβ1* will identify genes with the opposite effect. Those genes activated by the pro-inflammatory cytokines TNFα and IL-6 will provide new and highly useful insights into lncRNA-pcGene relationships in the vessel wall. In addition to novel lncRNAs, ET1 stimulation will identify lncRNAs associated with expression of *ETS1, MYC,* and *FOS* and other genes that promote matrix remodeling and proliferation.118 Finally, in keeping with similar comprehensive efforts in other cardiometabolic cell types101, 102, we expect to identify ~300 novel lncRNAs that are highly specific to HCASMC and expressed at low levels and to also link them to their connected pcGenes in a data-driven unbiased manner, thus further informing on the molecular basis of disease risk in this cell type. We anticipate that the noted number of DE lncRNAs in CAD loci will provide sufficient starting point for the studies in the following Aims and we will be able to study in detail the downstream pathways of 20-30 novel disease related lncRNAs in **Aim 3**, thus integrating these lncRNAs and related pcGenes into the disease networks and pathways that this lab and others are working to characterize. To guide work in the following Aims, we will prioritize the identified lncRNAs based on the following criteria: *i)* localization in replicated CAD loci, *ii)* association with a pcGene, *iii)* proximity to the lead CAD variant (majority of lncRNAs act on nearest gene46, 119), *iv)* relevance to unique HCASMC processes.

**Potential problems/alternative strategies**: We have established all of the experimental and analytical protocols and do not anticipate technical difficulties. If we have the opportunity to study additional candidate lncRNAs, e.g., because of improved CRISPR methods, we certainly hold open the option to study lncRNAs that are not differentially regulated by the stimuli employed, but are novel to HCASMC and/or reside in loci that are predicted to provide new insights into SMC disease processes, based on identification of their related pcGenes. Regarding the relationship between lncRNAs and pcGenes our correlation and coexpression analyses cannot directly infer directionality. If pcGenes are correlated with or are within the same modules as lncRNAs, this does not mean that lncRNAs will regulate the pcGenes. We will address this limitation through genetic analyses (eQTLs) where directions of effects can be inferred in **Aim 2** and in **Aim 3** via PerturbSeq where individual lncRNAs will be perturbed and downstream PCGs identified.

**Specific Aim 2.** To associate HCASMC lncRNA expression and function with allelic variation that mediates genomic molecular traits and risk for CAD.

**Rationale:** Analyses with modest datasets have identified lncRNA eQTLs in enhancer regions, and shown them to be enriched among SNPs associated with various complex human diseases.36, 46, 119, 120 Thus, given the large number of lncRNAs and their greater number of regulatory regions46, we hypothesize that they are a common primary target of complex disease causal variation.46, 119, 120 Regarding the molecular mechanism by which lncRNAs regulate pcGene expression, and thus convey disease risk to the proteome, there is ample evidence that this class of regulatory RNAs interact with chromatin remodeling complexes21, 121 and regulate intrachromosomal association or looping.35, 36, 39 Thus, studies proposed here will investigate the hypothesis that disease associated variation is enriched for lncRNA eQTLs that in turn alter disease causal pcGene expression, and secondarily that this is mediated in part by lncRNA regulation of chromosomal configuration. We will map allele-specific expression (ASE) to identify eQTLs for lncRNAs identified in **Aim 1** that are differentially expressed and colocalize in CAD associated loci. To develop an informative large number of reads for lowly expressed lncRNAs, we will use the microfluidics-based multiplex PCR sequencing (mmPCR-Seq) approach122 focused on lncRNA genes to get sufficient reads to allow definitive ASE determination of eQTLs. The Montgomery (Co-I) lab has been applying mmPCR-seq to thousands of GTEx samples.123 By layering the identified lncRNA eQTLs onto extensive genomic and epigenomic HCASMC datasets along with causal CAD variants we have mapped10, 93, we have the opportunity to identify lncRNA eQTLs that are likely causal variants in the CAD loci. Further, while the precise mechanisms for how allelic variation directs looping is not well understood, we have mapped chromosomal looping QTLs (clQTLs) in HCASMC, and will investigate whether they are enriched in lncRNA or lncRNA/pcGene eQTLs, and whether there is convergence of clQTLs/eQTLs with CAD causal variation. To validate the lncRNA eQTLs that are also CAD SNPs, we will use CRISPR/Cas9 genome editing methods to verify they regulate lncRNA expression. These experiments may also perturb expression of the related pcGene, proving/confirming its identity, and if the lncRNA eQTL colocalizes with a CAD causal variant will implicate this pcGene and its encoded protein as causal, and those eQTLs that colocalize with clQTLs support looping as a mechanism for pcGene expression and disease risk. Studies proposed here will thus inform on causal CAD mechanisms by elucidating how CAD SNPs affect HCASMC lncRNAs and how lncRNAs regulate chromosomal configuration and pcGene expression.

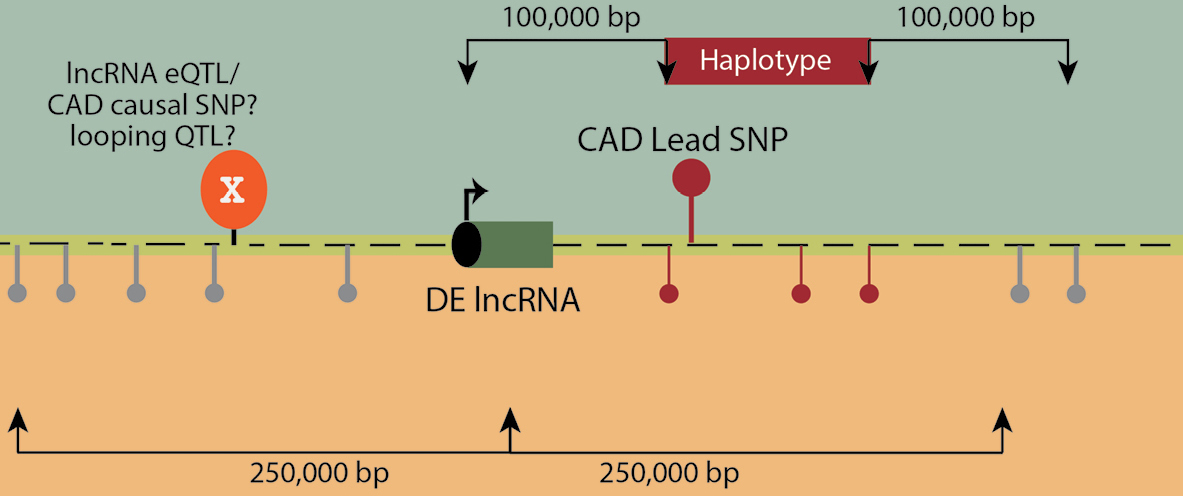
**Preliminary data:** *lncRNA eQTL and CAD GWAS overlap at TARID-TCF21 locus -* Our studies in the *TCF21* locus support the relevance of experiments proposed in this Aim. We are currently investigating the relationship between lncRNAs, pcGenes, and causal CAD SNPs, employing eQTLs and chromosomal looping as identified with HiChIP in HCASMC.10, 41 The CAD associated gene *TCF21* has a related antisense lncRNA TARID, that stretches ~300,000 bp upstream, to the *EYA4* gene. As visualized with Juicebox124, chromosomal interactions are restricted within a block stretching from *TCF21* to *EYA4*, suggesting multiple types of interactions related to TARID function **(Fig. 6A)**. FITHIC125 analysis of contacts identifies interactions between *TCF21* and the 3’ end of *TARID*, promoter-enhancer interactions, along with contacts among numerous locations within the *TARID* structural gene, presumably enhancer-enhancer interactions **(Fig. 6B)**. Epigenome and chromatin accessibility mapping show open and active chromatin in this region, as indicated by ATACseq peaks and H3K4me1 and H3K27Ac modifications, consistent with numerous enhancers in this region in HCASMC **(Fig. 6B)**. CAD GWAS studies have identified five independent CAD associated SNPs within and 5’ of *TCF21*. We have mapped each of these variants as highly significant eQTLs for both *TCF21* and *TARID* in HCASMC, consistent with TARID functioning as an intermediary, coupling the *eQTL* to *TCF21* expression. This hypothesis is consistent with molecular studies indicating that TARID tightly controls *TCF21* expression through regulation of methylation at its promoter region30, and GETx data showing that *TCF21* and *TARID* have identical patterns of tissue-specific gene expression. In addition, there are GWAS identified SNPs that are associated with pyloric stenosis126, blood pressure and hypertension127, visceral fat128, kidney disease129, osteoporosis130, and lung function131 within *TARID*. *TCF21* is expressed in all of these tissues (GTEx), especially during embryogenesis132, and thus is a likely candidate pcGene for each trait. Associated variants for these phenotypes are eQTLs, primarily for *TARID*, and presumably also affect *TCF21* expression, but the lncRNA *trans*-effect is not strong enough to be identified in HCASMC or the tested GTEx tissues. Although confirmatory genome editing studies remain to be completed, analysis of this instructive locus highlights the complexity of lncRNA – pcGene interactions in GWAS loci, and underscores the need for eQTL and CAD causal gene mapping as well as genomic, epigenomic, and looping data in disease relevant tissues to gain insights into the functional roles of lncRNAs in disease risk.



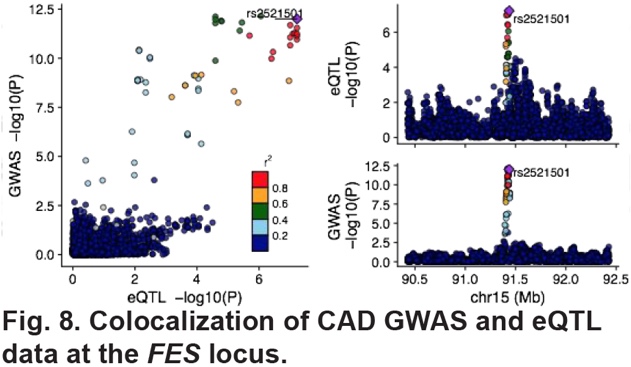
**Fig. 6. lncRNA *TARID* eQTLs, HiChIP looping, and disease variants at *TCF21* locus. *A)*** JUICEBOX visualization of HCASMC H3K27Ac HiChIP data at 6q23.2, showing interactions within the ~300,000 bp covered by *TARID* (purple box), including *TCF21* (star). ***B)*** HCASMC looping structure based on FITHIC analysis of HiChIP data, and histone modifications and ATACseq annotation suggest promoter-enhancer and enhancer-enhancer interactions. Purple arrows represent CAD variants that are also HCASMC eQTLs for both *TARID* and *TCF21*, red arrows indicate association of disease variants with *TARID* expression (eQTLs, GETx), blue arrow indicates association with *TCF21* expression (eQTL, GETx). Grey arrows/text indicate disease/phenotype associations without eQTL signals. Abbreviations: CAD, coronary artery disease; CAD EAS, CAD in East Asians; BP, blood pressure; FEV1, forced expiratory volume (lung fxn).

**Experimental details:** *Mapping lncRNA eQTLs in HCASMC*– Associating variation in RNAseq expression of lncRNAs with genotypes of 60 HCASMC, we will find eQTLs for these genes as we have done for pcGenes (see Significance). Another approach for mapping genetic regulation of gene expression is allele-specific expression (ASE) that compares relative lncRNA levels for the two alleles of the candidate gene in informative heterozygous individuals. As both alleles are exposed to the same environmental and trans-genetic factors, relative abundance of allele specific transcripts reflects *cis*-acting effects, and significant correlations can be identified with much smaller sample sizes than for classical eQTL analysis when influenced by technical noise. Correlation of ASE with genotype has been employed to detect eQTLs for lncRNAs that are expressed at lower levels than protein-coding genes119, as well as numerous human traits and diseases.133-147 Thus ASE will complement the classical eQTL mapping described above, and we will further strengthen the ASE approach with a high-throughput method, microfluidics-based multiplex PCR and sequencing (mmPCR-Seq).122 This system allows us to target and sequence the RNA of thousands of heterozygous positions and assess allelic imbalance with high fidelity.

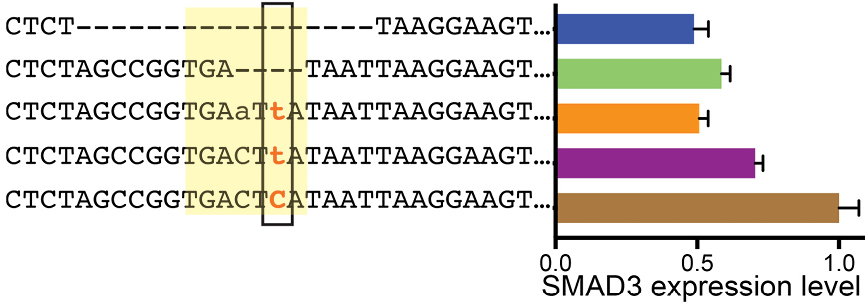
ASE looks for variants that are in exclusive heterozygosity with an ASE effect at a site – such that when a causal variant is heterozygous, we expect to observe an ASE effect.148 Importantly, the exonic (i.e., ASE-testing) variant is not expected to be causal – it merely tests the ASE effect from which we can map the causal variant. For every DE lncRNA localized in CAD loci (**Aim 1**), we will select lncRNA exonic variants for assessing ASE, choosing at least two common SNPs per gene. Thus, we will identify and test at least 300 variants for the ~150 lncRNAs identified in **Aim 1**. The microfluidics chips allow assessment of 48 samples x 48 PCR reactions, so with multiplexing 5 PCR assays per reaction, equals 11,520 assays/ chip and even with numerous control replicates all of the variants can be assessed for all of the lncRNAs with 2 chips. 122 Design of the multiplex PCR reactions is done with purpose-designed software, using published parameters. A significant innovation compared to conventional eQTL analysis, we will be able to run all of these reactions with a mid-capacity MiSeq run, and obtain ~1.0e4 reads per ASE variant. At this depth, we have an ability to detect allelic effects of 1.56-fold with a binomial probability of 0.05. In GTEx, the vast majority of lncRNA eQTLs exhibited larger effect sizes than 1.5-fold. eQTLs will be defined within 500 kb from the genomic coordinates of lncRNAs using a collection of 60 HCASMC lines from unrelated donors on which we have genotype and expression data, as described above (Significance) **(Fig. 7)**. We will identify an association between lncRNA-SNP pairs using the restricted ASE feature in the Robust Allele Specific QUAntification and quality controL (RASQUAL) algorithm.94 RASQUAL models read counts as a negative binomial and integrates between-individual differences, allele-specific signals and technical biases in sequencing-based cell phenotypes to provide identification of ASE associations. We will also use RASQUAL to combine the lncRNA ASE data with lncRNA eQTL data from standard RNAseq, for the subset of lncRNAs that are sufficiently expressed to support integrated analysis.



