HHS Public Access

Author manuscript

Circulation. Author manuscript; available in PMC 2023 February 08.

Published in final edited form as:

Circulation. 2022 February 08; 145(6): 469–485. doi:10.1161/CIRCULATIONAHA.121.057789.

ZEB2 Shapes the Epigenetic Landscape of Atherosclerosis

Paul Cheng, MD, PhD[1,#], **Robert C. Wirka, MD**[2,#], **Lee Shoa Clarke, MD**[1], **Quanyi Zhao, PhD**[1], **Ramendra Kundu, PhD**[1], **Trieu Nguyen, MS**[1], **Surag Nair, PhD**[3], **Disha Sharma, PhD**[1], **Hyun-jung Kim, PhD**[1], **Huitong Shi, PhD**[1], **Themistocles Assimes, MD, PhD**[1], **Juyong Brian Kim, MD**[1], **Anshul Kundaje, PhD**[3], **Thomas Quertermous, MS, MD**[1,*]

[1]Division of Cardiovascular Medicine and the Cardiovascular Institute, Stanford, CA

[2]Division of Cardiology, Departments of Medicine and Cell Biology and Physiology, McAllister Heart Institute, University of North Carolina, Chapel Hill, NC

[3]Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305;

Abstract

Background: Smooth muscle cells (SMC) transition into a number of different phenotypes during atherosclerosis, including those that resemble fibroblasts and chondrocytes, and make up the majority of cells in the atherosclerotic plaque. To better understand the epigenetic and transcriptional mechanisms that mediate these cell state changes, and how they relate to risk for coronary artery disease (CAD), we have investigated the causality and function of transcription factors (TFs) at genome wide associated loci.

Methods: We employed CRISPR-Cas 9 genome and epigenome editing to identify the causal gene and cell(s) for a complex CAD GWAS signal at 2q22.3. Subsequently, single-cell epigenetic and transcriptomic profiling in murine models and human coronary artery smooth muscle cells were employed to understand the cellular and molecular mechanism by which this CAD risk gene exerts its function.

Results: CRISPR-Cas 9 genome and epigenome editing showed that the complex CAD genetic signals within a genomic region at 2q22.3 lie within smooth muscle long-distance enhancers for *ZEB2*, a TF extensively studied in the context of epithelial mesenchymal transition (EMT) in development and cancer. *ZEB2* regulates SMC phenotypic transition through chromatin remodeling that obviates accessibility and disrupts both Notch and TGFβ signaling, thus altering the epigenetic trajectory of SMC transitions. SMC specific loss of *ZEB2* resulted in an inability of transitioning SMCs to turn off contractile programing and take on a fibroblast-like phenotype,

[*]**Materials and Correspondence:** Thomas Quertermous MD, 300 Pasteur Dr., Falk CVRC, Stanford, CA 94305, tomq1@stanford.edu, Tel: 650-723-5012, Fax: 650-725-2178.

[#]Co-First Authors

Author Contributions

PC: Designed research studies, conducted experiments, acquired and analyzed data, and wrote the manuscript.
TQ: Designed research studies, helped with data analysis, and wrote the manuscript.
RW/JK/TN/RK: Conducted experiments and acquired data.
LSC/TA: Designed and analyzed enrichment analysis from GWAS data
QZ/HS/SN/HK/DS/AK: Helped with analyzing data and other critical scientific input.

Disclosures: The authors have no competing interests to declare.

but accelerated the formation of chondromyocytes, mirroring features of high-risk atherosclerotic plaques in human coronary arteries.

Conclusions: These studies identify *ZEB2* as a new CAD GWAS gene that affects features of plaque vulnerability through direct effects on the epigenome, providing a new thereapeutic approach to target vascular disease.

Keywords

smooth muscle; atherosclerosis; genetics; epigenome; single cell

Introduction

Mapping gene regulatory variation and advances in highly multiplexed genotyping have allowed us to definitively identify a large number of genetic loci that affect risk for complex human diseases through genome-wide association studies (GWAS). Translating GWAS signals into mechanistic understanding, however, has been limited as most genetic signals lie within non-coding regions of the genome, which remain poorly annotated[1, 2]. Thus, identification of the proper gene and cell type is complex and translation of genetic discoveries into disease mechanisms is difficult[2]. Human coronary artery disease (CAD), for example, has over 160 associated genome-wide significant loci, many with uncertain target genes[3–5]. Critical first steps toward realizing the therapeutic potential related to CAD genetics is to identify causal genes in associated loci, the relevant causal cell type(s), and functionally link these genes into regulatory networks and targetable biological processes [6–9].

There is mounting evidence that smooth muscle cell (SMC) state changes play an important role in vascular diseases[3, 10–14]. The process by which SMC undergo phenotypic modulation in the face of vascular stress has previously been studied at the cellular and molecular level[15], but how genetic variation might modulate specific aspects of this process and the directionality of these effects has not been investigated. Indeed, it has been postulated that genes affecting different aspects of these cell state changes can have either disease promoting or inhibitory effects[14, 16]. For instance, renewed expression of the embryonic factor *TCF21*, which has a role in coronary artery SMC development, has been shown to promote de-differentiation of SMC into a fibroblast-like fibromyocyte (FMC) phenotype that may be protective toward CAD risk[13, 14]. SMC can also adopt an endochondral-like "chondromyocyte" SMC phenotype (CMC)[14, 17, 18], with pseudotime trajectories suggesting these cells potentially arise from FMC. Identification and characterization of additional CAD associated genes that regulate SMC phenotypic transitions promises to be informative regarding molecular effects on cellular anatomy and shed light on the direction of effect toward disease risk.

Allelic variation in a 3 million basepair (Mb) genomic region at 2q22.3 has reached genome-wide significance as a disease modifier for coronary artery disease risk through numerous meta-analyses of CAD GWAS data [5, 19–21]. Although distinguished by the absence of protein coding genes, this area contains evolutionarily conserved elements that likely harbor enhancers. These putative enhancers encompass SNPs in tight linkage disequilibrium to

the different lead CAD GWAS variants in this region of the genome. *ZEB2* is the gene physically located closest to the cluster of tightly LD-linked lead variants identified through GWAS, rs2252641 [5, 19], rs35500812 [20], and rs957293 [21]. *ZEB2,* a zinc finger homeodomain transcription factor, shares significant homology to *ZEB1* which is known to modulate SMC behavior[22]. Furthermore, *ZEB2* is a binding partner for another CAD associated gene, *SMAD3* [23]. Interestingly, ZEB2 has been primarily characterized as a key regulator of epithelial-mesenchymal transition (EMT), a phenotypic switch with related increased cellular proliferation and migration, but *ZEB2*'s involvement in atherosclerosis is unknown.

In the studies reported here, we combined genome and epigenome editing to establish that the genetic signal at 2q22 encodes a long-distance SMC enhancer for *ZEB2*. SMC-specific deletion of *Zeb2* coupled with single-cell transcriptomic and epigenetic profiling of SMC-specific lineage traced cells revealed that this CAD associated EMT gene dramatically alters cell state trajectories of SMC through epigenetic regulation of TGFβ and Notch signaling, resulting in altered SMC phenotypes contributing to atherosclerotic plaque composition.

Methods

Data availability

All data and materials are currently being deposited to the NCBI Gene Expression Omnibus and will be available for general public access upon acceptance of this manuscript for publication. (Accession numbers currently pending.)

Mouse gene manipulation

SMC-specific lineage tracing and gene knockout in atherosclerotic model as previously described[13]. The *Zeb2* conditional knockout mice were obtained from the Murphy lab from Washington Univ[57]. The animal study protocol was approved by the Administrative Panel on Laboratory Animal Care at Stanford University, and procedures followed were in accordance with institutional guidelines.

Mouse aortic tissue dissociation, cell capture, and sequencing

Please see a detailed description of these methods provided in the Supplemental Methods.

Human coronary artery tissue collection

Human coronary arteries used in this study were dissected from explanted hearts of transplant recipients, and were obtained from the Human Biorepository Tissue Research Bank under the Department of Cardiothoracic Surgery from consenting patients with approval from the Stanford University Institutional Review Board.

Murine aortic tissue processing and histology

Immediately after sacrifice, mice were perfused with 0.4% paraformaldehyde (PFA), aortic tissue harvested, embedded in OCT and sectioned. Immunofluorescence and immunohistochemistry were performed, and lesion areas proflied as previously described [13, 17]. For RNAscope, slides were processed and hybridized according to the manufacturer's

instructions, with reagents from ACD Bio. Please see a full description of staining and quantification in the Supplemental Methods.