**Fig. 7. Mapping eQTLs for lncRNAs in CAD loci.** Differentially expressed lncRNAs identified within 200,000 bp surrounding the CAD linkage disequilibrium block (haplotype) will be associated with SNPs within 500,000 bp of the lncRNA as assessed with allele-specific expression. The lncRNA eQTLs (X) will be evaluated as possible causal variants for CAD, and as clQTLs for chromosomal looping.

****** *Intersection of lncRNA eQTLs with pcGene eQTLs, CAD GWAS SNPs, and looping QTLs* - We will use HCASMC epigenomic and chromosomal feature data to prioritize the most likely SNPs in loci where a number of SNPs in linkage disequilibrium show association with lncRNA expression. The first step to prioritize the lncRNA eQTLs will be: *i)* to identify those lncRNA eQTLs that are located in ATACseq localized regions of open chromatin, and *ii)* in particular colocalize in enhancer regions with H3K27Ac and/or H3K4me1 histone marks. This can be accomplished with the BEDTools *intersect* algorithm as we have done previously **(Fig. 3)**.10 *iii)* We will then investigate which of the prioritized lncRNA eQTLs are likely CAD causal variants. Our high density imputed genotyping data will be used with colocalization methods that evaluate the two types of association data in a statistical framework that quantifies the probability of each variant to be causal. Specifically, we will combine the eQTL/ASE data with two large CAD GWAS datasets5, 58, apply two colocalization methods with different statistical assumptions (eCAVIAR95 and SMR96) and use the union of their findings. We have used this approach to investigate the colocalization of HCASMC pcGene eQTLs and CAD GWAS variants **(Fig. 8)**.93 Further, we will use these colocalization tools and our pcGene eQTL data to search for lncRNA eQTLs that are also pcGene eQTLs. Studies in other systems have found only 20-30% of lncRNA eQTLs are also pcGene eQTLs119, 120, and while we expect much better colocalization, those lncRNAs without a related pcGene will not be excluded from further study. *iv)* Further, we can use the HiChIP looping data to better connect lncRNA eQTL variation with promoter regions **(Fig. 6)**. This can be accomplished with BEDTools. *vi)* Using these tools and approaches we will also look for colocalization of lncRNA eQTLs with HiChIP derived clQTLs.

*CRISPR/Cas9 genome editing to confirm lncRNA eQTL and related pcGene identification* - Experiments described here will be supported by Chong Park, PhD, former director of the UCSF CRISPR Core facility, who has worked extensively in this field to optimize each of the methods proposed here and is now a Senior Scientist in this Division (see Biosketch and letter), and by Research Associate Trieu Nguyen who has extensive experience with CRISPR use for genome editing **(Fig. 9)**.149 The goal of these experiments is to use genome editing with the CRISPR/Cas9 approach to validate the functional effect of mapped lncRNA eQTLs on gene expression, primarily, for lncRNAs but also for related pcGenes. We will formally show in HCASMC that genome editing of mapped lncRNA eQTLs results in modification of expression of the gene identified to be *cis*-regulated by the SNP associated with ASE. We will source HCASMC from our collection that are homozygous for the higher expressing eQTL allele, and determine by qPCR if conversion of one or both alleles is associated with decreased lncRNA expression. We will use the CRISPR-ERA algorithm150 to design three sgRNAs for each targeted lncRNA gene. We will include two validated sgRNAs as a positive control and two negative control sgRNAs that are not expected to cause any significant genetic perturbation. We will generate three pools of sgRNAs manufactured by Integrated DNA Technologies, with each pool containing one sgRNA per lncRNA, and each pool will be complexed with Cas9-NLS purified protein (QB3 MacroLab) and electroporated into HCASMC using the AMAXA 4D-Nucleofector System, providing high efficiency uptake of these ribonucleoproteins in a medium-throughput format. Each pool will be evaluated by qPCR for each targeted gene. For those lncRNA eQTLs where a pcGene was found to also share the eQTL, we will validate the identification of the pcGene by performing qRT-PCR to determine whether CRISPR downregulates expression of the related pcGene.



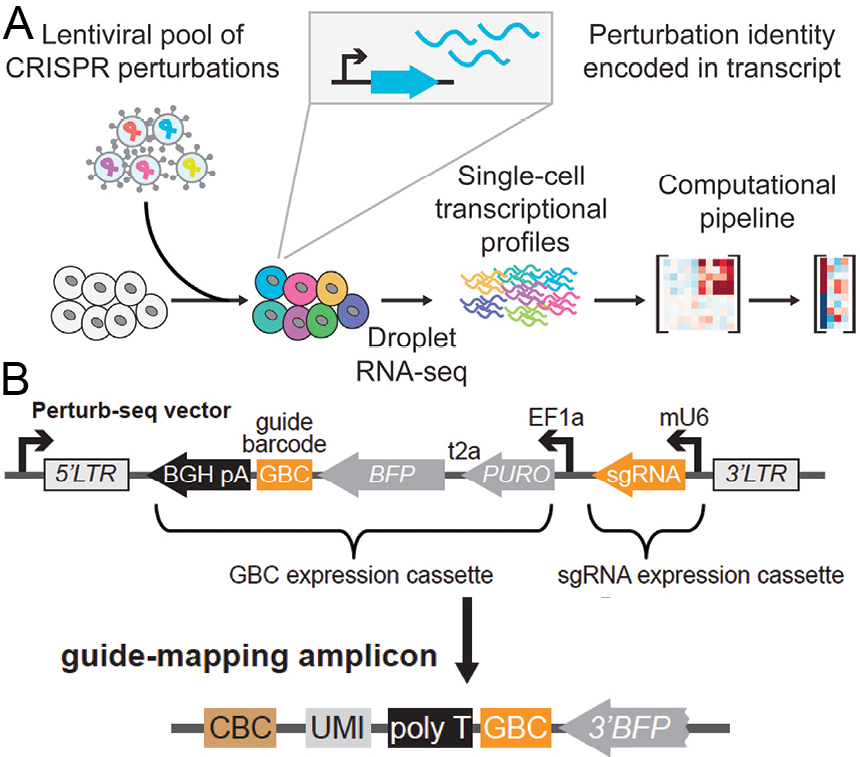
**Fig. 9. CRISPR validation that CAD causal variant rs17293632 and the encompassing AP1 motif regulate *SMAD3* expression.** Two guide RNAs along with a Cas9+GFP expression construct were transfected, and GFP positive cells selected by FACS, clonally expanded and evaluated by genome sequencing and *SMAD3* RNA expression. Changing the CC genotype to CT and deleting various basepairs in the AP-1 motif (yellow) decreased SMAD3 expression.

**Expected results:** Based on our previous experience and published data119, we expect to find informative SNPs for 100-120 of the 150 DE CAD locus lncRNAs and 500-600 putative regulatory variants (~5 SNPs/lncRNA) across all loci using eQTL and ASE tests alone. The list of putative eQTLs will be winnowed down using bioinformatic application of genome annotation and epigenetic data. We anticipate that we will find at least ~60 eQTLs that are high probability, as identified by significant ASE and residence in open chromatin as assessed by existing HCASMC ATACseq data, and in enhancer regions that are marked with H3K4m1 and H3K27Ac histone modifications119, 120 as previously mapped in this cell type10, and provide high confidence signals by the colocalization algorithms. Experimental evaluation of the transcriptional activity of these variants with CRISPR/Cas9 editing will validate at least half as bona fide regulatory variation. Similar to *cis* mRNA eQTLs, lncRNA eQTLs will be significantly enriched at transcription factor binding sites, splicing sites, microRNA binding sites, etc.151 Some portion of these lncRNA eQTLs are expected to also colocalize with CAD causal variants and clQTLs, and will be given priority in follow-up functional experiments in **Aim 3**. Because of our extensive mapping of mRNA eQTLs in this cell type93, we anticipate that up to 50% of the lncRNA eQTLs will also be mRNA eQTLs119, 120, verifying many of the lncRNA-pcGene connections identified in **Aim 1** and providing new avenues of study for mechanisms of CAD risk. Taken together, these highly unique data will further advance our understanding of lncRNA functional mechanisms for regulating pcGene expression, and provide compelling details regarding the role of these regulatory RNAs as intermediate links in the chain of complex disease risk from allelic variation to the transcriptome and ultimately the proteome.

**Potential problems/alternative strategies**: In those loci where we have difficulty identifying the single causal eQTL variant, because of tight linkage disequilibrium, CRISPR failure, etc., we can still pursue functional studies of the relevant lncRNA in **Aim 3**, which will be identified as the single lncRNA related to the candidate eQTL variants. In addition, combinations of epigenomic and QTL data have been shown to increase power for QTL mapping and we will investigate combining ASE and epigenome data where possible.152

**Specific Aim 3.** To map the downstream genes that are regulated by HCASMC lncRNAs and determine how their gene expression program alters SMC functions.

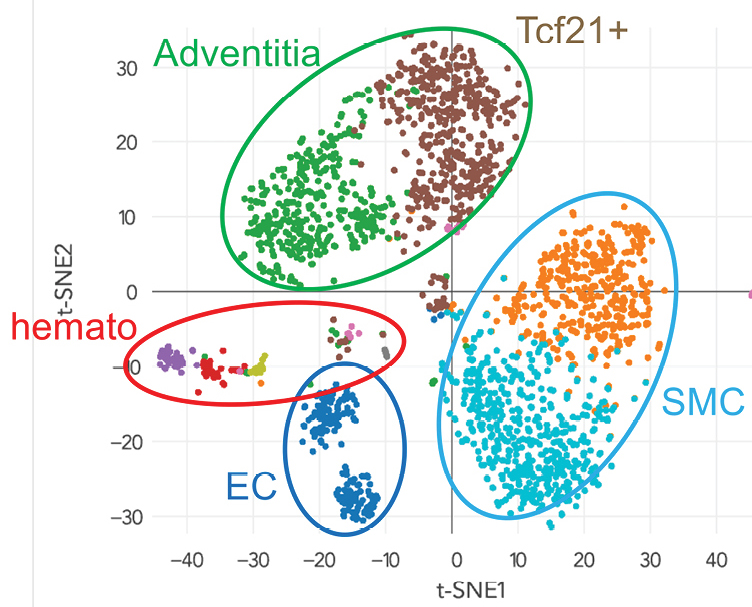
**Rationale:** For complex diseases such as CAD the primary risk is of genetic origin, mediated by small effect size common variation, and while there are hundreds of these variants and their related genes for each particular disease, there are likely to be a much smaller number of pathways.153-160 Thus, the identification of disease associated lncRNAs and their related pcGenes, and study of their relationships with other CAD genes in signaling networks will reduce the dimensionality of GWAS data, contribute to our understanding of the role of lncRNAs in disease pathophysiology, and point to drug targeting opportunities. We have made significant progress connecting CAD associated functional HCASMC genes **(Fig. 2)**, thus far highlighting a likely disease role for SMC differentiation and phenotypic modulation11, 13, as well as apoptosis and efferocytosis.56, 57 While these pathways have been implicated previously and studied in various model systems, the human data is now providing causal information that links them to human disease risk, and identifying new genetic interdependence. Thus, a primary goal of this Aim is to characterize the gene networks that are regulated by the disease lncRNAs identified in **Aims 1** **and 2**. We propose here to use highly-multiplexed CRISPR inhibition (CRISPRi) coupled with single-cell RNAseq (PerturbSeq) to characterize downstream signaling networks of CAD lncRNAs **(Fig. 10)**. Although this is a relatively new methodology, we have extensive experience with the 10x Genomics system for inDrop single cell RNAseq (scRNAseq)161-163 **(Fig. 11)** and are currently using PerturbSeq to investigate TCF21 downstream signaling.162 We will intersect networks from CAD related lncRNAs with those identified for CAD causal pcGenes such as *TCF21* and *SMAD3*, as well as diseased human coronary artery and mouse *ApoE-/-* aorta scRNAseq data, to identify disease relevant functional pathways in HCASMC. We will study 20-30 lncRNAs that are: *i)* differentially regulated in stimulated HCASMC as per **Aim 1**, *ii)* reside in a CAD associated locus, *iii)* are regulated by variation that has been identified as a specific lncRNA eQTL, *iv)* and preferably with an eQTL that is also causal toward CAD, and/or *v)* also functions as a pcGene eQTL and/or clQTL. The second goal of this Aim is to investigate the functional role of the CAD lncRNAs and their regulated pcGenes, investigating lncRNA phenotypic effects on HCASMC, including migration, proliferation, apoptosis,inflammatory profile (cytokine and chemokine expression), phenotypic modulation (lineage marker expression), etc. Together, the proposed studies will test the working hypothesis that lncRNAs contribute to HCASMC-specific signaling networks, and show that signals through these lncRNAs regulate basic HCASMC phenotype and function.

****

**Fig. 10. Overview of PerturbSeq. *A)*** Experimental workflow. ***B)*** With single cell RNAseq, every transcript from a given cell is tagged with the same cell barcode (CBC). In PerturbSeq, each sgRNA is also expressed along with a unique guide barcode (GBC), which allows determination of which guide was transduced into a given cell. Adapted from Adamson, et al., Cell 167, 1867.

**Experimental details: *Aim 3.1. Employ multiplexed targeting of disease associated lncRNAs to profile their cellular downstream signaling in HCASMC that regulates CAD risk and/or chromosomal looping -*** Collaborator Stanley Qi, PhD, Stanford Faculty in Bioengineering, helped develop PerturbSeq and serves as consultant on this grant (see Biosketch, letter of support), to advise on methodology and data interpretation.150, 164 CRISPR interference (CRISPRi) uses a catalytically-inactive Cas9 enzyme (dCas9) fused to the transcriptional repressor Kruppel-associated box (KRAB) domain.164 Upon targeting of dCas9-KRAB to a gene by a single-guide RNA (sgRNA), the KRAB domain recruits chromatin modifying complexes to silence transcription. This system has been shown to produce highly effective and uniform gene knockdown with minimal off-target effects. We will perform multiplexed CRISPRi by transducing HCASMCs with pooled lentivirus containing sgRNAs targeting *TCF21*, *SMAD3*, and *PDGFD* as control genes, and lncRNAs that have been identified through studies in **Aims 1 and 2**. During transduction, each cell will receive on average ≤1 sgRNA. Because each sgRNA is paired with a polyadenylated transcript containing a unique guide barcode (GBC) (**Fig. 10B**), and because every transcript from a given cell (including the GBC-containing transcript) is tagged with the same cell barcode (CBC), droplet-based single-cell RNAseq can be used to demultiplex these perturbations by determining which sgRNA (GBC) was transduced into a given cell (CBC) (**Fig 10B**). In addition to facilitating high-throughput interrogation of genetic perturbations, this approach also minimizes technical variations between different gene knockdowns because all perturbations occur in the same population of cells, leading to identical conditions during cell culture, lysis, reverse transcription, PCR, library preparation and sequencing.

*Design of CRISPRi guide vectors* – For known lncRNAs we will reference the published sgRNA sequences for the recent CRISPRi screen for cellular proliferation for sgRNA design165, otherwise we will design primers with the CRISPR-ERA algorithm.150 We will design three sgRNAs for each targeted lncRNA gene, and include two validated sgRNAs. We will clone these guide RNA protospacer sequences into a commercially-available sgRNA expression vector library (Addgene #85968), which contains >100,000 unique 18-nt guide GBC sequences as well as the blue fluorescent protein BFP) and puromycin resistance genes. The final constructs will drive expression of the sgRNA and a transcript containing BFP followed by the unique 18-nt GBC and a polyadenylation signal (**Fig. 10**). For each construct, we will determine the relationship between the sgRNA and its corresponding GBC by Sanger sequencing. For dCas9-KRAB expression, we will use a commercially available dCas9-KRAB expression vector (Addgene #60954) that also expresses the fluorescent protein mCherry. All of these reagents are in hand and have been validated in HCASMC.



**Fig. 11. Single cell RNA-seq analysis, t-SNE plot of cells grouped by K-means clustering.** inDrop (10X Genomics) analysis of ~2500 cells harvested from the aortic root of an *ApoE-/-* mouse. RNA sequencing of individual cells allowed their clustering into groups of smooth muscle (SMC), endothelial (EC), hematopoietic (hemato) and adventitial cells. *Tcf21* lineage traced cells group together within the adventitial cluster. Two groups are seen for adventitial, endothelial, and smooth muscle cells, a finding of unknown significance.

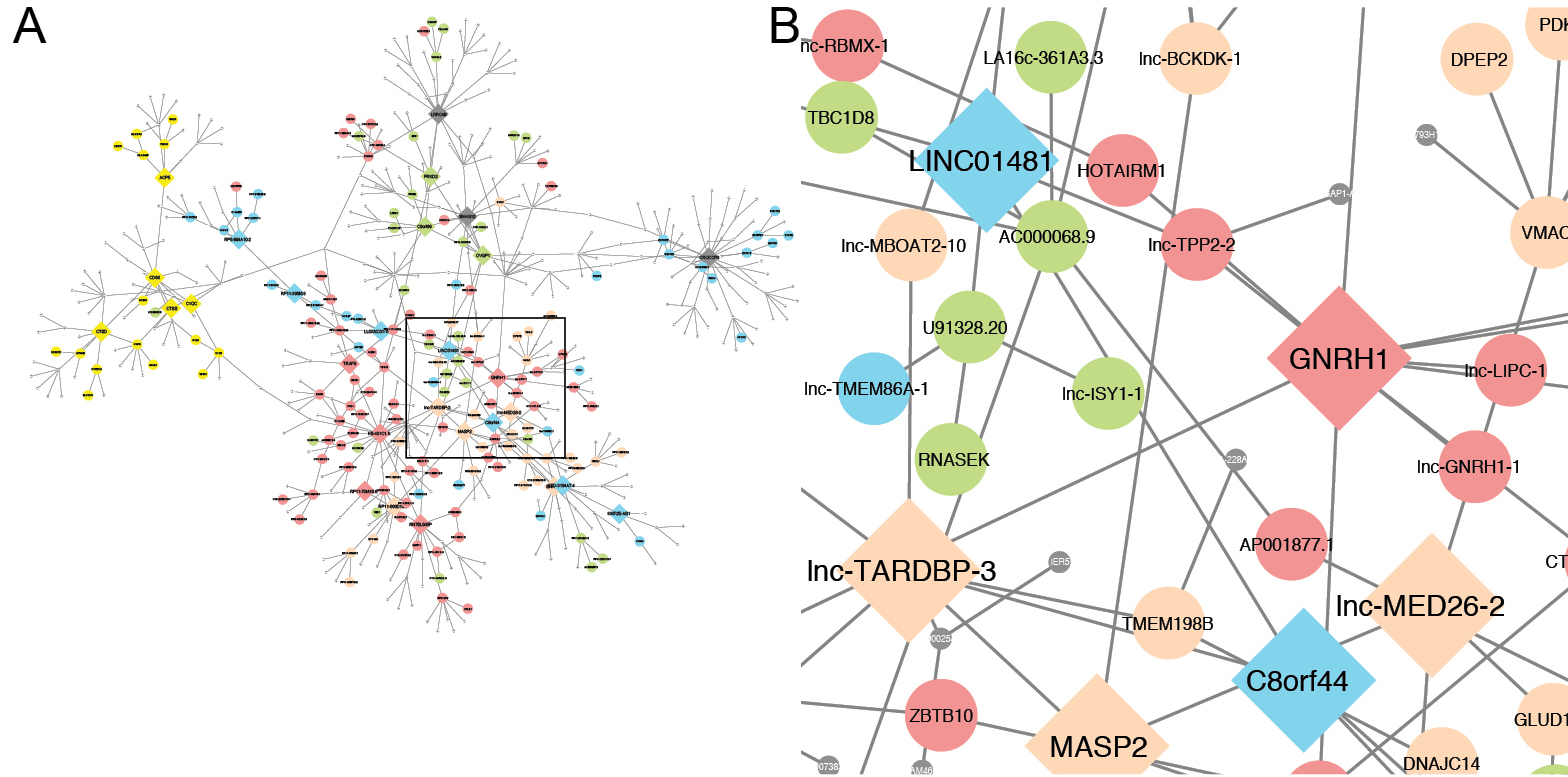
*Lentiviral transduction of SMCs* - All lentivirus will be produced by the Stanford Lentiviral core facility in HEK293T cells with standard packaging vectors. The targeted multiplicity of infection (MOI) for sgRNA-GBC constructs will be 0.5, and the MOI for dCas9-KRAB will be ~5-10, based on previous optimization experiments. After pooled transduction, HCASMCs will be grown in the presence of puromycin for 7 days to select cells that have received at least one sgRNA-GBC construct and to allow CRISPRi-mediated gene silencing to take effect. Cells will then be sorted for mCherry/BFP dual fluorescence to ensure receipt of both Cas9-KRAB and sgRNA-GBC constructs.

*Single-cell RNAseq readout* - We will perform droplet-based single-cell RNAseq with the 10X Genomics platform as we have described **(Fig. 11)**161, 162, with the addition of an extra PCR step designed to further amplify the transcript containing the GBC. This will ensure robust detection of this transcript and thus correct perturbation-to-cell assignment. Previous PerturbSeq studies have shown that confident assessment of a perturbation’s effect on individual genes (with 80% sensitivity and 90% specificity) is possible with as few as 100 cells/guide.164 Thus, with 90 guides targeting 30 genes, we will require approximately 10,000 cells.

*Characterizing differentially expressed genes/pathways downstream of candidate lncRNAs-*We will first assess the efficacy of each sgRNA in achieving knockdown of its target gene. For each guide, *mean* knockdown of the targeted gene will be calculated by comparing the distributions of the target gene expression in guide-transduced cells versus cells transduced with a control guide. 95% confidence intervals will be assigned to these knockdown estimates via bootstrapping. First, we will determine which genes are significantly affected by each perturbation, relative to the control cells transduced with non-targeting sgRNAs. Specifically, for each perturbation, we will compare the expression distribution of each gene in all cells receiving that perturbation with the expression distribution of that gene from the control cells using a maximum-likelihood test with a p-value threshold of 0.01. This will result in a list of genes that are significantly affected by each perturbation (guide). This list will be mapped initially against functional cellular pathways with Ingenuity, GSEA, and Gene Ontology algorithms. The results will also be compared with the pcGenes correlated or coexpressed with the lncRNAs as defined in **Aim 1** to assess consistency and differences between the different perturbation conditions.

Next, for each perturbation, we will assess the overlap of genes affected by that perturbation with genes affected by perturbation of known SMC CAD genes, *TCF21*, *SMAD3*, *PDGFD*, etc. This will allow a quantitative comparison of each perturbation’s relative similarity to the control perturbations. We will also determine whether each perturbation is more similar to one of the control CAD genes than expected by chance. To obtain a distribution of similarity scores under the null hypothesis that the perturbation assignment does not influence similarity to its nearest control gene, we will perform ~5000 permutations of the similarity analysis using groups of randomly mixed cells from the dataset. An experimental perturbation will be judged to be more similar to the control gene than expected by chance if its similarity score falls >2 standard deviations from the mean permuted similarity score. As an initial analysis, we will investigate similarity in regulated gene patterns among lncRNAs that responded to the same stimuli in **Aim 1**.

***Aim 3.2. Modeling of gene networks based on PerturbSeq dataset -*** We will explore the global relationships among lncRNAs and between lncRNAs and potential downstream pcGenes using various network modeling approaches. First, we will employ WGCNA and MEGENA, as described in **Aim 1**, to define modules of highly coexpressed genes (both lncRNAs and pcGenes) at different compactness scales across all cell transcriptome profiles in PerturbSeq. As scRNAseq typically measures a few thousand genes per cell, dropout transcripts that are expressed but not measured will be imputed using SAVER (bioRxiv 138677), an experimentally validated scRNAseq imputation method, before network construction. While the WGCNA and MEGENA coexpression networks can reflect the broad gene organization, these networks cannot reveal the detailed gene-gene regulatory relationships. Bayesian networks (BNs) are directed acyclic graphs that can demonstrate the directed causal relationships between gene pairs. In BNs, an edge is directed from a parent to a child node and state of each node is probabilistically predicted by the states of its parent nodes. We will use the Rimbanet package166, which is based on Monte Carlo Markov Chain, to construct BNs that best explain the observed PerturbSeq data. However, because of the Markov equivalence concept, many of the edge directions in such networks can be changed without affecting how well the model fits the data. Here we will take advantage of several priors that carry directional information to narrow down the searching-space in BN construction. The priors include *i)* the correlations among genes, *ii)* eQTL information (e.g. a *cis-*acting gene is more likely to be a parent node of *trans*-acting genes coinciding on the same eQTL), *iii)* directionality inferred by PerturbSeq by setting the lncRNAs perturbed as upstream nodes and their downstream differential pcGenes identified in **Aim 3.1** as target nodes, and *iv)* transcription factors and their known targets. We will incorporate these priors and construct 1000 plausible networks that fit our PerturbSeq data well, and will combine them to obtain a consensus network by taking only edges appearing in at least 30% of the 1000 networks. This analysis will reveal detailed regulatory information between lncRNAs and downstream genes and pathways, as we have done for lung using GTEx RNAseq data **(Fig. 12)**. The networks will be compared with those derived in **Aim 1** to assess consistency.



**Fig. 12. Example of a Bayesian network comprised of lncRNAs and pcGenes based on GTEx lung RNAseq data. *A)*** Bayesian subnetwork connecting genes in multiple pathways (colored nodes) related to pulmonary hypertension. ***B)*** An enlarged view of a segment of the Bayesian subnetwork (blue box in A) showing certain lncRNAs such as LINC01481 and lnc-TARDPB-3 and their neighboring genes.

***Aim 3.3. Investigate cellular functions of identified lncRNAs in HCASMC-*** For a limited number of lncRNAs, 10-15 total, we will examine the effects on HCASMC cell state and function, as guided by the downstream pathway, GO annotation, and network analyses from studies in **Aim 3.1 and 3.2**. We will be interested to investigate lncRNAs that have unique and previously uncharacterized signaling pathways in the context of SMC biology, but will also be interested to expand the phenotypic modulation paradigm that we have developed from CAD causal genes **(Fig. 2)**. For these studies, we will use the guide lentiviruses that were developed and employed in **Aim 3.1**. Use of dCas9KRAB and guide lentiviruses for PerturbSeq provides stable gene knockdown for at least two weeks, providing ample time for the following proposed assays.165 We will compare cellular effects of different stimuli, consider direction of effect, link lncRNAs to specific functions, and consider lncRNAs grouped by stimulus response.

lncRNA expression will be knocked down in HCASMC with lentivirus transduction as in **Aim 3.1**. Proliferation, apoptosis, migration, differentiation status of the genetically modified cells will be evaluated with methods we have employed extensively.11, 167-169 Further, we will analyze the control PerturbSeq data for expression of gene panels that we have identified with scRNAseq on SMC lineage traced cells in mouse undergoing phenotypic modulation in vivo, and validated in human scRNAseq studies, including *TCF21, LGALS3, TAZ, RUNX2*. Finally, we will assess the inflammatory state of the cells with qRT-PCR for inflammatory markers such as *IL1α, TNFα, and IL6*.