Analysis of single cell sequencing data

Please see a full description of methodology for the acquisition and analysis of both scRNAseq and scATACseq data in the Supplemental Methods.

In vitro genomic studies in HCASMCs and A549 cells

Please see a full description of methodology for the culture of HCASMC, A549 and in vitro genomic studies in the Supplemental Methods.

Statistical methods

Differentially expressed genes in the scRNA-Seq data were identified using a Wilcoxon rank-sum test. Significance determination of histological measurement, luciferase studies, qPCR results, and composite gene-score were done via unpaired two-tailed t-test for normally distributed variables, and Mann Whitney test for non-normally distributed values. Multiple comparisons were corrected via Bonferroni correction when necessary.

Results

CAD associated variation at chromosome 2q.22.3 localizes to SMC-specific enhancers that modulate expression of ZEB2

The genetic locus at 2q22.3 reached genome-wide significance for association with risk of coronary artery disease nearly a decade ago[19] and has been replicated in the past few years [5, 20] (Fig. 1A). This signal lies within a 3 Mb gene region with the closest protein coding gene, *ZEB2*, being more than 600,000 bp away. The most recent GWAS meta-analyses have demonstrated several large distinct haploblocks of highly linked single nucleotide polymorphisms (SNPs) in this locus, all significantly associated with risk for CAD [5, 19–21], with the most recent GWAS showing a lead SNP p-value of 5.4×10^{-13}. This region of the genome contains elements which are evolutionarily conserved even to lower vertebrates (Fig. 1B), suggesting the presence of regulatory functional elements [24]. We identified six conserved candidate enhancer regions in this area of the human genome (Fig. 1C). To gain information regarding the possible enhancer activity of these conserved regions in human coronaries, we generated enhancer-specific histone modification (histone 3 lysine 27 acetylation) (H3K27ac) ChIP-Seq data for human coronary artery SMC (HCASMC) (Fig. 1D). H3K27ac peaks overlapped with the conserved regions, suggesting these regulatory elements were likely functional enhancers. We then analyzed chromatin conformation capture (Hi-C) data [25] generated in HCASMC. Fit-Hi-C visualization revealed these evolutionarily conserved regions are topologically associated with the promoter of *ZEB2* through looping structures (Fig. 1D). Together, these data suggest that the genetic signal at 2q22.3 lies within a long-distance enhancer for *ZEB2*. It is important to note that lncRNAs *TEX41* and *LOC101928386* are located in this region, and *TEX41* linked to aortic valve stenosis [26], but we found no evidence for their being involved with the ZEB2 – enhancer interaction.

In order to ascertain whether the SNPs within and around these evolutionarily conserved regions influence *ZEB2* gene expression, and whether these gene expression changes are associated with alteration in risk of CAD in humans, we colocalized the CAD GWAS signal with human expression quantative trait locus (eQTL) data available from the human genotype-tissue expression (GTEx) database [4, 27]. This analysis revealed strong correlation between SNPs associated with CAD risk and those that influence arterial *ZEB2* expression (Fig. 1E) but not expression of other nearby genes (Sup Fig. 1A). Specifically, integration and combined analysis of genome wide transcriptomic and CAD association data indicated that variant alleles related to greater expression of *ZEB2* are protective against CAD risk[28].

To experimentally validate the functional significance of these epigenetic marks and looping chromatin interactions, we performed CRISPR-Cas9 mediated genome editing of this conserved region in A549 cells. A549 cells were chosen for their high *ZEB2* expression, their ability to model EMT, and to avoid HCASMC which senesce after a few cell divisions in culture and thus cannot be clonally expanded. Taking advantage of the different forms of genomic repair after Cas9 cutting, we generated distinct clones after treatment with targeted RNA guides (Fig. 1F). *ZEB2* expression was significantly lower in clones with deletion of the entire enhancer region than those that underwent homologous repair at the CRISPR sites (Fig. 1G). The expression of the next nearest gene was not altered (Sup Fig. 1B). To further investigate whether any of these conserved elements have enhancer activity alone, we employed epigenome silencing in each of these regions in HCASMC using dCas9-KRAB combined with specific guides targeting each conserved element [29]. These experiments revealed three distinct active enhancers for *ZEB2*, as evidence by decreased *ZEB2* expression level with the CRISPRi enhancer silencing (Figs. 1H, I). Taken together, these data suggest that CAD associated SNPs and their related enhancers regulate *ZEB2* expression in the arterial wall to modulate the risk for myocardial infarction.

Zeb2 expression is activated with SMC phenotypic transition and globally decreases chromatin accessibility around ZEB binding motifs

To determine which cells in atherosclerotic lesions express *Zeb2*, we performed single cell RNA sequencing (scRNAseq) on aortic root tissue from $ApoE^{-/-}$ mice with *Myh11*-Cre directed SMC lineage tracing ($Myh11^{CreERT2}$,$ROSA^{tdT/+}$, $ApoE^{-/-}$), at baseline and after 16 weeks of high fat diet (Fig. 2A). In the absence of atherosclerosis, *Zeb2* was expressed in macrophages and fibroblasts, but absent in lineage traced quiescent SMCs (Fig. 2B, C, D). Fibroblast-like “fibromyocyte” (FMC) and chondrocyte-like “chondromyocyte” (CMC) cells were identified, as previously described [13, 14, 17, 18]. *Zeb2* expression was noted to be activated in the process of phenotypic transition during the development of atherosclerotic plaques (Fig. 2C). In addition, RNA in situ hybridization identified *Zeb2* expression by a large number of lineage-traced SMC just below the lesion cap, and also by scattered macrophages and adventitial cells (Fig 2E, F, G). The same ZEB2 expression pattern was observed in human coronary arteries (Sup Fig. 2A).

To identify which cells utilize the distal ZEB2 enhancers, and to better understand how *Zeb2* regulates the epigenome to influence the biology of atherosclerotic plaque, we mapped single cell chromatin accessibility with the assay for single cell transposase-

accessible chromatin with sequencing (scATACseq) within murine aortic atherosclerotic tissue. The identity of each cluster was determined using a “pseudo-expression” metric based on chromatin accessibility around specific marker genes, with resulting clusters similar to those derived from the scRNAseq data (Fig. 2H). Mapping the CAD linked human enhancers (Figs. 1B, C) on chromosome 2 to conserved regions of the murine genome allowed assessment of the cell-specific accessibility at the orthologous murine chromosomal region for the different cell types in the diseased tissue (Fig. 2I). These data identified peaks representing closed chromatin for macrophages and endothelial cells, compared to open chromatin in transition smooth muscle cells. Comparing transitional to quiescent SMC revealed dynamic activation of this region of chromatin in the process of atherosclerosis. Open chromatin was also observed in pericytes and adventitial fibroblasts but not dynamically regulated, and these cells are not expected to impact lesion pathology. The observed disease related dynamic regulation of *Zeb2* expression and chromatin accessibility at regions in the murine genome that correspond to the location of the GWAS genetic signal suggest *ZEB2's* CAD risk modifying effect is mediated through SMC but not endothelial cells or macrophages.

In order to gain a better understanding of *Zeb2* function in atherosclerosis, we merged scRNAseq data with scATACseq data from SMC-derived cells in murine atherosclerotic lesions [30]. The combined dataset representing lineage traced cells was readily clustered into quiescent and transition SMC (Fig. 2J). We used the combined matrix with Monocle3 pseudo-temporal analysis[31] of the SMC cell state changes to develop a trajectory of the SMC transition states (Fig. 2K). The smooth muscle trajectory starts with high contractile gene expression, e.g. *Cnn1*, that diminishes as SMC lineage cells take on the fibromyocyte phenotype (Fig. 2L). *Zeb2* expression is present in the transition between quiescent and phenotypically modulated cells, with expression of *Zeb2* being inversely correlated with expression of quiescent SMC markers *Cnn1*, *Myocd,* etc. (Sup Fig. 2B). Importantly, initiation of *Zeb2* expression is accompanied by a decrease in accessibility around Zeb binding sites genome wide, as identified with ChromVAR (Fig. 2M, N, O) [32]. In the cells that appear to be further along the trajectory, toward the chondromyocyte (CMC) phenotype, *Zeb2* expression is lost and accessibility around Zeb motifs is partially recovered. Taken together, these findings suggest that *Zeb2* primarily functions as an epigenetic repressor, closing chromatin globally at sites where Zeb motifs are present. Depressed relative genome accessibility around Zeb motifs is further indicated by “footprinting” analysis at Zeb binding sites genome wide (Fig. 2P), demonstrating decreased probability of Tn5 insertion 200 bp either side of the Zeb binding motifs. This is not seen with other classical transcription factors (Sup Fig 2C).