**Expected findings:** We expect that our PerturbSeq approach that directly targets individual lncRNAs, coupled with advanced network modeling, will delineate the regulatory circuits of individual lncRNAs and group lncRNAs that share similar functions. We also expect that the effects of some of the lncRNAs on downstream signaling will closely recapitulate those of one or more of the control CAD genes. In particular, evidence suggests that TGFβ, SMAD3, ET1 and PDGF signaling are highly important in vascular wall pathology, and we expect downstream signaling networks for a number of the lncRNA perturbations, and the resulting perturbations of their related pcGenes, will show very significant overlap with these CAD gene perturbations. We also expect that the lncRNAs will be closely connected with the known CAD genes and pathways in gene networks. Importantly, this type of result will significantly expand the repertoire of both lncRNAs and pcGenes that contribute to specific pathways, opening up additional, possibly highly cell-specific, therapeutic targeting opportunities. Features of the downstream signaling pathways will enrich our ongoing disease transcriptional network characterization, for example allowing us to assign directionality to the different lncRNA-pcGene pairs. For example, we expect to provide evidence that knockdown of *SMAD3* related factors will have an opposite effect to the knockdown of *TCF21*. It is anticipated that alterations in lncRNA expression level will significantly impact cell fate decisions such as proliferation, differentiation, apoptosis, migration, inflammatory state, in a direction that is consistent with observed cellular behavior in the disease models. These studies will teach us much regarding the biology of SMC and the contribution of this cell type to the different phases of vascular disease, but more importantly inform on the role that lncRNAs play in the causal relationship between disease variation and protein genes, and the molecular mechanisms by which they accomplish this role.

**Potential Problems & Alternative Strategies**: Despite the fact that PerturbSeq is a recent development, it is a combination of two fairly well-established techniques, CRISPR gene targeting and Drop-seq, both of which are in common use in the lab.162 However, in the unlikely event that we are unable to perform these experiments in a multiplexed manner, we will perform CRISPRi or shRNA-mediated knockdown of these genes individually with bulk RNA sequencing. This approach lacks the advantages mentioned above, but would still achieve the overall objective of this Aim. In addition, simple hierarchical clustering of perturbations can be achieved by constructing synthetic bulk RNAseq profiles from all cells subjected to a given perturbation, as has been done previously.170 Also, there are multiple R packages available for analyzing PerturbSeq data.170, 171

**DELIVERABLES AND FUTURE DIRECTIONS**

These studies will integrate lncRNAs, their regulatory variation, and target pcGenes into the component of CAD risk that resides in the vascular wall, and more broadly will identify mechanisms by which this class of regulatory molecules interact with the genome to provide for exquisite control of pcGene expression. We will be able to integrate lncRNAs into existing causal CAD pathways such as HCASMC phenotypic modulation, identify new networks that will encompass new CAD loci and point to novel functional pathways. Future studies will build on these findings to further investigate at a mechanistic level how key disease lncRNAs mediate the connection between upstream epigenetic signaling and pcGene expression and function. Further, identified HCASMC CAD associated lncRNAs that appear to regulate intra-chromosomal looping will be investigated to better understand the mechanisms by which lncRNAs function to organize chromatin domains to coordinate long-range gene activation. There is every indication that the huge lncRNA repertoire serves to transmit genetic signaling to chromatin, and in concert with epigenetically mediated effects, regulates a wide range of human traits and diseases, and the expectation is that studies such as those proposed here will provide much deeper understanding of genomic and genetic cellular mechanisms of CAD.

**BIBLIOGRAPHY AND REFERENCES CITED**

1. Daviglus ML, Talavera GA, Aviles-Santa ML, Allison M, Cai J, Criqui MH, Gellman M, Giachello AL, Gouskova N, Kaplan RC, LaVange L, Penedo F, Perreira K, Pirzada A, Schneiderman N, Wassertheil-Smoller S, Sorlie PD and Stamler J. Prevalence of major cardiovascular risk factors and cardiovascular diseases among Hispanic/Latino individuals of diverse backgrounds in the United States. JAMA : the journal of the American Medical Association. 2012;308(17):1775. Epub 2012/11/03. PubMed PMID: 23117778.

2. Gupta R and Gupta KD. Coronary heart disease in low socioeconomic status subjects in India: "an evolving epidemic". Indian heart journal. 2009;61(4):358. Epub 2010/07/20. PubMed PMID: 20635739.

3. Ikem I and Sumpio BE. Cardiovascular disease: the new epidemic in sub-Saharan Africa. Vascular. 2011;19(6):301. Epub 2011/09/24. PubMed PMID: 21940758.

4. Almahmeed W, Arnaout MS, Chettaoui R, Ibrahim M, Kurdi MI, Taher MA and Mancia G. Coronary artery disease in Africa and the Middle East. Therapeutics and clinical risk management. 2012;8:65. Epub 2012/03/01. PubMed PMID: 22368447; PMCID: 3284217.

5. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikainen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, Konig IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, Hall AS, Hamsten A, Harris TB, Hazen SL, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ, Kooner JS, Kullo IJ, Lehtimaki T, Loos RJ, Melander O, Metspalu A, Marz W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M and Consortium CAD. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015;47(10):1121. PubMed PMID: 26343387; PMCID: PMC4589895.

6. Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Muhleisen TW, Muhlestein JB, Munzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nothen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schafer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Witteman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, Marz W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J and Samani NJ. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet. 2011;43:333. Epub 2011/03/08. PubMed PMID: 21378990.

7. Assimes TL, Knowles JW, Basu A, Iribarren C, Southwick A, Tang H, Absher D, Li J, Fair JM, Rubin GD, Sidney S, Fortmann SP, Go AS, Hlatky MA, Myers RM, Risch N and Quertermous T. Susceptibility locus for clinical and subclinical coronary artery disease at chromosome 9p21 in the multi-ethnic ADVANCE study. Hum Mol Genet. 2008;17(15):2320. Epub 2008/04/30. PubMed PMID: 18443000.

8. Knowles JW, Assimes TL, Boerwinkle E, Fortmann SP, Go A, Grove ML, Hlatky M, Iribarren C, Li J, Myers R, Risch N, Sidney S, Southwick A, Volcik KA and Quertermous T. Failure to replicate an association of SNPs in the oxidized LDL receptor gene (OLR1) with CAD. BMC Med Genet. 2008;9:23. PubMed PMID: 18384690; PMCID: PMC2322963.

9. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups K, Konig IR, Cazier JB, Johansson A, Hall AS, Lee JY, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A, Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K, Lundmark P, Lyytikainen LP, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altshuler D, Balmforth AJ, Barroso I, Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R, Doney AS, Mokhtari NE, Eriksson P, Fischer K, Fontanillas P, Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han BG, Hunt SE, Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C, Leander K, Lokki ML, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S, Morris AD, Muller-Nurasyid M, Nikus K, Peden JF, Rayner NW, Rasheed A, Rosinger S, Rubin D, Rumpf MP, Schafer A, Sivananthan M, Song C, Stewart AF, Tan ST, Thorgeirsson G, Schoot CE, Wagner PJ, Wells GA, Wild PS, Yang TP, Amouyel P, Arveiler D, Basart H, Boehnke M, Boerwinkle E, Brambilla P, Cambien F, Cupples AL, de Faire U, Dehghan A, Diemert P, Epstein SE, Evans A, Ferrario MM, Ferrieres J, Gauguier D, Go AS, Goodall AH, Gudnason V, Hazen SL, Holm H, Iribarren C, Jang Y, Kahonen M, Kee F, Kim HS, Klopp N, Koenig W, Kratzer W, Kuulasmaa K, Laakso M, Laaksonen R, Lee JY, Lind L, Ouwehand WH, Parish S, Park JE, Pedersen NL, Peters A, Quertermous T, Rader DJ, Salomaa V, Schadt E, Shah SH, Sinisalo J, Stark K, Stefansson K, Tregouet DA, Virtamo J, Wallentin L, Wareham N, Zimmermann ME, Nieminen MS, Hengstenberg C, Sandhu MS, Pastinen T, Syvanen AC, Hovingh GK, Dedoussis G, Franks PW, Lehtimaki T, Metspalu A, Zalloua PA, Siegbahn A, Schreiber S, Ripatti S, Blankenberg SS, Perola M, Clarke R, Boehm BO, O'Donnell C, Reilly MP, Marz W, Collins R, Kathiresan S, Hamsten A, Kooner JS, Thorsteinsdottir U, Danesh J, Palmer CN, Roberts R, Watkins H, Schunkert H and Samani NJ. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet. 2012;45(1):25. Epub 2012/12/04. PubMed PMID: 23202125.

10. Miller CL, Pjanic M, Wang T, Nguyen T, Cohain A, Perisic L, Hedin U, Betsholtz C, Ruusalepp A, Franzen O, Assimes TL, Montgomery SB, Schadt EE, Bjorkegren JLM and Quertermous T. Integrative functional genomics identifies regulatory mechanisms at coronary artery disease loci. Nature communications. 2016(7):12092.

11. Nurnberg ST, Cheng K, Raiesdana A, Kundu R, Miller CL, Kim JB, Arora K, Carcamo-Orive I, Xiong Y, Tellakula N, Nanda V, Murthy N, Boisvert WA, Hedin U, Perisic L, Aldi S, Maegdefessel L, Pjanic M, Owens GK, Tallquist MD and Quertermous T. Coronary artery disease associated transcription factor TCF21 regulates smooth muscle precursor cells that contribute to the fibrous cap. PLoS Genet. 2015;11(5); PMCID: in process.

12. Sazonova O, Zhao Y, Nurnberg S, Miller C, Pjanic M, Castano VG, Kim JB, Salfati EL, Kundaje AB, Bejerano G, Assimes TL, Yang X and Quertermous T. Characterization of TCF21 downstream target regions identifies a transcriptional network linking multiple independent coronary artery disease loci. PLoS Genet. 2015;11(5); PMCID: in process.

13. Miller CL, Anderson DR, Kundu RK, Raiesdana A, Nurnberg ST, Diaz R, Cheng K, Leeper NJ, Chen CH, Chang IS, Schadt EE, Hsiung CA, Assimes TL and Quertermous T. Disease-Related Growth Factor and Embryonic Signaling Pathways Modulate an Enhancer of TCF21 Expression at the 6q23.2 Coronary Heart Disease Locus. PLoS Genet. 2013;9(7):e1003652. Epub 2013/07/23. PubMed PMID: 23874238; PMCID: 3715442.

14. Miller CL, Haas U, Diaz R, Leeper NJ, Kundu RK, Patlolla B, Assimes TL, Kaiser FJ, Perisic L, Hedin U, Maegdefessel L, Schunkert H, Erdmann J, Quertermous T and Sczakiel G. Coronary Heart Disease-Associated Variation in TCF21 Disrupts a miR-224 Binding Site and miRNA-Mediated Regulation. PLoS Genet. 2014;10(3):e1004263. Epub 2014/03/29. PubMed PMID: 24676100.

15. Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, Ding H, Butty VL, Torrey L, Haas S, Abo R, Tabebordbar M, Lee RT, Burge CB and Boyer LA. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. Cell. 2013;152(3):570. PubMed PMID: 23352431; PMCID: PMC3563769.

16. Meola N, Pizzo M, Alfano G, Surace EM and Banfi S. The long noncoding RNA Vax2os1 controls the cell cycle progression of photoreceptor progenitors in the mouse retina. RNA. 2012;18(1):111. PubMed PMID: 22128341; PMCID: PMC3261733.

17. Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y, Grant GD, Horlings HM, Shah N, Umbricht C, Wang P, Wang Y, Kong B, Langerod A, Borresen-Dale AL, Kim SK, van de Vijver M, Sukumar S, Whitfield ML, Kellis M, Xiong Y, Wong DJ and Chang HY. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet. 2011;43(7):621. PubMed PMID: 21642992; PMCID: PMC3652667.

18. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S and Chang HY. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010;464(7291):1071. PubMed PMID: 20393566; PMCID: PMC3049919.

19. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, Blencowe BJ, Prasanth SG and Prasanth KV. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell. 2010;39(6):925. PubMed PMID: 20797886; PMCID: PMC4158944.

20. Miao Y, Ajami NE, Huang TS, Lin FM, Lou CH, Wang YT, Li S, Kang J, Munkacsi H, Maurya MR, Gupta S, Chien S, Subramaniam S and Chen Z. Enhancer-associated long non-coding RNA LEENE regulates endothelial nitric oxide synthase and endothelial function. Nature communications. 2018;9(1):292. PubMed PMID: 29348663.

21. Alcid EA and Tsukiyama T. ATP-dependent chromatin remodeling shapes the long noncoding RNA landscape. Genes Dev. 2014;28(21):2348. PubMed PMID: 25367034; PMCID: PMC4215180.

22. Wang KC and Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell. 2011;43(6):904. PubMed PMID: 21925379; PMCID: PMC3199020.

23. Volders PJ, Helsens K, Wang X, Menten B, Martens L, Gevaert K, Vandesompele J and Mestdagh P. LNCipedia: a database for annotated human lncRNA transcript sequences and structures. Nucleic Acids Res. 2013;41(Database issue):D246. PubMed PMID: 23042674; PMCID: PMC3531107.

24. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A and Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 2011;25(18):1915. PubMed PMID: 21890647; PMCID: PMC3185964.

25. Ravasi T, Suzuki H, Pang KC, Katayama S, Furuno M, Okunishi R, Fukuda S, Ru K, Frith MC, Gongora MM, Grimmond SM, Hume DA, Hayashizaki Y and Mattick JS. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. Genome Res. 2006;16(1):11. PubMed PMID: 16344565; PMCID: PMC1356124.

26. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH and Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. Science. 2007;316(5830):1488. Epub 2007/05/05. PubMed PMID: 17478681.

27. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiasdottir S, Jonsdottir T, Palsson S, Einarsdottir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A and Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science. 2007;316(5830):1491. Epub 2007/05/05. PubMed PMID: 17478679.

28. Cunnington MS, Santibanez Koref M, Mayosi BM, Burn J and Keavney B. Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression. PLoS Genet. 2010;6(4):e1000899. Epub 2010/04/14. PubMed PMID: 20386740; PMCID: 2851566.

29. Holdt LM, Hoffmann S, Sass K, Langenberger D, Scholz M, Krohn K, Finstermeier K, Stahringer A, Wilfert W, Beutner F, Gielen S, Schuler G, Gabel G, Bergert H, Bechmann I, Stadler PF, Thiery J and Teupser D. Alu elements in ANRIL non-coding RNA at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks. PLoS Genet. 2013;9(7):e1003588. PubMed PMID: 23861667; PMCID: PMC3701717.

30. Arab K, Park YJ, Lindroth AM, Schafer A, Oakes C, Weichenhan D, Lukanova A, Lundin E, Risch A, Meister M, Dienemann H, Dyckhoff G, Herold-Mende C, Grummt I, Niehrs C and Plass C. Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. Mol Cell. 2014;55(4):604. PubMed PMID: 25087872.

31. Eidem TM, Kugel JF and Goodrich JA. Noncoding RNAs: Regulators of the Mammalian Transcription Machinery. J Mol Biol. 2016;428(12):2652. PubMed PMID: 26920110; PMCID: PMC4894004.

32. Benabdallah NS and Bickmore WA. Regulatory Domains and Their Mechanisms. Cold Spring Harb Symp Quant Biol. 2015;80:45. PubMed PMID: 26590168.

33. Krivega I and Dean A. Enhancer and promoter interactions-long distance calls. Current opinion in genetics & development. 2012;22(2):79. PubMed PMID: 22169023; PMCID: PMC3342482.

34. Matharu N and Ahituv N. Minor Loops in Major Folds: Enhancer-Promoter Looping, Chromatin Restructuring, and Their Association with Transcriptional Regulation and Disease. PLoS Genet. 2015;11(12):e1005640. PubMed PMID: 26632825; PMCID: PMC4669122.

35. Lai F, Gardini A, Zhang A and Shiekhattar R. Integrator mediates the biogenesis of enhancer RNAs. Nature. 2015;525(7569):399. PubMed PMID: 26308897; PMCID: PMC4718573.

36. Tan JY, Smith AAT, Ferreira da Silva M, Matthey-Doret C, Rueedi R, Sonmez R, Ding D, Kutalik Z, Bergmann S and Marques AC. cis-Acting Complex-Trait-Associated lincRNA Expression Correlates with Modulation of Chromosomal Architecture. Cell Rep. 2017;18(9):2280. PubMed PMID: 28249171.

37. Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, Zhang S, Wang HB, Ge J, Lu X, Yang L and Chen LL. Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. Cell research. 2014;24(5):513. PubMed PMID: 24662484; PMCID: PMC4011346.

38. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA and Chang HY. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature. 2011;472(7341):120. PubMed PMID: 21423168; PMCID: PMC3670758.

39. Hacisuleyman E, Goff LA, Trapnell C, Williams A, Henao-Mejia J, Sun L, McClanahan P, Hendrickson DG, Sauvageau M, Kelley DR, Morse M, Engreitz J, Lander ES, Guttman M, Lodish HF, Flavell R, Raj A and Rinn JL. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. Nat Struct Mol Biol. 2014;21(2):198. PubMed PMID: 24463464; PMCID: PMC3950333.

40. Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ and Chang HY. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Methods. 2016;13(11):919. PubMed PMID: 27643841.

41. Mumbach MR, Satpathy AT, Boyle EA, Dai C, Gowen BG, Cho SW, Nguyen ML, Rubin AJ, Granja JM, Kazane KR, Wei Y, Nguyen T, Greenside PG, Corces MR, Tycko J, Simeonov DR, Suliman N, Li R, Xu J, Flynn RA, Kundaje A, Khavari PA, Marson A, Corn JE, Quertermous T, Greenleaf WJ and Chang HY. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. Nat Genet. 2017;49(11):1602. PubMed PMID: 28945252.

42. Wapinski O and Chang HY. Long noncoding RNAs and human disease. Trends Cell Biol. 2011;21(6):354. PubMed PMID: 21550244.

43. Moran I, Akerman I, van de Bunt M, Xie R, Benazra M, Nammo T, Arnes L, Nakic N, Garcia-Hurtado J, Rodriguez-Segui S, Pasquali L, Sauty-Colace C, Beucher A, Scharfmann R, van Arensbergen J, Johnson PR, Berry A, Lee C, Harkins T, Gmyr V, Pattou F, Kerr-Conte J, Piemonti L, Berney T, Hanley N, Gloyn AL, Sussel L, Langman L, Brayman KL, Sander M, McCarthy MI, Ravassard P and Ferrer J. Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. Cell metabolism. 2012;16(4):435. PubMed PMID: 23040067; PMCID: PMC3475176.

44. Miller CL, Pjanic M and Quertermous T. From Locus Association to Mechanism of Gene Causality: The Devil Is in the Details. Arterioscler Thromb Vasc Biol. 2015;35(10):2079. PubMed PMID: 26399919; PMCID: PMC4594207.

45. Miller CL, Assimes TL, Montgomery SB and Quertermous T. Dissecting the causal genetic mechanisms of coronary heart disease. Current atherosclerosis reports. 2014;16(5):406. PubMed PMID: 24623178; PMCID: 4015632.

46. Popadin K, Gutierrez-Arcelus M, Dermitzakis ET and Antonarakis SE. Genetic and Epigenetic Regulation of Human lincRNA Gene Expression. Am J Hum Genet. 2013;93(6):1015. Epub 2013/11/26. PubMed PMID: 24268656; PMCID: 3852921.

47. Consortium GT, Laboratory DA, Coordinating Center -Analysis Working G, Statistical Methods groups-Analysis Working G, Enhancing Gg, Fund NIHC, Nih/Nci, Nih/Nhgri, Nih/Nimh, Nih/Nida, Biospecimen Collection Source Site N, Biospecimen Collection Source Site R, Biospecimen Core Resource V, Brain Bank Repository-University of Miami Brain Endowment B, Leidos Biomedical-Project M, Study E, Genome Browser Data I, Visualization EBI, Genome Browser Data I, Visualization-Ucsc Genomics Institute UoCSC, Lead a, Laboratory DA, Coordinating C, management NIHp, Biospecimen c, Pathology, e QTLmwg, Battle A, Brown CD, Engelhardt BE and Montgomery SB. Genetic effects on gene expression across human tissues. Nature. 2017;550(7675):204. PubMed PMID: 29022597; PMCID: PMC5776756.

48. !!! INVALID CITATION !!! 5, 6, 48, 49.

49. Pennisi E. The Biology of Genomes. Disease risk links to gene regulation. Science. 2011;332(6033):1031. PubMed PMID: 21617055.

50. McPherson R and Tybjaerg-Hansen A. Genetics of Coronary Artery Disease. Circ Res. 2016;118(4):564. PubMed PMID: 26892958.

51. Congrains A, Kamide K, Oguro R, Yasuda O, Miyata K, Yamamoto E, Kawai T, Kusunoki H, Yamamoto H, Takeya Y, Yamamoto K, Onishi M, Sugimoto K, Katsuya T, Awata N, Ikebe K, Gondo Y, Oike Y, Ohishi M and Rakugi H. Genetic variants at the 9p21 locus contribute to atherosclerosis through modulation of ANRIL and CDKN2A/B. Atherosclerosis. 2012;220(2):449. PubMed PMID: 22178423.

52. Holdt LM, Beutner F, Scholz M, Gielen S, Gabel G, Bergert H, Schuler G, Thiery J and Teupser D. ANRIL expression is associated with atherosclerosis risk at chromosome 9p21. Arterioscler Thromb Vasc Biol. 2010;30(3):620. Epub 2010/01/09. PubMed PMID: 20056914.

53. Liu Y, Sanoff HK, Cho H, Burd CE, Torrice C, Mohlke KL, Ibrahim JG, Thomas NE and Sharpless NE. INK4/ARF transcript expression is associated with chromosome 9p21 variants linked to atherosclerosis. PLoS ONE. 2009;4(4):e5027. Epub 2009/04/04. PubMed PMID: 19343170; PMCID: 2660422.

54. Jarinova O, Stewart AF, Roberts R, Wells G, Lau P, Naing T, Buerki C, McLean BW, Cook RC, Parker JS and McPherson R. Functional Analysis of the Chromosome 9p21.3 Coronary Artery Disease Risk Locus. Arterioscler Thromb Vasc Biol. 2009;29:1671. Epub 2009/07/14. PubMed PMID: 19592466.

55. Broadbent HM, Peden JF, Lorkowski S, Goel A, Ongen H, Green F, Clarke R, Collins R, Franzosi MG, Tognoni G, Seedorf U, Rust S, Eriksson P, Hamsten A, Farrall M and Watkins H. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. Hum Mol Genet. 2008;17(6):806. Epub 2007/12/01. PubMed PMID: 18048406.

56. Kojima Y, Downing K, Kundu R, Miller C, Dewey F, Lancero H, Raaz U, Perisic L, Hedin U, Schadt E, Maegdefessel L, Quertermous T and Leeper NJ. Cyclin-dependent kinase inhibitor 2B regulates efferocytosis and atherosclerosis. J Clin Invest. 2014;124(3):1083. Epub 2014/02/18. PubMed PMID: 24531546; PMCID: 3938254.

57. Leeper NJ, Raiesdana A, Kojima Y, Kundu RK, Cheng H, Maegdefessel L, Toh R, Ahn GO, Ali ZA, Anderson DR, Miller CL, Roberts SC, Spin JM, de Almeida PE, Wu JC, Xu B, Cheng K, Quertermous M, Kundu S, Kortekaas KE, Berzin E, Downing KP, Dalman RL, Tsao PS, Schadt EE, Owens GK and Quertermous T. Loss of CDKN2B Promotes p53-Dependent Smooth Muscle Cell Apoptosis and Aneurysm Formation. Arterioscler Thromb Vasc Biol. 2013;33(1):e1. Epub 2012/11/20. PubMed PMID: 23162013.

58. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, Zeng L, Ntalla I, Lai FY, Hopewell JC, Giannakopoulou O, Jiang T, Hamby SE, Di Angelantonio E, Assimes TL, Bottinger EP, Chambers JC, Clarke R, Palmer CNA, Cubbon RM, Ellinor P, Ermel R, Evangelou E, Franks PW, Grace C, Gu D, Hingorani AD, Howson JMM, Ingelsson E, Kastrati A, Kessler T, Kyriakou T, Lehtimaki T, Lu X, Lu Y, Marz W, McPherson R, Metspalu A, Pujades-Rodriguez M, Ruusalepp A, Schadt EE, Schmidt AF, Sweeting MJ, Zalloua PA, AlGhalayini K, Keavney BD, Kooner JS, Loos RJF, Patel RS, Rutter MK, Tomaszewski M, Tzoulaki I, Zeggini E, Erdmann J, Dedoussis G, Bjorkegren JLM, Consortium E-C, CardioGramplusC4D, group UKBCCCw, Schunkert H, Farrall M, Danesh J, Samani NJ, Watkins H and Deloukas P. Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat Genet. 2017;49(9):1385. PubMed PMID: 28714975.

59. Russell MR, Penikis A, Oldridge DA, Alvarez-Dominguez JR, McDaniel L, Diamond M, Padovan O, Raman P, Li Y, Wei JS, Zhang S, Gnanchandran J, Seeger R, Asgharzadeh S, Khan J, Diskin SJ, Maris JM and Cole KA. CASC15-S Is a Tumor Suppressor lncRNA at the 6p22 Neuroblastoma Susceptibility Locus. Cancer Res. 2015;75(15):3155. PubMed PMID: 26100672; PMCID: PMC4526355.

60. Betts JA, Moradi Marjaneh M, Al-Ejeh F, Lim YC, Shi W, Sivakumaran H, Tropee R, Patch AM, Clark MB, Bartonicek N, Wiegmans AP, Hillman KM, Kaufmann S, Bain AL, Gloss BS, Crawford J, Kazakoff S, Wani S, Wen SW, Day B, Moller A, Cloonan N, Pearson J, Brown MA, Mercer TR, Waddell N, Khanna KK, Dray E, Dinger ME, Edwards SL and French JD. Long Noncoding RNAs CUPID1 and CUPID2 Mediate Breast Cancer Risk at 11q13 by Modulating the Response to DNA Damage. Am J Hum Genet. 2017;101(2):255. PubMed PMID: 28777932; PMCID: PMC5544418.

61. Guo H, Ahmed M, Zhang F, Yao CQ, Li S, Liang Y, Hua J, Soares F, Sun Y, Langstein J, Li Y, Poon C, Bailey SD, Desai K, Fei T, Li Q, Sendorek DH, Fraser M, Prensner JR, Pugh TJ, Pomerantz M, Bristow RG, Lupien M, Feng FY, Boutros PC, Freedman ML, Walsh MJ and He HH. Modulation of long noncoding RNAs by risk SNPs underlying genetic predispositions to prostate cancer. Nat Genet. 2016;48(10):1142. PubMed PMID: 27526323.

62. Luo Z, Rhie SK, Lay FD and Farnham PJ. A Prostate Cancer Risk Element Functions as a Repressive Loop that Regulates HOXA13. Cell Rep. 2017;21(6):1411. PubMed PMID: 29117547; PMCID: PMC5726543.

63. Mathy NW and Chen XM. Long non-coding RNAs (lncRNAs) and their transcriptional control of inflammatory responses. J Biol Chem. 2017;292(30):12375. PubMed PMID: 28615453; PMCID: PMC5535013.

64. Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, Rinn JL and Raj A. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol. 2015;16:20. PubMed PMID: 25630241; PMCID: PMC4369099.

65. Zuber V, Bettella F, Witoelar A, Consortium P, Cruk G, Consortium B, Consortium T, Andreassen OA, Mills IG and Urbanucci A. Bromodomain protein 4 discriminates tissue-specific super-enhancers containing disease-specific susceptibility loci in prostate and breast cancer. BMC genomics. 2017;18(1):270. PubMed PMID: 28359301; PMCID: PMC5374680.

66. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutyavin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, Bates D, Hansen RS, Neph S, Sabo PJ, Heimfeld S, Raubitschek A, Ziegler S, Cotsapas C, Sotoodehnia N, Glass I, Sunyaev SR, Kaul R and Stamatoyannopoulos JA. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012;337(6099):1190. Epub 2012/09/08. PubMed PMID: 22955828; PMCID: 3771521.

67. Lage K, Hansen NT, Karlberg EO, Eklund AC, Roque FS, Donahoe PK, Szallasi Z, Jensen TS and Brunak S. A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. Proc Natl Acad Sci U S A. 2008;105(52):20870. PubMed PMID: 19104045; PMCID: PMC2606902.

68. Bohmdorfer G and Wierzbicki AT. Control of Chromatin Structure by Long Noncoding RNA. Trends Cell Biol. 2015;25(10):623. PubMed PMID: 26410408; PMCID: PMC4584417.

69. Rinn JL and Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012;81:145. PubMed PMID: 22663078; PMCID: PMC3858397.