The pseudo-temporal anti-correlation between *Zeb2* expression and mature SMC gene expression, along with its apparent repressive effect on chromatin accessibility, suggest it plays a role in phenotypic transition as a repressor of the SMC program to allow for mature SMC to transition into an FMC phenotype. Given that *Zeb2* is a prominent factor in promoting epithelial mesenchymal transition (EMT), we used the single cell transcriptomic and chromatin accessibility data to investigate the expression of EMT-associated genes in SMC phenotypic transitions. Transition and cell-specific expression of classic EMT-associated genes, *Twist1, Snail1, Snail2, Gsc*, and, *Fn1* were examined within the combined

scRNAseq and scATACseq object, and were found to be dynamically activated in the process of phenotypic modulation (Sup Fig 2D). In combination with human genetic data, these findings suggest that altered *ZEB2* expression may modulate SMC phenotype and likely represents the mechanism by which this causal gene regulates the risk for coronary artery disease in humans.

SMC-specific Zeb2 knockout alters the epigenetic trajectory specifically in transition SMC

The specific role that *Zeb2* plays in SMC transition during atherosclerosis was investigated with conditional *Zeb2* knockout mice. In 8-week old mice, prior to the onset of atherosclerosis, SMC-specific *Zeb2* deletion and concurrent activation of the *tdTomato* lineage reporter were accomplished using an inducible *Myh11-Cre* allele in $Myh11^{CreERT2}$, $Zeb2^{\Delta SMC/\Delta SMC}$, $ROSA^{tdT/+}$, $ApoE^{-/-}$ ($Zeb2^{\Delta SMC}$) mice. Tamoxifen was administered to activate Cre function and animals were then maintained for 16 weeks on high fat diet to induce atherosclerosis. Atherosclerotic lesions in the aortic root were then dissected and processed for scATACseq, scRNAseq, or histological analysis (Fig. 3A) [13, 17, 33]. scATACseq data was processed, clustered based on differentially accessible peaks, and visualized using dimensional reduction with UMAP (Fig. 3B). The identity of each cluster was determined using the pseudo-expression metric based on chromatin accessibility around specific marker genes (Sup Fig. 3). Loss of *Zeb2* resulted in marked global alteration in accessibility specifically in transition SMC (Fig 3C). Importantly, the epigenetic landscape for endothelial cells, inflammatory cells, macrophages, and *Myh11*-expressing mature SMC was unchanged (Fig. 3C).

To better visualize the epigenetic alterations specifically in $Zeb2^{\Delta SMC}$ cells, we evaluated only the SMC lineage-traced cells, and repeated dimensionality reduction using principal components that defined the epigenetic variance in this population. Differentiated SMC expressed sarcomeric proteins readily separated from transition smooth muscle cells based on their marker gene chromatin accessibility (Fig. 3C, D, E). While the control and $Zeb2^{\Delta SMC}$ cells overlapped entirely in their epigenome in the quiescent stage, the modulated SMC showed drastic global epigenetic alterations. Consistent with our hypothesis that *Zeb2* suppresses mature SMC marker gene expression, we saw an increase in chromatin accessibility at mature SMC genes in modulated $Zeb2^{\Delta SMC}$ cells, compared to cells from control mice (Fig. 3F). These data suggest that *Zeb2* has a role in transition SMC to suppress mature SMC marker genes, and may also inhibit modulated SMC from adopting the chondromycyte phenotype.

To better understand the transcriptional program that is specifically repressed by *Zeb2*, we compiled regions across the genome that become preferentially open after the loss of *Zeb2*. As expected, these regions were enriched for Zeb binding sites (Fig. 3G). $Zeb2^{\Delta SMC}$ cells revealed differentially open chromatin at mature SMC markers such as *Cnn1, Myh11,* and *Myocd*, and there appeared to be earlier and more common chromatin accessibility in the trajectory around chrondrogenesis-related genes such as *Spp1, Sox9* and *Col2a1* [17] (Figs. 3F, G). Further, differential open chromatin was enriched for both Notch effector Rbpj binding motifs, as well as TGFβ effector, Smad, binding motifs (Fig. 3H). Interestingly the degree of enrichment for Rbpj and Smad binding was similar to that observed for Zeb motifs. These

findings suggest that Zeb2 specifically modulates the chromatin accessibility around regions which mediate Notch and TGFβ signaling. Consistent with Zeb2 being antagonistic to Notch signaling, there appeared to be a strong negative correlation between *Zeb2* expression and accessibility at Notch target genes such as *Hes1*, and an overall inverse relationship between expression of Zeb2 and Notch target genes in the process of SMC transition (Fig. 3G, I). Finally, to investigate the pathways that are differentially regulated by *Zeb2* expression, we used the GREAT algorithm [34] to identify genes that are the likely target of genomic regions that become accessible with *Zeb2* knockout and investigated the pathways enriched with these genes (Fig. 3J). Terms related to EMT, including "cell-substrate adhesion," "adherens junction organization", and "cell junction organization" were identified. Also, consistent with Zeb2 being a Smad3 interacting and TGFβ signaling regulatory factor [23], multiple terms specifically related to TGFβ signaling, including "response to TGFβ signaling" and "negative regulation of cellular response to TGFβ stimulus" were enriched for *Zeb2* target genes.

SMC-specific Zeb2 knockout results in decreased plaque transition SMC, disrupted cap organization and accelerated chondrogenic transformation

Single cell RNA sequencing data from control and $Zeb2^{\Delta SMC}$ mice were processed as previously described [13, 30, 33], and unbiased UMAP clustering identified different cell groups that contribute to the lesion (Fig. 4A). As with scATACseq data, loss of *Zeb2* resulted in transcriptional alteration and phenotypic change in only the transition SMC population, as marked by lineage tracing (Fig.4A, B, C). Consistent with Zeb2's epigenetic role, these changes in SMC cluster composition indicated a substantial change in the SMC transition trajectory across transcriptome-wide UMAP space. To provide detailed investigation of the transcriptomic changes associated with *Zeb2* loss, we focused analyses on the lineage traced SMC (Fig. 4B, C). The loss of *Zeb2* significantly decreased the fraction of transition SMC that take on the FMC phenotype and increased the quiescent mature SMC population (Fig. 4D–I). Although the total fraction of lineage traced SMC that takes on the chondromyocyte fate is unchanged, the total number of CMC represented an increased percentage of the total transition cells due to the decrease in FMC (Fig. 4H, I, J). The loss of *Zeb2*, therefore, appears to impair the ability of quiescent SMC to undergo phenotypic transition in the setting of atherosclerosis, and inhibit SMC from assuming the fibroblast-like proliferative intermediate FMC phenotype, shunting transition cells toward the CMC phenotype. This is consistent with the single-cell epigenetic findings where loss of *Zeb2* resulted in transition SMC that appeared epigenetically further away from the fibroblast population (Fig. 3C). The FMC population exhibits the highest expression of cell cycle genes such as *CyclinD1*, mirroring "transit amplifying" populations identified in other tissues in the context of regeneration and repair processes [35], as well as cells that enter the mesenchymal transition in EMT.

Consistent with our epigenetic finding that loss of *Zeb2* increases chromatin accessibility in regions of the genome that harbor Smad binding sites, *Zeb2* loss likely leads to increased overall TGFβ-Smad signaling. The SMC phenotype seen here appears to be opposite to the phenotype observed with SMC loss of *Smad3* as we have previously observed, i.e., increased phenotypic modulation and increased transition to the most fibroblast-like *Mmp3*

and *Cxcl12* expressing remodeling SMC (R-SMC) phenotype cells [33]. With $Zeb2^{\Delta SMC}$ cells we observed decreased phenotypic modulation as well as decreased expression of characteristic R-SMC genes *Mmp3* and *Cxcl12* and loss of almost all R-SMC cluster cells (Fig. 4K–N).

Consistent with our hypothesis that *Zeb2* is critical for suppressing expression of mature SMC genes in the setting of phenotypic transition, with histological analysis we observed frequent ectopic expression of the mature SMC gene calponin 1 (*Cnn1*) in isolated smooth muscle transition cells in the central portion of the plaque below the fibrous cap (Fig. 4O–R). Such displaced and isolated individual cells were rarely observed in the control lesions, where these mature markers were seen almost exclusively in the media and the fibrous cap in an organized layer. In addition, while chondromyocytes and their endochondral marker genes such as *Col2a1* were normally observed only in the base of the plaque near the internal elastic lamina [17, 33], we saw expression of *Col2a1* in individual cells in the body of the plaque close to the region of the fibrous cap (Fig. 4S, T). The latter finding suggests that transition SMC are turning into chondromyocytes prematurely in the SMC transition trajectory. This is consistent with the observed lineage preference for chondromyocytes in our scRNAseq data.