70. Clarke M and Bennett M. The emerging role of vascular smooth muscle cell apoptosis in atherosclerosis and plaque stability. American journal of nephrology. 2006;26(6):531. PubMed PMID: 17159340.

71. Falk E, Nakano M, Bentzon JF, Finn AV and Virmani R. Update on acute coronary syndromes: the pathologists' view. Eur Heart J. 2013;34(10):719. PubMed PMID: 23242196.

72. Gomez D and Owens GK. Smooth muscle cell phenotypic switching in atherosclerosis. Cardiovasc Res. 2012;95(2):156. Epub 2012/03/13. PubMed PMID: 22406749; PMCID: 3388816.

73. Sakakura K, Nakano M, Otsuka F, Ladich E, Kolodgie FD and Virmani R. Pathophysiology of atherosclerosis plaque progression. Heart, lung & circulation. 2013;22(6):399. PubMed PMID: 23541627.

74. Shah PK. Mechanisms of plaque vulnerability and rupture. J Am Coll Cardiol. 2003;41(4 Suppl S):15S. PubMed PMID: 12644336.

75. Shankman LS, Gomez D, Cherepanova OA, Salmon M, Alencar GF, Haskins RM, Swiatlowska P, Newman AA, Greene ES, Straub AC, Isakson B, Randolph GJ and Owens GK. KLF4-dependent phenotypic modulation of smooth muscle cells has a key role in atherosclerotic plaque pathogenesis. Nat Med. 2015;21(6):628. PubMed PMID: 25985364.

76. Milewicz DM, Kwartler CS, Papke CL, Regalado ES, Cao J and Reid AJ. Genetic variants promoting smooth muscle cell proliferation can result in diffuse and diverse vascular diseases: evidence for a hyperplastic vasculomyopathy. Genet Med. 2010;12(4):196. PubMed PMID: 20130469.

77. Guo DC, Papke CL, Tran-Fadulu V, Regalado ES, Avidan N, Johnson RJ, Kim DH, Pannu H, Willing MC, Sparks E, Pyeritz RE, Singh MN, Dalman RL, Grotta JC, Marian AJ, Boerwinkle EA, Frazier LQ, LeMaire SA, Coselli JS, Estrera AL, Safi HJ, Veeraraghavan S, Muzny DM, Wheeler DA, Willerson JT, Yu RK, Shete SS, Scherer SE, Raman CS, Buja LM and Milewicz DM. Mutations in smooth muscle alpha-actin (ACTA2) cause coronary artery disease, stroke, and Moyamoya disease, along with thoracic aortic disease. Am J Hum Genet. 2009;84(5):617. PubMed PMID: 19409525; PMCID: PMC2680995.

78. Mani A, Radhakrishnan J, Wang H, Mani A, Mani MA, Nelson-Williams C, Carew KS, Mane S, Najmabadi H, Wu D and Lifton RP. LRP6 mutation in a family with early coronary disease and metabolic risk factors. Science. 2007;315(5816):1278. PubMed PMID: 17332414; PMCID: PMC2945222.

79. Srivastava R, Zhang J, Go GW, Narayanan A, Nottoli TP and Mani A. Impaired LRP6-TCF7L2 Activity Enhances Smooth Muscle Cell Plasticity and Causes Coronary Artery Disease. Cell Rep. 2015;13(4):746. PubMed PMID: 26489464; PMCID: PMC4626307.

80. Howson JMM, Zhao W, Barnes DR, Ho WK, Young R, Paul DS, Waite LL, Freitag DF, Fauman EB, Salfati EL, Sun BB, Eicher JD, Johnson AD, Sheu WHH, Nielsen SF, Lin WY, Surendran P, Malarstig A, Wilk JB, Tybjaerg-Hansen A, Rasmussen KL, Kamstrup PR, Deloukas P, Erdmann J, Kathiresan S, Samani NJ, Schunkert H, Watkins H, CardioGramplusC4D, Do R, Rader DJ, Johnson JA, Hazen SL, Quyyumi AA, Spertus JA, Pepine CJ, Franceschini N, Justice A, Reiner AP, Buyske S, Hindorff LA, Carty CL, North KE, Kooperberg C, Boerwinkle E, Young K, Graff M, Peters U, Absher D, Hsiung CA, Lee WJ, Taylor KD, Chen YH, Lee IT, Guo X, Chung RH, Hung YJ, Rotter JI, Juang JJ, Quertermous T, Wang TD, Rasheed A, Frossard P, Alam DS, Majumder AAS, Di Angelantonio E, Chowdhury R, Epic CVD, Chen YI, Nordestgaard BG, Assimes TL, Danesh J, Butterworth AS and Saleheen D. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. Nat Genet. 2017;49(7):1113. PubMed PMID: 28530674; PMCID: PMC5555387.

81. Klarin D, Zhu QM, Emdin CA, Chaffin M, Horner S, McMillan BJ, Leed A, Weale ME, Spencer CCA, Aguet F, Segre AV, Ardlie KG, Khera AV, Kaushik VK, Natarajan P, Consortium CAD and Kathiresan S. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. Nat Genet. 2017;49(9):1392. PubMed PMID: 28714974; PMCID: PMC5577383.

82. van der Harst P and Verweij N. The Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circ Res. 2017. PubMed PMID: 29212778.

83. Nanda V and Miano JM. Leiomodin 1, a new serum response factor-dependent target gene expressed preferentially in differentiated smooth muscle cells. J Biol Chem. 2012;287(4):2459. PubMed PMID: 22157009; PMCID: PMC3268406.

84. Darrah R, McKone E, O'Connor C, Rodgers C, Genatossio A, McNamara S, Gibson R, Stuart Elborn J, Ennis M, Gallagher CG, Kalsheker N, Aitken M, Wiese D, Dunn J, Smith P, Pace R, Londono D, Goddard KA, Knowles MR and Drumm ML. EDNRA variants associate with smooth muscle mRNA levels, cell proliferation rates, and cystic fibrosis pulmonary disease severity. Physiol Genomics. 2010;41(1):71. PubMed PMID: 20028935; PMCID: PMC2841492.

85. Thomas JA, Deaton RA, Hastings NE, Shang Y, Moehle CW, Eriksson U, Topouzis S, Wamhoff BR, Blackman BR and Owens GK. PDGF-DD, a novel mediator of smooth muscle cell phenotypic modulation, is upregulated in endothelial cells exposed to atherosclerosis-prone flow patterns. Am J Physiol Heart Circ Physiol. 2009;296(2):H442. Epub 2008/11/26. PubMed PMID: 19028801; PMCID: 2643880.

86. Cheong A, Wood IC and Beech DJ. Less REST, more vascular disease? Regulation of cell cycle and migration of vascular smooth muscle cells. Cell Cycle. 2006;5(2):129. PubMed PMID: 16357530.

87. Miano JM. Vascular smooth muscle cell differentiation-2010. Journal of biomedical research. 2010;24(3):169. PubMed PMID: 23554628; PMCID: 3596552.

88. Chappell J, Harman JL, Narasimhan VM, Yu H, Foote K, Simons BD, Bennett MR and Jorgensen HF. Extensive Proliferation of a Subset of Differentiated, yet Plastic, Medial Vascular Smooth Muscle Cells Contributes to Neointimal Formation in Mouse Injury and Atherosclerosis Models. Circ Res. 2016;119(12):1313. PubMed PMID: 27682618; PMCID: PMC5149073.

89. Jacobsen K, Lund MB, Shim J, Gunnersen S, Fuchtbauer EM, Kjolby M, Carramolino L and Bentzon JF. Diverse cellular architecture of atherosclerotic plaque derives from clonal expansion of a few medial SMCs. JCI Insight. 2017;2(19). PubMed PMID: 28978793.

90. Feil S, Fehrenbacher B, Lukowski R, Essmann F, Schulze-Osthoff K, Schaller M and Feil R. Transdifferentiation of vascular smooth muscle cells to macrophage-like cells during atherogenesis. Circ Res. 2014;115(7):662. PubMed PMID: 25070003.

91. Iyer D, Naravane A, Wirks R, Nguyen T and Quertermous T. SMAD3 opposes diifferentiation and opposes phenotypic modulation in coronary artery smooth muscle cells. Submitted. 2018.

92. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR, Anttila V, Xu H, Zang C, Farh K, Ripke S, Day FR, ReproGen C, Schizophrenia Working Group of the Psychiatric Genomics C, Consortium R, Purcell S, Stahl E, Lindstrom S, Perry JR, Okada Y, Raychaudhuri S, Daly MJ, Patterson N, Neale BM and Price AL. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet. 2015;47(11):1228. PubMed PMID: 26414678; PMCID: PMC4626285.

93. Liu B, Iyer D, Wang T, Castano VG, Nguyen T, Montgomery S, Quertermous T and Miller C. Rgulatory variation identification in primary human coronary artery smooth muscle cells provides localization of coronary artery disease causal variation Submitted. 2018.

94. Kumasaka N, Knights AJ and Gaffney DJ. Fine-mapping cellular QTLs with RASQUAL and ATAC-seq. Nat Genet. 2015. PubMed PMID: 26656845.

95. Hormozdiari F, van de Bunt M, Segre AV, Li X, Joo JWJ, Bilow M, Sul JH, Sankararaman S, Pasaniuc B and Eskin E. Colocalization of GWAS and eQTL Signals Detects Target Genes. Am J Hum Genet. 2016;99(6):1245. PubMed PMID: 27866706; PMCID: PMC5142122.

96. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM and Yang J. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481. PubMed PMID: 27019110.

97. Buenrostro JD, Giresi PG, Zaba LC, Chang HY and Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10(12):1213. PubMed PMID: 24097267; PMCID: PMC3959825.

98. Kim BJ, Pjanic M, Nguyen T, Miller CL, Liu B, Wang T, Sazonova O, Carcamo-Orive I, Perisic L, Maegdefessel L, Hedin U and Quertermous T. TCF21 and the aryl-hydrocarbon receptor cooperate to activate a pro-atherosclerotic gene expression program. PLoS Genet. 2017;13.(5):1006750. PubMed PMID: 28481916; PMCID: PMC5439967

99. Tehranchi AK, Myrthil M, Martin T, Hie BL, Golan D and Fraser HB. Pooled ChIP-Seq Links Variation in Transcription Factor Binding to Complex Disease Risk. Cell. 2016;165(3):730. PubMed PMID: 27087447; PMCID: PMC4842172.

100. Zhang H, Xue C, Wang Y, Shi J, Zhang X, Li W, Nunez S, Foulkes AS, Lin J, Hinkle CC, Yang W, Morrisey EE, Rader DJ, Li M and Reilly MP. Deep RNA Sequencing Uncovers a Repertoire of Human Macrophage Long Intergenic Noncoding RNAs Modulated by Macrophage Activation and Associated With Cardiometabolic Diseases. J Am Heart Assoc. 2017;6(11). PubMed PMID: 29133519; PMCID: PMC5721798.

101. Ballantyne RL, Zhang X, Nunez S, Xue C, Zhao W, Reed E, Salaheen D, Foulkes AS, Li M and Reilly MP. Genome-wide interrogation reveals hundreds of long intergenic noncoding RNAs that associate with cardiometabolic traits. Hum Mol Genet. 2016;25(14):3125. PubMed PMID: 27288454; PMCID: PMC5181595.

102. Xue C, Zhang X, Zhang H, Ferguson JF, Wang Y, Hinkle CC, Li M and Reilly MP. De novo RNA sequence assembly during in vivo inflammatory stress reveals hundreds of unannotated lincRNAs in human blood CD14(+) monocytes and in adipose tissue. Physiol Genomics. 2017;49(6):287. PubMed PMID: 28389524; PMCID: PMC5495909.

103. Ounzain S, Burdet F, Ibberson M and Pedrazzini T. Discovery and functional characterization of cardiovascular long noncoding RNAs. J Mol Cell Cardiol. 2015;89(Pt A):17. PubMed PMID: 26408097.

104. Zhao J, Zhang W, Lin M, Wu W, Jiang P, Tou E, Xue M, Richards A, Jourd'heuil D, Asif A, Zheng D, Singer HA, Miano JM and Long X. MYOSLID Is a Novel Serum Response Factor-Dependent Long Noncoding RNA That Amplifies the Vascular Smooth Muscle Differentiation Program. Arterioscler Thromb Vasc Biol. 2016;36(10):2088. PubMed PMID: 27444199; PMCID: PMC5033703.

105. Dechamethakun S and Muramatsu M. Long noncoding RNA variations in cardiometabolic diseases. J Hum Genet. 2017;62(1):97. PubMed PMID: 27305986.

106. Boon RA, Jae N, Holdt L and Dimmeler S. Long Noncoding RNAs: From Clinical Genetics to Therapeutic Targets? J Am Coll Cardiol. 2016;67(10):1214. PubMed PMID: 26965544.

107. Bennett MR, Sinha S and Owens GK. Vascular Smooth Muscle Cells in Atherosclerosis. Circ Res. 2016;118(4):692. PubMed PMID: 26892967; PMCID: PMC4762053.

108. Sun L, Zhang Z, Bailey TL, Perkins AC, Tallack MR, Xu Z and Liu H. Prediction of novel long non-coding RNAs based on RNA-Seq data of mouse Klf1 knockout study. BMC Bioinformatics. 2012;13:331. PubMed PMID: 23237380; PMCID: PMC3577497.

109. Tsoi LC, Iyer MK, Stuart PE, Swindell WR, Gudjonsson JE, Tejasvi T, Sarkar MK, Li B, Ding J, Voorhees JJ, Kang HM, Nair RP, Chinnaiyan AM, Abecasis GR and Elder JT. Analysis of long non-coding RNAs highlights tissue-specific expression patterns and epigenetic profiles in normal and psoriatic skin. Genome Biol. 2015;16:24. PubMed PMID: 25723451; PMCID: PMC4311508.

110. Ushakov K, Koffler-Brill T, Rom A, Perl K, Ulitsky I and Avraham KB. Genome-wide identification and expression profiling of long non-coding RNAs in auditory and vestibular systems. Scientific reports. 2017;7(1):8637. PubMed PMID: 28819115; PMCID: PMC5561060.

111. Lin MF, Jungreis I and Kellis M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics. 2011;27(13):i275. PubMed PMID: 21685081; PMCID: PMC3117341.

112. Sun K, Chen X, Jiang P, Song X, Wang H and Sun H. iSeeRNA: identification of long intergenic non-coding RNA transcripts from transcriptome sequencing data. BMC genomics. 2013;14 Suppl 2:S7. PubMed PMID: 23445546; PMCID: PMC3582448.

113. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Do R, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkila K, Hypponen E, Isaacs A, Jackson AU, Johansson A, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikainen LP, Magnusson PKE, Mangino M, Mihailov E, Montasser ME, Muller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney ASF, Doring A, Elliott P, Epstein SE, Ingi Eyjolfsson G, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJP, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimaki T, Lin SY, Lindstrom J, Loos RJF, Mach F, McArdle WL, Meisinger C, Mitchell BD, Muller G, Nagaraja R, Narisu N, Nieminen TVM, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruokonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stancakova A, Stirrups K, Swift AJ, Tiret L, Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YI, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrieres J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllensten U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Jarvelin MR, Jula A, Kahonen M, Kaprio J, Kesaniemi A, Kivimaki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, Marz W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njolstad I, Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PEH, Sheu WH, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BHR, Ordovas JM, Boerwinkle E, Palmer CNA, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Kathiresan S, Mohlke KL, Ingelsson E, Abecasis GR and Global Lipids Genetics C. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013;45(11):1274. PubMed PMID: 24097068; PMCID: PMC3838666.

114. Liu DJ, Peloso GM, Yu H, Butterworth AS, Wang X, Mahajan A, Saleheen D, Emdin C, Alam D, Alves AC, Amouyel P, Di Angelantonio E, Arveiler D, Assimes TL, Auer PL, Baber U, Ballantyne CM, Bang LE, Benn M, Bis JC, Boehnke M, Boerwinkle E, Bork-Jensen J, Bottinger EP, Brandslund I, Brown M, Busonero F, Caulfield MJ, Chambers JC, Chasman DI, Chen YE, Chen YI, Chowdhury R, Christensen C, Chu AY, Connell JM, Cucca F, Cupples LA, Damrauer SM, Davies G, Deary IJ, Dedoussis G, Denny JC, Dominiczak A, Dube MP, Ebeling T, Eiriksdottir G, Esko T, Farmaki AE, Feitosa MF, Ferrario M, Ferrieres J, Ford I, Fornage M, Franks PW, Frayling TM, Frikke-Schmidt R, Fritsche LG, Frossard P, Fuster V, Ganesh SK, Gao W, Garcia ME, Gieger C, Giulianini F, Goodarzi MO, Grallert H, Grarup N, Groop L, Grove ML, Gudnason V, Hansen T, Harris TB, Hayward C, Hirschhorn JN, Holmen OL, Huffman J, Huo Y, Hveem K, Jabeen S, Jackson AU, Jakobsdottir J, Jarvelin MR, Jensen GB, Jorgensen ME, Jukema JW, Justesen JM, Kamstrup PR, Kanoni S, Karpe F, Kee F, Khera AV, Klarin D, Koistinen HA, Kooner JS, Kooperberg C, Kuulasmaa K, Kuusisto J, Laakso M, Lakka T, Langenberg C, Langsted A, Launer LJ, Lauritzen T, Liewald DCM, Lin LA, Linneberg A, Loos RJF, Lu Y, Lu X, Magi R, Malarstig A, Manichaikul A, Manning AK, Mantyselka P, Marouli E, Masca NGD, Maschio A, Meigs JB, Melander O, Metspalu A, Morris AP, Morrison AC, Mulas A, Muller-Nurasyid M, Munroe PB, Neville MJ, Nielsen JB, Nielsen SF, Nordestgaard BG, Ordovas JM, Mehran R, O'Donnell CJ, Orho-Melander M, Molony CM, Muntendam P, Padmanabhan S, Palmer CNA, Pasko D, Patel AP, Pedersen O, Perola M, Peters A, Pisinger C, Pistis G, Polasek O, Poulter N, Psaty BM, Rader DJ, Rasheed A, Rauramaa R, Reilly DF, Reiner AP, Renstrom F, Rich SS, Ridker PM, Rioux JD, Robertson NR, Roden DM, Rotter JI, Rudan I, Salomaa V, Samani NJ, Sanna S, Sattar N, Schmidt EM, Scott RA, Sever P, Sevilla RS, Shaffer CM, Sim X, Sivapalaratnam S, Small KS, Smith AV, Smith BH, Somayajula S, Southam L, Spector TD, Speliotes EK, Starr JM, Stirrups KE, Stitziel N, Strauch K, Stringham HM, Surendran P, Tada H, Tall AR, Tang H, Tardif JC, Taylor KD, Trompet S, Tsao PS, Tuomilehto J, Tybjaerg-Hansen A, van Zuydam NR, Varbo A, Varga TV, Virtamo J, Waldenberger M, Wang N, Wareham NJ, Warren HR, Weeke PE, Weinstock J, Wessel J, Wilson JG, Wilson PWF, Xu M, Yaghootkar H, Young R, Zeggini E, Zhang H, Zheng NS, Zhang W, Zhang Y, Zhou W, Zhou Y, Zoledziewska M, Charge Diabetes Working G, Consortium EP-I, Consortium E-C, Consortium G, Program VAMV, Howson JMM, Danesh J, McCarthy MI, Cowan CA, Abecasis G, Deloukas P, Musunuru K, Willer CJ and Kathiresan S. Exome-wide association study of plasma lipids in >300,000 individuals. Nat Genet. 2017;49(12):1758. PubMed PMID: 29083408; PMCID: PMC5709146.

115. Langfelder P and Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. PubMed PMID: 19114008; PMCID: PMC2631488.

116. Song WM and Zhang B. Multiscale Embedded Gene Co-expression Network Analysis. PLoS computational biology. 2015;11(11):e1004574. PubMed PMID: 26618778; PMCID: PMC4664553.

117. Chella Krishnan K, Kurt Z, Barrere-Cain R, Sabir S, Das A, Floyd R, Vergnes L, Zhao Y, Che N, Charugundla S, Qi H, Zhou Z, Meng Y, Pan C, Seldin MM, Norheim F, Hui S, Reue K, Lusis AJ and Yang X. Integration of Multi-omics Data from Mouse Diversity Panel Highlights Mitochondrial Dysfunction in Non-alcoholic Fatty Liver Disease. Cell Syst. 2018;6(1):103. PubMed PMID: 29361464; PMCID: PMC5799036.

118. Ivey ME, Osman N and Little PJ. Endothelin-1 signalling in vascular smooth muscle: pathways controlling cellular functions associated with atherosclerosis. Atherosclerosis. 2008;199(2):237. PubMed PMID: 18436225.

119. Almlof JC, Lundmark P, Lundmark A, Ge B, Pastinen T, Cardiogenics C, Goodall AH, Cambien F, Deloukas P, Ouwehand WH and Syvanen AC. Single nucleotide polymorphisms with cis-regulatory effects on long non-coding transcripts in human primary monocytes. PLoS One. 2014;9(7):e102612. PubMed PMID: 25025429; PMCID: PMC4099216.

120. Kumar V, Westra HJ, Karjalainen J, Zhernakova DV, Esko T, Hrdlickova B, Almeida R, Zhernakova A, Reinmaa E, Vosa U, Hofker MH, Fehrmann RS, Fu J, Withoff S, Metspalu A, Franke L and Wijmenga C. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. PLoS Genet. 2013;9(1):e1003201. Epub 2013/01/24. PubMed PMID: 23341781; PMCID: 3547830.

121. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES and Rinn JL. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009;106(28):11667. PubMed PMID: 19571010; PMCID: PMC2704857.

122. Zhang R, Li X, Ramaswami G, Smith KS, Turecki G, Montgomery GW and Li JB. Quantifying RNA allelic ratios by microfluidics-based multiplex PCR and sequencing. Nature Methods. 2013;11(1):51.

123. Project G. Enhancing GTEx by bridging the gaps between genotype, gene expression, and disease. Nat Genet. 2017;49(12):1664. PubMed PMID: 29019975.

124. Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES and Aiden EL. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst. 2016;3(1):99. PubMed PMID: 27467250; PMCID: PMC5596920.

125. Ay F, Bailey TL and Noble WS. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. Genome Res. 2014;24(6):999. PubMed PMID: 24501021; PMCID: PMC4032863.

126. Feenstra B, Geller F, Krogh C, Hollegaard MV, Gortz S, Boyd HA, Murray JC, Hougaard DM and Melbye M. Common variants near MBNL1 and NKX2-5 are associated with infantile hypertrophic pyloric stenosis. Nat Genet. 2012;44(3):334. PubMed PMID: 22306654; PMCID: PMC3693399.

127. Liang J, Le TH, Edwards DRV, Tayo BO, Gaulton KJ, Smith JA, Lu Y, Jensen RA, Chen G, Yanek LR, Schwander K, Tajuddin SM, Sofer T, Kim W, Kayima J, McKenzie CA, Fox E, Nalls MA, Young JH, Sun YV, Lane JM, Cechova S, Zhou J, Tang H, Fornage M, Musani SK, Wang H, Lee J, Adeyemo A, Dreisbach AW, Forrester T, Chu PL, Cappola A, Evans MK, Morrison AC, Martin LW, Wiggins KL, Hui Q, Zhao W, Jackson RD, Ware EB, Faul JD, Reiner AP, Bray M, Denny JC, Mosley TH, Palmas W, Guo X, Papanicolaou GJ, Penman AD, Polak JF, Rice K, Taylor KD, Boerwinkle E, Bottinger EP, Liu K, Risch N, Hunt SC, Kooperberg C, Zonderman AB, Laurie CC, Becker DM, Cai J, Loos RJF, Psaty BM, Weir DR, Kardia SLR, Arnett DK, Won S, Edwards TL, Redline S, Cooper RS, Rao DC, Rotter JI, Rotimi C, Levy D, Chakravarti A, Zhu X and Franceschini N. Single-trait and multi-trait genome-wide association analyses identify novel loci for blood pressure in African-ancestry populations. PLoS Genet. 2017;13(5):e1006728. PubMed PMID: 28498854; PMCID: PMC5446189.

128. Fox CS, Liu Y, White CC, Feitosa M, Smith AV, Heard-Costa N, Lohman K, Consortium G, Consortium M, Consortium G, Johnson AD, Foster MC, Greenawalt DM, Griffin P, Ding J, Newman AB, Tylavsky F, Miljkovic I, Kritchevsky SB, Launer L, Garcia M, Eiriksdottir G, Carr JJ, Gudnason V, Harris TB, Cupples LA and Borecki IB. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. PLoS Genet. 2012;8(5):e1002695. PubMed PMID: 22589738; PMCID: PMC3349734 competing interests exist.

129. Gorski M, van der Most PJ, Teumer A, Chu AY, Li M, Mijatovic V, Nolte IM, Cocca M, Taliun D, Gomez F, Li Y, Tayo B, Tin A, Feitosa MF, Aspelund T, Attia J, Biffar R, Bochud M, Boerwinkle E, Borecki I, Bottinger EP, Chen MH, Chouraki V, Ciullo M, Coresh J, Cornelis MC, Curhan GC, d'Adamo AP, Dehghan A, Dengler L, Ding J, Eiriksdottir G, Endlich K, Enroth S, Esko T, Franco OH, Gasparini P, Gieger C, Girotto G, Gottesman O, Gudnason V, Gyllensten U, Hancock SJ, Harris TB, Helmer C, Hollerer S, Hofer E, Hofman A, Holliday EG, Homuth G, Hu FB, Huth C, Hutri-Kahonen N, Hwang SJ, Imboden M, Johansson A, Kahonen M, Konig W, Kramer H, Kramer BK, Kumar A, Kutalik Z, Lambert JC, Launer LJ, Lehtimaki T, de Borst M, Navis G, Swertz M, Liu Y, Lohman K, Loos RJF, Lu Y, Lyytikainen LP, McEvoy MA, Meisinger C, Meitinger T, Metspalu A, Metzger M, Mihailov E, Mitchell P, Nauck M, Oldehinkel AJ, Olden M, Wjh Penninx B, Pistis G, Pramstaller PP, Probst-Hensch N, Raitakari OT, Rettig R, Ridker PM, Rivadeneira F, Robino A, Rosas SE, Ruderfer D, Ruggiero D, Saba Y, Sala C, Schmidt H, Schmidt R, Scott RJ, Sedaghat S, Smith AV, Sorice R, Stengel B, Stracke S, Strauch K, Toniolo D, Uitterlinden AG, Ulivi S, Viikari JS, Volker U, Vollenweider P, Volzke H, Vuckovic D, Waldenberger M, Jin Wang J, Yang Q, Chasman DI, Tromp G, Snieder H, Heid IM, Fox CS, Kottgen A, Pattaro C, Boger CA and Fuchsberger C. 1000 Genomes-based meta-analysis identifies 10 novel loci for kidney function. Scientific reports. 2017;7:45040. PubMed PMID: 28452372; PMCID: PMC5408227.

130. Kemp JP, Morris JA, Medina-Gomez C, Forgetta V, Warrington NM, Youlten SE, Zheng J, Gregson CL, Grundberg E, Trajanoska K, Logan JG, Pollard AS, Sparkes PC, Ghirardello EJ, Allen R, Leitch VD, Butterfield NC, Komla-Ebri D, Adoum AT, Curry KF, White JK, Kussy F, Greenlaw KM, Xu C, Harvey NC, Cooper C, Adams DJ, Greenwood CMT, Maurano MT, Kaptoge S, Rivadeneira F, Tobias JH, Croucher PI, Ackert-Bicknell CL, Bassett JHD, Williams GR, Richards JB and Evans DM. Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. Nat Genet. 2017;49(10):1468. PubMed PMID: 28869591; PMCID: PMC5621629.

131. Lutz SM, Cho MH, Young K, Hersh CP, Castaldi PJ, McDonald ML, Regan E, Mattheisen M, DeMeo DL, Parker M, Foreman M, Make BJ, Jensen RL, Casaburi R, Lomas DA, Bhatt SP, Bakke P, Gulsvik A, Crapo JD, Beaty TH, Laird NM, Lange C, Hokanson JE, Silverman EK, Investigators E and Investigators CO. A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. BMC Genet. 2015;16:138. PubMed PMID: 26634245; PMCID: PMC4668640.