Further histological analysis of the plaque identified a decrease in total tdTomato positive SMC contribution to the atherosclerotic plaque (Fig. 5A, B), although the overall lesion size was not significantly decreased (Fig. 5C). Trichrome staining of the lesions revealed that while plaque burden was unchanged, there was an increase in acellular areas in the lesions (Fig. 5 D, E) in regions usually occupied by modulated SMC. Plaque lipid content as assessed with Oil Red O staining as well as plaque macrophage content (CD68 staining) and cap thickness were not different between wildtype and knockout mice (Fig 5F–J). These findings are consistent with recent data from other groups suggesting that decreased lesion SMC increases high-risk plaque features but does not affect plaque volume [36].

Further, scRNAseq data was employed to identify fundamental pathways differentially represented in wildtype versus $Zeb2^{\Delta SMC}$ cells. The Seurat FindMarker algorithm was used to identify differentially regulated genes that were employed to search for enrichment in biological function gene ontology genesets. This analysis identified terms that were highly consistent with observations from the single cell scATACseq data (Sup Table 1) including those related to cell migration, smooth muscle proliferation, and Tgfβ signaling, all relevant for EMT, as well as endochondral bone formation terms, including “cartilage development,” and “regulation of ossification”. This latter finding was consistent with the data showing an accelerated and relative increase in CMC among transition cells in the $Zeb2^{\Delta SMC}$ mice.

ZEB2 modulates human coronary artery smooth muscle cell fate, in part through epigenetic inhibition of TGFβ and NOTCH signaling

To determine whether SMC cell fate modulation by *ZEB2* is preserved in human SMC, we knocked down *ZEB2* (*ZEB2*-KD) in cultured primary human coronary artery smooth muscle cells (HCASMC). Loss of *ZEB2* resulted in increased expression of mature SMC genes such as *CNN1* and *TAGLN,* and a decrease in FMC genes such as *TCF21* and *KLF4* (Fig 6A), consistent with the *in vivo* phenotype. Over-expression of *ZEB2* resulted in the

opposite effect on *CNN1* gene expression and mirrored the effect of NOTCH inhibition (Fig. 6B), as observed *in vivo*. Consistent with loss of the more proliferative transient-amplifying population of FMC observed *in vivo*, loss of *ZEB2* resulted in a decrease in the rate of HCASMC division as assayed by EDU uptake, which is consistent with the decrease in SMC lineage contribution to plaque in the murine models (Fig. 6C). Bulk RNA sequencing in knockdown cells showed loss of *ZEB2* resulted in increased expression of many canonical NOTCH responsive genes, including *HEY2, HES1,* and *HES2*. (Sup. Table 2). Also, the global effect of *ZEB2*-KD appeared to phenocopy the effect of TGFβ stimulation on gene expression at a whole genome level. This was evident from the strong correlations between *ZEB2*-KD and TGFβ stimulation, with 84% of significantly differentially regulated genes altered in identical directions (Fig. 6D). To further investigate the interaction of ZEB2 and SMAD3 characterized in other cell types, we performed co-immunoprecipitation of ZEB2 protein and verified that SMAD3 was co-precipitated from HCASMC nuclear extracts (Sup. Fig. 4A). Over-expression of *ZEB2* in SMCs appeared to decrease the luciferase activity of both SMAD responsive elements as well as conserved sequence upstream of *CNN1* that contain conserved ZEB, SMAD, and RBPJ binding sites (Sup. Fig. 4B). These data suggest that ZEB2 may have a direct effect on transcription in addition to the characterized epigenetic effects.

To determine whether the Zeb2 effects on chromatin accessibility observed in the mouse model were also observed in HCASMC, we performed bulk ATAC sequencing with these cells. A similar dip in accessibility was seen globally around ZEB binding sites genomewide, as seen in our *in vivo* model (Fig. 6E), in a pattern distinct from other classical transcription factors (Sup Fig 4C). We trained a base-resolution neural network to predict chromatin accessibility based on the in vitro ATACseq data [37] to determine the effect of adding a ZEB binding site *in silico*. The model predicted that adding ZEB binding sites would decrease the expected accessibility in HCASMC (Sup Fig 4D). Given that multiple TFs have motifs that may be similar to ZEB2, to further test specifically the chromatin-closing ability of ZEB2, we performed perturb-ATACseq with *ZEB2*-KD in these cells. Consistent with our *in vivo* findings that Zeb2 functions as a repressor of chromatin accessibility, loss of *ZEB2* resulted in a global increase by greater than 50% for 15% of all ATAC peaks (14,258 of 93,391 total peaks), whereas only 2% of peaks (2,276 of 9,3391 total peaks) had decreased accessibility (Fig. 6F). Motif analysis for *ZEB2*-KD differentially open peaks revealed significant enrichment in ZEB binding sites. SMAD and RBPJ binding sites were also enriched as noted with the *in vivo* murine disease model, though to a lesser extent (Fig. 6G, and Supplementary Table 3).

To investigate whether NOTCH inhibition is a direct result of ZEB2 mediated epigenetic regulation or simply an indirect effect from alteration of cell-fate, we specifically investigated chromatin accessibility around regulatory regions of NOTCH responsive genes *HES1, HES2, HEY1,* and *HEY2* as well as the SMC lineage determining factor *MYOCD*. Consistent with a role for ZEB2 in epigenetically regulating NOTCH-RBPJ signaling, loss of *ZEB2* resulted in a significant increase in accessibility around the promoters of all these genes (Fig. 6H). NOTCH, specifically Notch3, activation is critical for arterial SMC lineage determination during development [38]. Further, NOTCH3 is known to be activated by SMC at the fibrous cap in human atherosclerotic lesions [39], and inhibition of Notch

signaling appears to be necessary for murine SMC to undergo phenotypic modulation in atherosclerosis [40]. We hypothesized, therefore, that ZEB2 may be regulating this NOTCH mediated lineage determination process and that ZEB2-mediated inhibition of NOTCH signaling promotes phenotypic modulation. Consistent with this hypothesis, activation of NOTCH signaling in HCASMC by over-expression of NOTCH3 intracellular domain resulted in an increase in both NOTCH-responsive genes such as *HEY2*, and mature SMC markers such as *CNN1*. Both of these NOTCH-mediated effects were blocked by simultaneous over-expression of *ZEB2* (Figure 6I). While HCASMC do not easily form chondromyocytes in culture, an increase in chromatin accessibility was also observed around endochondral calcification genes such as *SOX9* and *SPP1*, suggesting that ZEB2 target genes also contribute to the lineage preference for chondromyocyte cell fate (Fig. 6H).

Functional analysis using the GREAT algorithm demonstrated that differentially regulated ATACseq peaks were enriched for genes with biological processes related to SMC biology and angiogenesis, as well as epithelial biology, despite this cell type not really being a single layered epithelium (Fig. 6J). These findings were consistent with our in vivo results that *ZEB2* epigenetically regulates smooth muscle phenotypic modulation in an EMT-like process.

Discussion

In this study, we used CRISPR-Cas9 genome and epigenome editing to demonstrate that the genome wide CAD association signals at chromosome 2q22.3 are located in distal vascular smooth muscle enhancer(s) for the *ZEB2* gene. Single cell RNAseq and ATACseq studies in mouse atherosclerotic tissues indicated that *Zeb2* is expressed transiently in SMC as they de-differentiate and undergo transition to a modulated phenotype and localize near the fibrous cap. Pseudotime analysis of the scATACseq data mapped the trajectory of contractile SMC transitioning to fibromyocytes which give rise to chondromyocytes. For contractile SMCs positioned early in the trajectory before *Zeb2* expression, chromatin regions containing Zeb2 binding motifs were open, but these regions were closed in FMC expressing *Zeb2*. These data are consistent with an epigenetic repressive effect toward gene expression of SMC lineage-determining genes. SMC-specific knockout of *Zeb2* led to a marked shift in the SMC transition trajectory with a significant decrease in the number of FMC and a relative increase in premature differentiation to CMC. Interestingly, a number of the transition knockout SMCs continued to ectopically express contractile markers in the middle of the plaque and failed to proliferate. Motif enrichments in regions of altered accessibility and gene expression suggested modulation of TGFβ and NOTCH signaling as key drivers in Zeb2 mediated epigenetic and transcriptional processes.