132. Hidai H, Bardales R, Goodwin R, Quertermous T and Quertermous EE. Cloning of capsulin, a basic helix-loop-helix factor expressed in progenitor cells of the pericardium and the coronary arteries. Mech Dev. 1998;73(1):33. Epub 1998/05/28. PubMed PMID: 9545526.

133. Wei QX, Claus R, Hielscher T, Mertens D, Raval A, Oakes CC, Tanner SM, de la Chapelle A, Byrd JC, Stilgenbauer S and Plass C. Germline allele-specific expression of DAPK1 in chronic lymphocytic leukemia. PLoS One. 2013;8(1):e55261. Epub 2013/02/06. PubMed PMID: 23383130; PMCID: 3557246.

134. Valle L, Serena-Acedo T, Liyanarachchi S, Hampel H, Comeras I, Li Z, Zeng Q, Zhang HT, Pennison MJ, Sadim M, Pasche B, Tanner SM and de la Chapelle A. Germline allele-specific expression of TGFBR1 confers an increased risk of colorectal cancer. Science. 2008;321(5894):1361. Epub 2008/08/16. PubMed PMID: 18703712; PMCID: 2672914.

135. Tomsic J, Guda K, Liyanarachchi S, Hampel H, Natale L, Markowitz SD, Tanner SM and de la Chapelle A. Allele-specific expression of TGFBR1 in colon cancer patients. Carcinogenesis. 2010;31(10):1800. Epub 2010/08/14. PubMed PMID: 20705955; PMCID: 2950937.

136. Rohlin A, Engwall Y, Fritzell K, Goransson K, Bergsten A, Einbeigi Z, Nilbert M, Karlsson P, Bjork J and Nordling M. Inactivation of promoter 1B of APC causes partial gene silencing: evidence for a significant role of the promoter in regulation and causative of familial adenomatous polyposis. Oncogene. 2011;30(50):4977. Epub 2011/06/07. PubMed PMID: 21643010; PMCID: 3240859.

137. Ria M, Lagercrantz J, Samnegard A, Boquist S, Hamsten A and Eriksson P. A common polymorphism in the promoter region of the TNFSF4 gene is associated with lower allele-specific expression and risk of myocardial infarction. PLoS One. 2011;6(3):e17652. Epub 2011/03/30. PubMed PMID: 21445270; PMCID: 3060868.

138. Moore AE, Young LE and Dixon DA. A common single-nucleotide polymorphism in cyclooxygenase-2 disrupts microRNA-mediated regulation. Oncogene. 2012;31(12):1592. Epub 2011/08/09. PubMed PMID: 21822307; PMCID: 3454533.

139. Milani L, Lundmark A, Nordlund J, Kiialainen A, Flaegstad T, Jonmundsson G, Kanerva J, Schmiegelow K, Gunderson KL, Lonnerholm G and Syvanen AC. Allele-specific gene expression patterns in primary leukemic cells reveal regulation of gene expression by CpG site methylation. Genome Res. 2009;19(1):1. Epub 2008/11/11. PubMed PMID: 18997001; PMCID: 2612957.

140. Matera I, Musso M, Griseri P, Rusmini M, Di Duca M, So MT, Mavilio D, Miao X, Tam PH, Ravazzolo R, Ceccherini I and Garcia-Barcelo M. Allele-specific expression at the RET locus in blood and gut tissue of individuals carrying risk alleles for Hirschsprung disease. Human mutation. 2013;34(5):754. Epub 2013/02/27. PubMed PMID: 23441071.

141. Lefebvre JF, Vello E, Ge B, Montgomery SB, Dermitzakis ET, Pastinen T and Labuda D. Genotype-based test in mapping cis-regulatory variants from allele-specific expression data. PLoS One. 2012;7(6):e38667. Epub 2012/06/12. PubMed PMID: 22685595; PMCID: 3369843.

142. Gao C, Devarajan K, Zhou Y, Slater CM, Daly MB and Chen X. Identifying breast cancer risk loci by global differential allele-specific expression (DASE) analysis in mammary epithelial transcriptome. BMC genomics. 2012;13:570. Epub 2012/10/31. PubMed PMID: 23107584; PMCID: 3532379.

143. Castellsague E, Gonzalez S, Guino E, Stevens KN, Borras E, Raymond VM, Lazaro C, Blanco I, Gruber SB and Capella G. Allele-specific expression of APC in adenomatous polyposis families. Gastroenterology. 2010;139(2):439. Epub 2010/05/04. PubMed PMID: 20434453; PMCID: 2910837.

144. Campino S, Forton J, Raj S, Mohr B, Auburn S, Fry A, Mangano VD, Vandiedonck C, Richardson A, Rockett K, Clark TG and Kwiatkowski DP. Validating discovered Cis-acting regulatory genetic variants: application of an allele specific expression approach to HapMap populations. PLoS One. 2008;3(12):e4105. Epub 2009/01/01. PubMed PMID: 19116668; PMCID: 2605564.

145. Li Y, Grupe A, Rowland C, Nowotny P, Kauwe JS, Smemo S, Hinrichs A, Tacey K, Toombs TA, Kwok S, Catanese J, White TJ, Maxwell TJ, Hollingworth P, Abraham R, Rubinsztein DC, Brayne C, Wavrant-De Vrieze F, Hardy J, O'Donovan M, Lovestone S, Morris JC, Thal LJ, Owen M, Williams J and Goate A. DAPK1 variants are associated with Alzheimer's disease and allele-specific expression. Hum Mol Genet. 2006;15(17):2560. Epub 2006/07/19. PubMed PMID: 16847012.

146. Serre D, Gurd S, Ge B, Sladek R, Sinnett D, Harmsen E, Bibikova M, Chudin E, Barker DL, Dickinson T, Fan JB and Hudson TJ. Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic cis-acting mechanisms regulating gene expression. PLoS Genet. 2008;4(2):e1000006. Epub 2008/05/06. PubMed PMID: 18454203; PMCID: 2265535.

147. Conde L, Bracci PM, Richardson R, Montgomery SB and Skibola CF. Integrating GWAS and Expression Data for Functional Characterization of Disease-Associated SNPs: An Application to Follicular Lymphoma. Am J Hum Genet. 2013;92(1):126. Epub 2012/12/19. PubMed PMID: 23246294.

148. Montgomery SB, Lappalainen T, Gutierrez-Arcelus M and Dermitzakis ET. Rare and common regulatory variation in population-scale sequenced human genomes. PLoS Genet. 2011;7(7):e1002144. Epub 2011/08/04. PubMed PMID: 21811411; PMCID: 3141000.

149. Miyaoka Y, Chan AH, Judge LM, Yoo J, Huang M, Nguyen TD, Lizarraga PP, So PL and Conklin BR. Isolation of single-base genome-edited human iPS cells without antibiotic selection. Nat Methods. 2014;11(3):291. PubMed PMID: 24509632; PMCID: PMC4063274.

150. Liu H, Wei Z, Dominguez A, Li Y, Wang X and Qi LS. CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. Bioinformatics. 2015;31(22):3676. PubMed PMID: 26209430; PMCID: PMC4757951.

151. Xia W, Zhu XW, Mo XB, Wu LF, Wu J, Guo YF, Zeng KQ, Wang MJ, Lin X, Qiu YH, Wang L, He P, Xie FF, Bing PF, Lu X, Liu YZ, Yi NJ, Deng FY and Lei SF. Integrative multi-omics analysis revealed SNP-lncRNA-mRNA (SLM) networks in human peripheral blood mononuclear cells. Hum Genet. 2017;136(4):451. PubMed PMID: 28243742.

152. Wen X. Molecular QTL discovery incorporating genomic annotations using Bayesian false discovery rate control. Ann of Appl Stat. 2016;10:1619.

153. Wang SS, Schadt EE, Wang H, Wang X, Ingram-Drake L, Shi W, Drake TA and Lusis AJ. Identification of pathways for atherosclerosis in mice: integration of quantitative trait locus analysis and global gene expression data. Circ Res. 2007;101(3):e11. Epub 2007/07/21. PubMed PMID: 17641228.

154. Califano A, Butte AJ, Friend S, Ideker T and Schadt E. Leveraging models of cell regulation and GWAS data in integrative network-based association studies. Nat Genet. 2012;44(8):841. Epub 2012/07/28. PubMed PMID: 22836096; PMCID: 3593099.

155. Schadt EE, Friend SH and Shaywitz DA. A network view of disease and compound screening. Nature reviews Drug discovery. 2009;8(4):286. Epub 2009/04/02. PubMed PMID: 19337271.

156. Schadt EE. Molecular networks as sensors and drivers of common human diseases. Nature. 2009;461(7261):218. Epub 2009/09/11. PubMed PMID: 19741703.

157. Shu L, Chan KHK, Zhang G, Huan T, Kurt Z, Zhao Y, Codoni V, Tregouet DA, Cardiogenics C, Yang J, Wilson JG, Luo X, Levy D, Lusis AJ, Liu S and Yang X. Shared genetic regulatory networks for cardiovascular disease and type 2 diabetes in multiple populations of diverse ethnicities in the United States. PLoS Genet. 2017;13(9):e1007040. PubMed PMID: 28957322; PMCID: PMC5634657.

158. Zhao Y, Chen J, Freudenberg JM, Meng Q, Rajpal DK and Yang X. Network-Based Identification and Prioritization of Key Regulators of Coronary Artery Disease Loci. Arterioscler Thromb Vasc Biol. 2016;36(5):928. PubMed PMID: 26966275; PMCID: PMC5576868.

159. Ghosh S, Vivar J, Nelson CP, Willenborg C, Segre AV, Makinen VP, Nikpay M, Erdmann J, Blankenberg S, O'Donnell C, Marz W, Laaksonen R, Stewart AF, Epstein SE, Shah SH, Granger CB, Hazen SL, Kathiresan S, Reilly MP, Yang X, Quertermous T, Samani NJ, Schunkert H, Assimes TL and McPherson R. Systems Genetics Analysis of Genome-Wide Association Study Reveals Novel Associations Between Key Biological Processes and Coronary Artery Disease. Arterioscler Thromb Vasc Biol. 2015;35(7):1712. PubMed PMID: 25977570; PMCID: PMC4841833.

160. Makinen VP, Civelek M, Meng Q, Zhang B, Zhu J, Levian C, Huan T, Segre AV, Ghosh S, Vivar J, Nikpay M, Stewart AF, Nelson CP, Willenborg C, Erdmann J, Blakenberg S, O'Donnell CJ, Marz W, Laaksonen R, Epstein SE, Kathiresan S, Shah SH, Hazen SL, Reilly MP, Coronary ADG-WR, Meta-Analysis C, Lusis AJ, Samani NJ, Schunkert H, Quertermous T, McPherson R, Yang X and Assimes TL. Integrative genomics reveals novel molecular pathways and gene networks for coronary artery disease. PLoS Genet. 2014;10(7):e1004502. PubMed PMID: 25033284; PMCID: PMC4102418.

161. Li G, Xu A, Sim S, Priest JR, Tian X, Khan T, Quertermous T, Zhou B, Tsao PS, Quake SR and Wu SM. Transcriptomic Profiling Maps Anatomically Patterned Subpopulations among Single Embryonic Cardiac Cells. Developmental cell. 2016;39(4):491. PubMed PMID: 27840109; PMCID: PMC5130110.

162. Wirka R, Pjanic M and Quertermous T. Advances in Transcriptomics: Investigating cardiovascular disease at unprecedented resolution. Circulation research. 2018; in press.

163. Paik DT, Tian L, J. L, Sayed N, Rhee S-R, Wirka R, Red-Horse K and Quertermous T. Large-Scale Single-Cell RNA-Seq Reveals Molecular Signatures of Heterogeneous Populations of Human Induced Pluripotent Stem Cell-Derived Endothelial Cells. Circ Res. 2018;In review. .

164. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP and Lim WA. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013;152(5):1173. PubMed PMID: 23452860; PMCID: PMC3664290.

165. Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, Attenello FJ, Villalta JE, Cho MY, Chen Y, Mandegar MA, Olvera MP, Gilbert LA, Conklin BR, Chang HY, Weissman JS and Lim DA. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science. 2017;355(6320). PubMed PMID: 27980086; PMCID: PMC5394926.

166. Zhu J, Zhang B, Smith EN, Drees B, Brem RB, Kruglyak L, Bumgarner RE and Schadt EE. Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. Nat Genet. 2008;40(7):854. PubMed PMID: 18552845; PMCID: PMC2573859.

167. Deng DX, Spin JM, Tsalenko A, Vailaya A, Ben-Dor A, Yakhini Z, Tsao P, Bruhn L and Quertermous T. Molecular signatures determining coronary artery and saphenous vein smooth muscle cell phenotypes: distinct responses to stimuli. Arterioscler Thromb Vasc Biol. 2006;26(5):1058. Epub 2006/02/04. PubMed PMID: 16456091.

168. Ishida T, Kundu RK, Yang E, Hirata K, Ho YD and Quertermous T. Targeted disruption of endothelial cell-selective adhesion molecule inhibits angiogenic processes in vitro and in vivo. J Biol Chem. 2003;278(36):34598. Epub 2003/06/24. PubMed PMID: 12819200.

169. Kojima Y, Kundu RK, Cox CM, Leeper NJ, Anderson JA, Chun HJ, Ali ZA, Ashley EA, Krieg PA and Quertermous T. Upregulation of the apelin-APJ pathway promotes neointima formation in the carotid ligation model in mouse. Cardiovasc Res. 2010;87(1):156. Epub 2010/02/24. PubMed PMID: 20176814; PMCID: 2883899.

170. Adamson B, Norman TM, Jost M, Cho MY, Nunez JK, Chen Y, Villalta JE, Gilbert LA, Horlbeck MA, Hein MY, Pak RA, Gray AN, Gross CA, Dixit A, Parnas O, Regev A and Weissman JS. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. Cell. 2016;167(7):1867. PubMed PMID: 27984733; PMCID: PMC5315571.

171. Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, Marjanovic ND, Dionne D, Burks T, Raychowdhury R, Adamson B, Norman TM, Lander ES, Weissman JS, Friedman N and Regev A. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. Cell. 2016;167(7):1853. PubMed PMID: 27984732; PMCID: PMC5181115.