Macroscopic plaque features observed in SMC-specific loss of *Zeb2* mirrored several high-risk features of human atherosclerotic plaques. Decreased SMC content in plaques and increased acellular regions have long been associated with an increased risk of plaque rupture and are observed in these $Zeb2^{\Delta SMC}$ lesions. Furthermore, the disorganized pattern of calponin staining is akin to heterogenous plaques with disrupted fibrous caps, a feature associated with higher risk of plaque rupture [41–43]. Also important is the observed early chondromyocyte differentiation in more superficial regions of the intima in a disorganized

or dispersed pattern resembling spotty calcifications seen on both IVUS and CT that are characteristic signs of plaques that may rupture [44–46]. The presence of these high-risk features may explain the increased risk of myocardial infarction in humans with genetic variation linked to lower in vivo *ZEB2* expression in SMC and suggests *ZEB2* is an epigenetic regulator of these high-risk features. Of note, the same genomic region where the CAD variation is localized also contains lead SNPs discovered by GWAS for human aortic stenosis [26], an interstitial calcification process heavily regulated by Notch signaling [47, 48]. In these studies, *ZEB2* and *TEX41* were identified as possible causal genes. This observation suggests our findings likely extend to other important cardiovascular conditions. Finally, studies with HCASMC validated the murine model findings, showing that *ZEB2* opposes TGFβ and NOTCH signaling, and suppresses chromatin accessibility at sites containing *ZEB* binding sites in the vicinity of genes that regulate EMT like processes, NOTCH signaling, and chondrogenesis.

This work represents the first genome-wide study of vascular SMC epigenomics at the single cell level and significantly extends our understanding of how SMC transitions are regulated by targeted global modification of chromatin accessibility. Previous studies have identified other chromatin modifiers such as KLF4 [49] and TET2 [50] as regulators of plaque SMC behavior, suggesting that much of SMC phenotypic modulation is orchestrated by precise epigenetic modifications. Disruption of this process, such as those induced by one of many FDA-approved epigenetic modifying therapies, including hypomethylating agents and HDAC inhibitors, may harbor previously unrecognized risks or benefits for cardiovascular events such as myocardial infarction. Targeted therapeutics that augment the epigenetic effect of *ZEB2* could also be explored as a possible treatment for CAD. Further analysis of the scATACseq data will allow identification of novel biological mechanisms of disease risk.

The discovery of an EMT regulator, *ZEB2*, as a modulator of SMC phenotype and CAD risk is revealing. Although not previously discussed or investigated in detail, the de-differentiation, proliferation and migratory activity of epithelial cells during EMT is analogous to activities exhibited by transition SMC. Both processes are associated with the down-regulation of mature lineage marker genes, and the activation of a similar complement of signature genes. This connection is further supported by numerous CAD GWAS signals near other EMT related genes, including *TGFβ1* and *TWIST1*. Further, It is important to consider the implications of *Zeb2* expression in SMC just below the fibrous cap. Through this observation and other results described above, a compelling argument can be made that the migration of transition SMC occurs first from the media to the fibrous cap, with subsequent migration into the inner plaque. This possibility contradicts common belief that the fibrous cap forms from plaque cells regaining SMC gene expression and contractile phenotype [40]. This possibility is further supported by the observation that fibrous cap cells expressing mature SMC markers are frequently seen in early lesions without SMC-derived plaque cells [51]. The *Zeb2* expression pattern and its correlation to SMC-trajectory offers convincing support for this hypothesis (Fig. 7). The movement of SMC into the plaque is analogous to an epithelial cell EMT process and is likely regulated by *Zeb2*, given the disruption of the disease SMC single cell trajectory phenotypes. For instance, we identified individual transition SMC below the fibrous cap that express *Cnn1* and *Myh11*, apparently unable to adopt the FMC phenotype. Further, we also identified *Col2a1* expressing cells,

presumably CMC, in the body of the lesion instead of the base of the plaque near the internal elastic lamina where we have characterized their normal location [17,33]. This finding suggests that these cells are prematurely adopting the CMC phenotype while migrating in an abluminal direction. Through better understanding of these complex SMC transitions we may be better able to understand the contribution of each of these cell types to CAD risk.

Despite the highly significant effect of *Zeb2* deletion on the epigenetic profile of disease cells, there was no change in the magnitude of plaque burden in the *Zeb2* knockout murine disease model. There was instead a significant increase in lesion features that mirror human high-risk plaques. These contradictions between human risk and murine plaque size have been consistently found in non-lipid CAD modifier genes [13, 36, 40, 52–54], many of which modulate intrinsic lesion cell biology. These observations mirror both decades old and emerging human epidemiologic data whereby a variety of higher risk plaque features, rather than plaque size or lumen narrowing, are much better predictors of myocardial infarction [45, 55, 56]. Thus, a convergence of different types of information points to the critical importance of utilizing finer molecular phenotyping at higher resolution to better understand the biology that influences risk for vascular events related to atherosclerosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Special thanks to the Krista Hennig, Yana Ryan, Peter Mcguire and Hassan Chaib at the Stanford Genomic Sequencing and Service Center (GSSC) for performing 10x capture, library construction, and sequencing. We also thank the Stanford shared FACS facility for required FACS analysis and experiments. Illustrations were made with BioRender software. We thank Yujiro Higashi (Department of Perinatology, Institute for Developmental Research, Aichi Human Service Center, Japan) and Kenneth Murphy, University of Washington for supplying the *Zeb2* knockout mice. We also thank Dr. Janet Mertz at University of Wisconsin for supplying the *ZEB2* expression plasmids.

Sources of funding

This work was supported by National Institutes of Health grants K08HL153798 (PC), F32HL143847 (PC), K08HL152308 (RW), K08HL133375 (JBK), R01HL109512 (TQ), R01HL134817 (TQ), R33HL120757 (TQ), R01HL139478 (TQ), R01HL156846 (TQ), R01HL151535 (TQ), R01HL145708 (TQ), as well as a Human Cell Atlas grant from the Chan Zuckerberg Foundation. This work was also supported by American Heart Association grant 20CDA35310303 (PC) and 18CDA34110206 (RW).

Non-standard Abbreviations and Acronyms

scRNAseq	single cell RNA sequencing
scATACseq	single cell chromatin accessibility with the assay for single cell transposase-accessible chromatin with sequencing
FMC	fibromyocyte
CMC	chondromyocyte
GWAS	genome wide association study

EMT	epithelial mesenchymal transition
H3K27ac	histone 3 lysine 27 acetylation
Hi-C	chromatin conformation capture
eQTL	expression quantitative locus
GTEx	genotype-tissue expression database
R-SMC	remodeling smooth muscle cells

References

1. Klein RJ, Xu X, Mukherjee S, Willis J and Hayes J. Successes of genome-wide association studies. Cell 2010;142:350–351. [PubMed: 20691890]
2. Musunuru K and Kathiresan S. Genetics of Common, Complex Coronary Artery Disease. Cell 2019;177:132–145. [PubMed: 30901535]
3. Miller CL, Pjanic M, Wang T, Nguyen T, Cohain A, Lee JD, Perisic L, Hedin U, Kundu RK, Majmudar D, et al. T. Integrative functional genomics identifies regulatory mechanisms at coronary artery disease loci. Nat Commun 2016;7:12092. [PubMed: 27386823]
4. Liu B, Pjanic M, Wang T, Nguyen T, Gloudemans M, Rao A, Castano VG, Nurnberg S, Rader DJ, Elwyn S, et al. Genetic Regulatory Mechanisms of Smooth Muscle Cells Map to Coronary Artery Disease Risk Loci. Am J Hum Genet 2018;103:377–388. [PubMed: 30146127]
5. van der Harst P and Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circ Res 2018;122:433–443. [PubMed: 29212778]
6. Acharya A, Baek ST, Huang G, Eskiocak B, Goetsch S, Sung CY, Banfi S, Sauer MF, Olsen GS, Duffield JS, et al. The bHLH transcription factor Tcf21 is required for lineage-specific EMT of cardiac fibroblast progenitors. Development 2012;139:2139–49. [PubMed: 22573622]
7. Stryjewska A, Dries R, Pieters T, Verstappen G, Conidi A, Coddens K, Francis A, Umans L, van IJcken WFJ, Berx G, et al. Zeb2 Regulates Cell Fate at the Exit from Epiblast State in Mouse Embryonic Stem Cells. Stem Cells 2017;35:611–625. [PubMed: 27739137]
8. Timmerman LA, Grego-Bessa J, Raya A, Bertran E, Perez-Pomares JM, Diez J, Aranda S, Palomo S, McCormick F, Izpisua-Belmonte JC, et al. Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. Genes Dev 2004;18:99–115. [PubMed: 14701881]
9. Xu X, Zhou Y, Xie C, Wei SM, Gan H, He S, Wang F, Xu L, Lu J, Dai W, et al. Genome-wide screening reveals an EMT molecular network mediated by Sonic hedgehog-Gli1 signaling in pancreatic cancer cells. PLoS One 2012;7:e43119. [PubMed: 22900095]
10. Erbilgin A, Civelek M, Romanoski CE, Pan C, Hagopian R, Berliner JA and Lusis AJ. Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. J Lipid Res 2013;54:1894–1905. [PubMed: 23667179]
11. Nurnberg ST, Cheng K, Raiesdana A, Kundu R, Miller CL, Kim JB, Arora K, Carcamo-Oribe I, Xiong Y, Tellakula N, et al. Coronary Artery Disease Associated Transcription Factor TCF21 Regulates Smooth Muscle Precursor Cells that Contribute to the Fibrous Cap. Genom Data 2015;5:36–37. [PubMed: 26090325]
12. Shankman LS, Gomez D, Cherepanova OA, Salmon M, Alencar GF, Haskins RM, Swiatlowska P, Newman AA, Greene ES, Straub AC, et al. KLF4-dependent phenotypic modulation of smooth muscle cells has a key role in atherosclerotic plaque pathogenesis. Nat Med 2015;21:628–637. [PubMed: 25985364]
13. Wirka RC, Wagh D, Paik DT, Pjanic M, Nguyen T, Miller CL, Kundu R, Nagao M, Coller J, Koyano TK, et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. Nat Med 2019;25:1280–1289. [PubMed: 31359001]

14. Alencar GF, Owsiany KM, K S, Sukhavasi K, Mocci G, Nguyen A, Williams CM, Shamsuzzaman S, Mokry M, Henderson CA, Haskins R, et al. The Stem Cell Pluripotency Genes Klf4 and Oct4 Regulate Complex SMC Phenotypic Changes Critical in Late-Stage Atherosclerotic Lesion Pathogenesis. Circulation 2020;142:2045–2059. [PubMed: 32674599]
15. Bennett MR, Sinha S and Owens GK. Vascular Smooth Muscle Cells in Atherosclerosis. Circ Res 2016;118:692–702. [PubMed: 26892967]
16. Cherepanova OA, Gomez D, Shankman LS, Swiatlowska P, Williams J, Sarmento OF, Alencar GF, Hess DL, Bevard MH, Greene ES, et al. Activation of the pluripotency factor OCT4 in smooth muscle cells is atheroprotective. Nat Med 2016;22:657–665. [PubMed: 27183216]
17. Kim JB, Zhao Q, Nguyen T, Pjanic M, Cheng P, Wirka R, Travisano S, Nagao M, Kundu R and Quertermous T. Environment-Sensing Aryl Hydrocarbon Receptor Inhibits the Chondrogenic Fate of Modulated Smooth Muscle Cells in Atherosclerotic Lesions. Circulation 2020;142:575–590. [PubMed: 32441123]
18. Pan H, Xue C, Auerbach BJ, Fan J, Bashore AC, Cui J, Yang DY, Trignano SB, Liu W, Shi J, et al. Single-Cell Genomics Reveals a Novel Cell State During Smooth Muscle Cell Phenotypic Switching and Potential Therapeutic Targets for Atherosclerosis in Mouse and Human. Circulation 2020;142:2060–2075. [PubMed: 32962412]
19. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 2013;45:25–33. [PubMed: 23202125]
20. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, et al. Comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015;47:1121–1130. [PubMed: 26343387]
21. Erdmann J, Kessler T, Munoz Venegas L and Schunkert H. A decade of genome-wide association studies for coronary artery disease: the challenges ahead. Cardiovasc Res 2018;114:1241–1257. [PubMed: 29617720]
22. Nishimura G, Manabe I, Tsushima K, Fujiu K, Oishi Y, Imai Y, Maemura K, Miyagishi M, Higashi Y, Kondoh H, et al. DeltaEF1 mediates TGF-beta signaling in vascular smooth muscle cell differentiation. Dev Cell 2006;11:93–104. [PubMed: 16824956]
23. Postigo AA. Opposing functions of ZEB proteins in the regulation of the TGFbeta/BMP signaling pathway. EMBO J 2003;22:2443–2452. [PubMed: 12743038]
24. Ovcharenko I, Nobrega MA, Loots GG and Stubbs L. ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. Nucleic Acids Res 2004;32:W280–6. [PubMed: 15215395]
25. Zhao Q, Dacre M, Nguyen T, Pjanic M, Liu B, Iyer D, Cheng P, Wirka R, Kim JB, Fraser HB, et al. Molecular mechanisms of coronary disease revealed using quantitative trait loci for TCF21 binding, chromatin accessibility, and chromosomal looping. Genome Biol 2020;21:135. [PubMed: 32513244]
26. Helgadottir A, Thorleifsson G, Gretarsdottir S, Stefansson OA, Tragante V, Thorolfsdottir RB, Jonsdottir I, Bjornsson T, Steinthorsdottir V, Verweij N, et al. Genome-wide analysis yields new loci associating with aortic valve stenosis. Nat Commun 2018;9:987. [PubMed: 29511194]
27. Bahcall OG. Human genetics: GTEx pilot quantifies eQTL variation across tissues and individuals. Nat Rev Genet 2015;16:375. [PubMed: 26077371]
28. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, Torstenson ES, Shah KP, Garcia T, Edwards TL, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun 2018;9:1825. [PubMed: 29739930]
29. Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, Reddy TE, Crawford GE and Gersbach CA. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat Methods 2015;12:1143–1149. [PubMed: 26501517]
30. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, Hao Y, Stoeckius M, Smibert P and Satija R. Comprehensive Integration of Single-Cell Data. Cell 2019;177:1888–1902 e21. [PubMed: 31178118]

31. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, et al. Nature 2019;566:496–502. [PubMed: 30787437]
32. Schep AN, Wu B, Buenrostro JD and Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat Methods 2017;14:975–978. [PubMed: 28825706]
33. Cheng P, Wirka RC, Kim JB, Nguyen T, Kundu R, Zhao Q, Pedroza A, Nagao M, Iyer D, Pjanic M, et al. Smad3 regulates smooth muscle cell fate and governs adverse remodeling and calcification of atherosclerotic plaque bioRxiv Preprint posted online September 09, 2020.doi.10.1101/2020.09.15.299131
34. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM and Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 2010;28:495–501. [PubMed: 20436461]
35. Hsu YC, Li L and Fuchs E. Transit-amplifying cells orchestrate stem cell activity and tissue regeneration. Cell 2014;157:935–949. [PubMed: 24813615]
36. Newman AAC, Serbulea V, Baylis RA, Shankman LS, Bradley X, Alencar GF, Owsiany K, Deaton RA, Karnewar S, Shamsuzzaman S, et al. Multiple cell types contribute to the atherosclerotic lesion fibrous cap by PDGFRbeta and bioenergetic mechanisms. Nat Metab 2021;3:166–181. [PubMed: 33619382]
37. Avsec Z, Weilert M, Shrikumar A, Krueger S, Alexandari A, Dalal K, Fropf R, McAnany C, Gagneur J, Kundaje A et al. Base-resolution models of transcription-factor binding reveal soft motif syntax. Nat Genet 2021;53:354–366. [PubMed: 33603233]
38. Domenga V, Fardoux P, Lacombe P, Monet M, Maciazek J, Krebs LT, Klonjkowski B, Berrou E, Mericskay M, Li Z, Tournier-Lasserve E, et al. Notch3 is required for arterial identity and maturation of vascular smooth muscle cells. Genes Dev 2004;18:2730–2735. [PubMed: 15545631]
39. Davis-Knowlton J, Turner JE, Turner A, Damian-Loring S, Hagler N, Henderson T, Emery IF, Bond K, Duarte CW, Vary CPH, et al. Characterization of smooth muscle cells from human atherosclerotic lesions and their responses to Notch signaling. Lab Invest 2019;99:290–304. [PubMed: 29795127]
40. Martos-Rodriguez CJ, Albarran-Juarez J, Morales-Cano D, Caballero A, MacGrogan D, de la Pompa JL, Carramolino L and Bentzon JF. Fibrous Caps in Atherosclerosis Form by Notch-Dependent Mechanisms Common to Arterial Media Development. Arterioscler Thromb Vasc Biol 2021;41:e427–e439 [PubMed: 34261328]
41. Ward MR, Pasterkamp G, Yeung AC and Borst C. Arterial remodeling. Mechanisms and clinical implications. Circulation 2000;102:1186–1191. [PubMed: 10973850]
42. AbuRahma AF, Wulu JT Jr. and Crotty B. Carotid plaque ultrasonic heterogeneity and severity of stenosis. Stroke 2002;33:1772–1775. [PubMed: 12105350]
43. Falk E, Shah PK and Fuster V. Coronary plaque disruption. Circulation 1995;92:657–71. [PubMed: 7634481]
44. Puchner SB, Liu T, Mayrhofer T, Truong QA, Lee H, Fleg JL, Nagurney JT, Udelson JE, Hoffmann U and Ferencik M. High-risk plaque detected on coronary CT angiography predicts acute coronary syndromes independent of significant stenosis in acute chest pain: results from the ROMICAT-II trial. J Am Coll Cardiol 2014;64:684–692. [PubMed: 25125300]
45. Ferencik M, Mayrhofer T, Bittner DO, Emami H, Puchner SB, Lu MT, Meyersohn NM, Ivanov AV, Adami EC, Patel MR, et al. Use of High-Risk Coronary Atherosclerotic Plaque Detection for Risk Stratification of Patients With Stable Chest Pain: A Secondary Analysis of the PROMISE Randomized Clinical Trial. JAMA Cardiol 2018;3:144–152. [PubMed: 29322167]
46. Mintz GS. Intravascular imaging of coronary calcification and its clinical implications. JACC Cardiovasc Imaging 2015;8:461–471. [PubMed: 25882575]
47. Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, Grossfeld PD and Srivastava D. Mutations in NOTCH1 cause aortic valve disease. Nature 2005;437:270–274. [PubMed: 16025100]
48. Theodoris CV, Mourkioti F, Huang Y, Ranade SS, Liu L, Blau HM and Srivastava D. Long telomeres protect against age-dependent cardiac disease caused by NOTCH1 haploinsufficiency. J Clin Invest 2017;127:1683–1688. [PubMed: 28346225]

49. Yoshida T, Kaestner KH and Owens GK. Conditional deletion of Kruppel-like factor 4 delays downregulation of smooth muscle cell differentiation markers but accelerates neointimal formation following vascular injury. Circ Res 2008;102:1548–1557. [PubMed: 18483411]
50. Liu R, Jin Y, Tang WH, Qin L, Zhang X, Tellides G, Hwa J, Yu J and Martin KA. Ten-eleven translocation-2 (TET2) is a master regulator of smooth muscle cell plasticity. Circulation 2013;128:2047–2057. [PubMed: 24077167]
51. Misra A, Feng Z, Chandran RR, Kabir I, Rotllan N, Aryal B, Sheikh AQ, Ding L, Qin L, Fernandez-Hernando C, et al. Integrin beta3 regulates clonality and fate of smooth muscle-derived atherosclerotic plaque cells. Nature communications 2018;9:2073.
52. Gomez D, Baylis RA, Durgin BG, Newman AAC, Alencar GF, Mahan S, St Hilaire C, Muller W, Waisman A, Francis SE, et al. Interleukin-1beta has atheroprotective effects in advanced atherosclerotic lesions of mice. Nat Med 2018;24:1418–1429. [PubMed: 30038218]
53. Alexander MR, Moehle CW, Johnson JL, Yang Z, Lee JK, Jackson CL and Owens GK. Genetic inactivation of IL-1 signaling enhances atherosclerotic plaque instability and reduces outward vessel remodeling in advanced atherosclerosis in mice. J Clin Invest 2012;122:70–79. [PubMed: 22201681]
54. Christersdottir T, Pirault J, Gistera A, Bergman O, Gallina AL, Baumgartner R, Lundberg AM, Eriksson P, Yan ZQ, Paulsson-Berne G, et al. Prevention of radiotherapy-induced arterial inflammation by interleukin-1 blockade. Eur Heart J 2019;40:2495–2503. [PubMed: 31081038]
55. Ambrose JA, Tannenbaum MA, Alexopoulos D, Hjemdahl-Monsen CE, Leavy J, Weiss M, Borrico S, Gorlin R and Fuster V. Angiographic progression of coronary artery disease and the development of myocardial infarction. J Am Coll Cardiol 1988;12:56–62. [PubMed: 3379219]
56. Williams MC, Moss AJ, Dweck M, Adamson PD, Alam S, Hunter A, Shah ASV, Pawade T, Weir-McCall JR, Roditi G, et al. Coronary Artery Plaque Characteristics Associated With Adverse Outcomes in the SCOT-HEART Study. J Am Coll Cardiol 2019;73:291–301. [PubMed: 30678759]
57. Wu LM, Wang J, Conidi A, Zhao C, Wang H, Ford Z, Zhang L, Zweier C, Ayee BG, Maurel P, et al. Zeb2 recruits HDAC-NuRD to inhibit Notch and controls Schwann cell differentiation and remyelination. Nat Neurosci 2016;19:1060–1072. [PubMed: 27294509]
58. van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, Burdziak C, Moon KR, Chaffer CL, Pattabiraman D, et al. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. Cell 2018;174:716–729 e27. [PubMed: 29961576]
59. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 2017;14:959–962. [PubMed: 28846090]

Clinical Perspective

What is new?

- From human genetics, we discovered that *ZEB2*, a master regulator of epithelial to mesenchymal transition associated with many cancers, is a coronary artery disease associated gene.
- Lower *Zeb2* in SMC results in plaques with high risk features.
- Zeb2 epigenetically regulates whether plaque SMC proliferate or calcify, through global suppression of specific transcriptomic programs in conjunction with Notch and TGFβ signaling.

What are the clinical implications?

- Therapies that specifically regulate smooth muscle behavior can alter risk of plaque rupture and may be used to further reduce risk of myocardial infarction.
- Existing chemotherapies, and additional drugs in development, that modulate epigenetic silencing (i.e. HDAC inhibitors and hypomethylating agents) may increase risk of myocardial infarction.

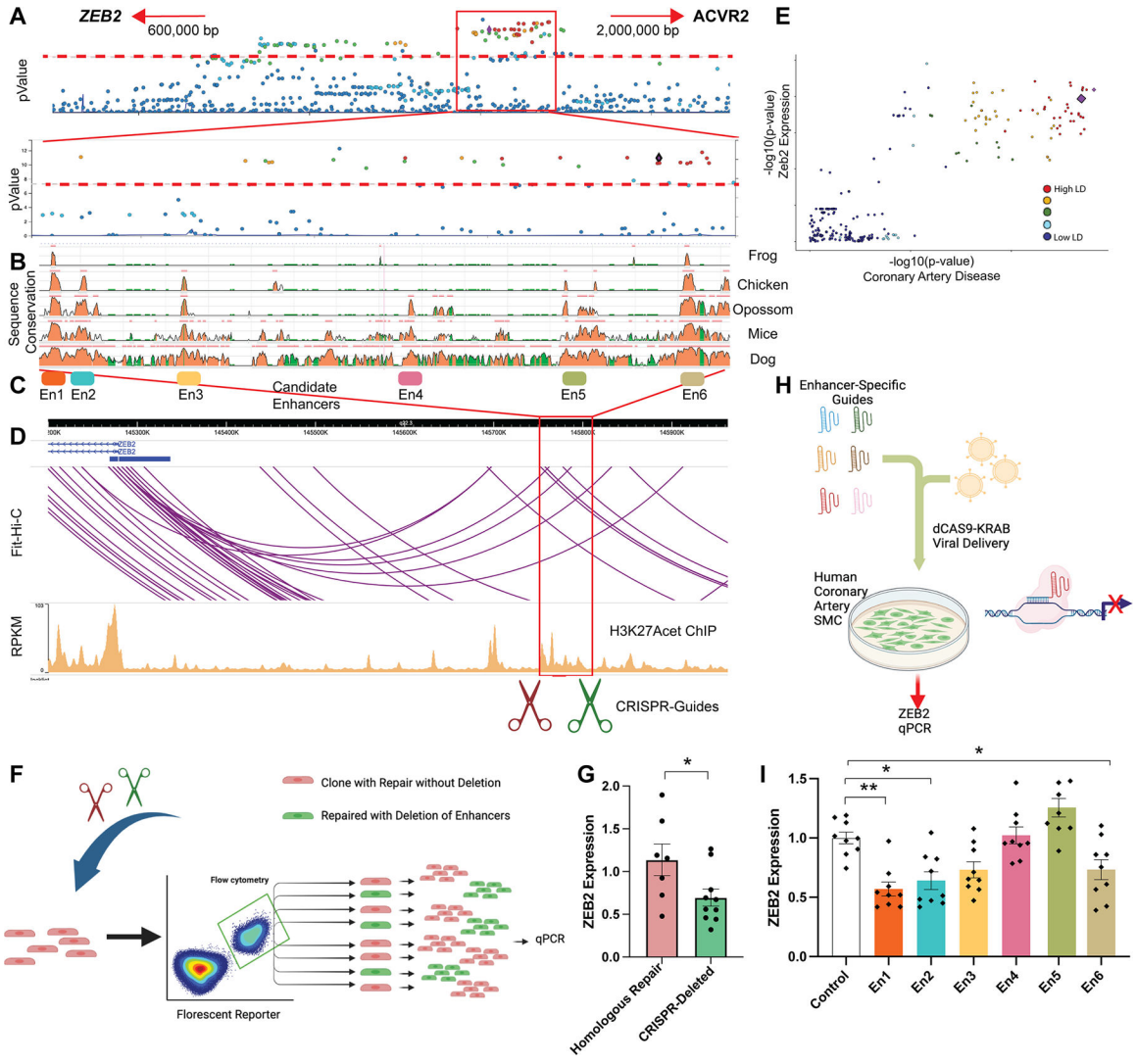

Figure 1: Human CAD GWAS signals at 2q22.3 localize to a distal enhancer for *ZEB2*.

(A) Locus-Zoom plot of CAD GWAS signal at 2q22.3 from CARDIOGRAM+C4D-UK Biobank meta-analysis [5], displaying a region with a large number of high LD-linked SNPs that span an area with high evolutionary conservation. (B) Sequence conservation among different vertebrates, Y axis represents percent sequence conservation. Data and track display from ECR browser [24]. (C) Representation of putative enhancers within this area based on sequence conservation. (D) Top: Fit-Hi-C representation of HiC data generated from HCASMC [25], demonstrating that the evolutionarily conserved regions are topologically associated with the *ZEB2* promoter. Bottom: H3K27ac Chip-Seq data generated in HCASMC, demonstrating regions of enhancer histone modification overlie regions of evolutionary conservation. (E) Co-localization of p-value of SNPs associated with altered arterial *ZEB2* expression, plotted against p-value of GWAS for CAD[20] demonstrating SNPs associated with CAD risk colocalize with those associated with altered *ZEB2* expression from human aorta GTex data as visualized using LocusCompare [4]. Lead SNP indicated here had most recent CAD GWAS pValue of 5.4×10^{-13} [5]. (F, G) CRISPR-Cas9 genome editing with clonal selection to test functionality of this evolutionarily conserved region in A549

cells, with *ZEB2* expression of clones of different genotype quantified. (H, I) Enhancer-inhibition using dCas9-KRAB along with CRISPR guides designed against enhancers from (C), with expression of *ZEB2* after transduction of CRISPRi or putative enhancer regions quantified via qPCR. *Statistical tests with $p<0.05$.

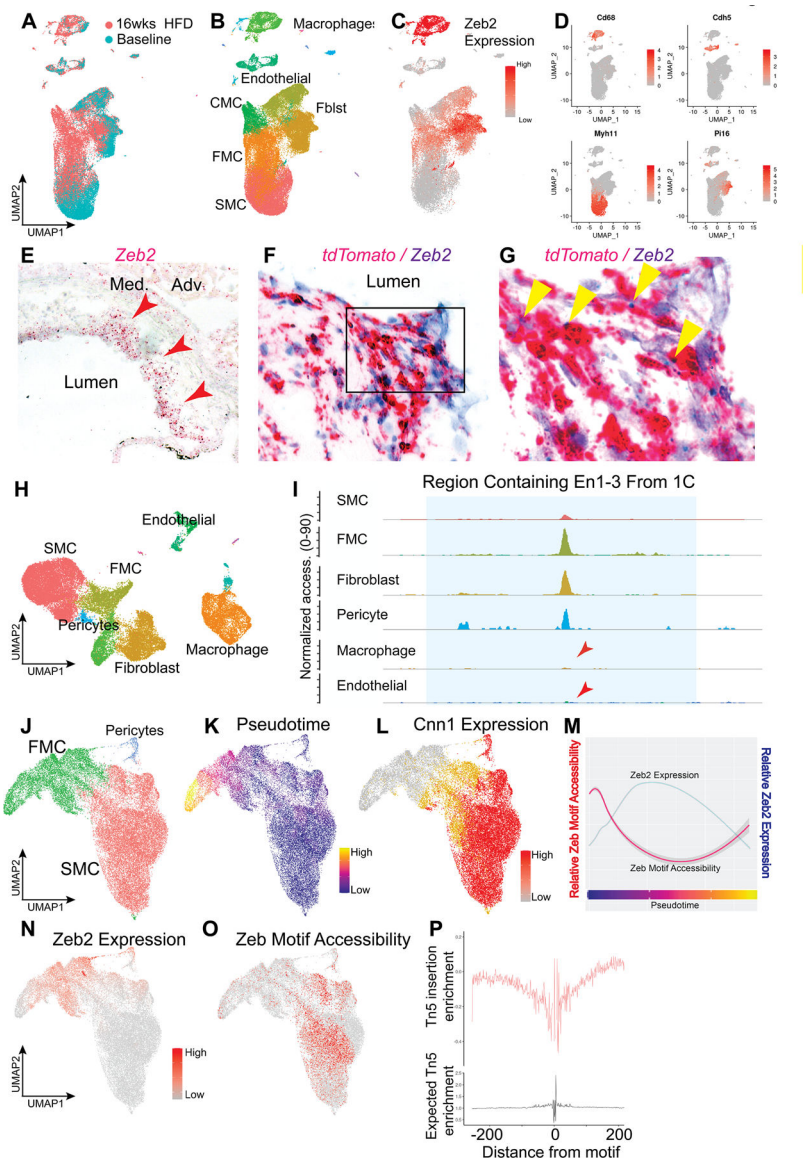

Figure 2: *Zeb2* is expressed in SMC during phenotypic transition and decreases chromatin accessibility around Zeb2 motifs.

(A) UMAP of scRNAseq data for all cells from the aortic root of *ApoE*$^{-/-}$ mice before (baseline, blue) and after 16 weeks of high-fat diet (HFD) (16 wks HFD, red). (B) Unbiased clustering of data from A into distinct cell types. (C) Imputed *Zeb2* expression in the different cell clusters in A. (D) Cell-specific cluster markers for the different populations of cells labeled in B. (E) RNA-scope of *Zeb2* (red color) in the aortic root of *ApoE*$^{-/-}$ mice after 16 weeks of high-fat diet. Arrows point to Zeb2 expressing cells near the fibrous cap. Adv, adventitia; med, media. (F, G) Two color RNA-scope of *Zeb2* (blue) and *tdTomato* (red) in *Myh11*^{CreERT2}, *ROSA*$^{tdT/+}$, *ApoE*$^{-/-}$ mice after 16 weeks of high-fat diet. Yellow arrows indicate *Zeb2* expressing lineage traced *tdTomato* positive cells. (H) UMAP of scATACseq data for all cells from the aortic root of *ApoE*$^{-/-}$ mice after 16 weeks of high-fat diet, clustered to match distinct cell types identified in B. (I) Pooled (pseudo-bulk) accessibility of different cell types in H in the region of CRISPRi validated

active conserved enhancer elements (1 and 2). (J) UMAP of combined scRNAseq and scATACseq data from all lineage-traced cells from *ApoE*$^{-/-}$ mice after 16 weeks of high-fat diet, with cells clustered into quiescent (SMC), transition (FMC) cells, and pericytes. (K) Pseudotime trajectory of SMC transition during atherosclerosis, as determined by Monocle3. (L) Gradient of *Cnn1* expression for cells from J. (M) *Zeb2* expression vs Zeb motif accessibility for cells across the pseudotime trajectory, demonstrating a negative correlation between accessibility and expression. (N) Featureplot of imputed *Zeb2* expression and (O) Zeb motif accessibility plotted on the same UMAP. (P) Normalized probability of Tn5 insertion across the genome centered around predicted Zeb motifs demonstrating negative enrichment of Tn5 insertion genome-wide +/− 100 bp around predicted Zeb motifs.

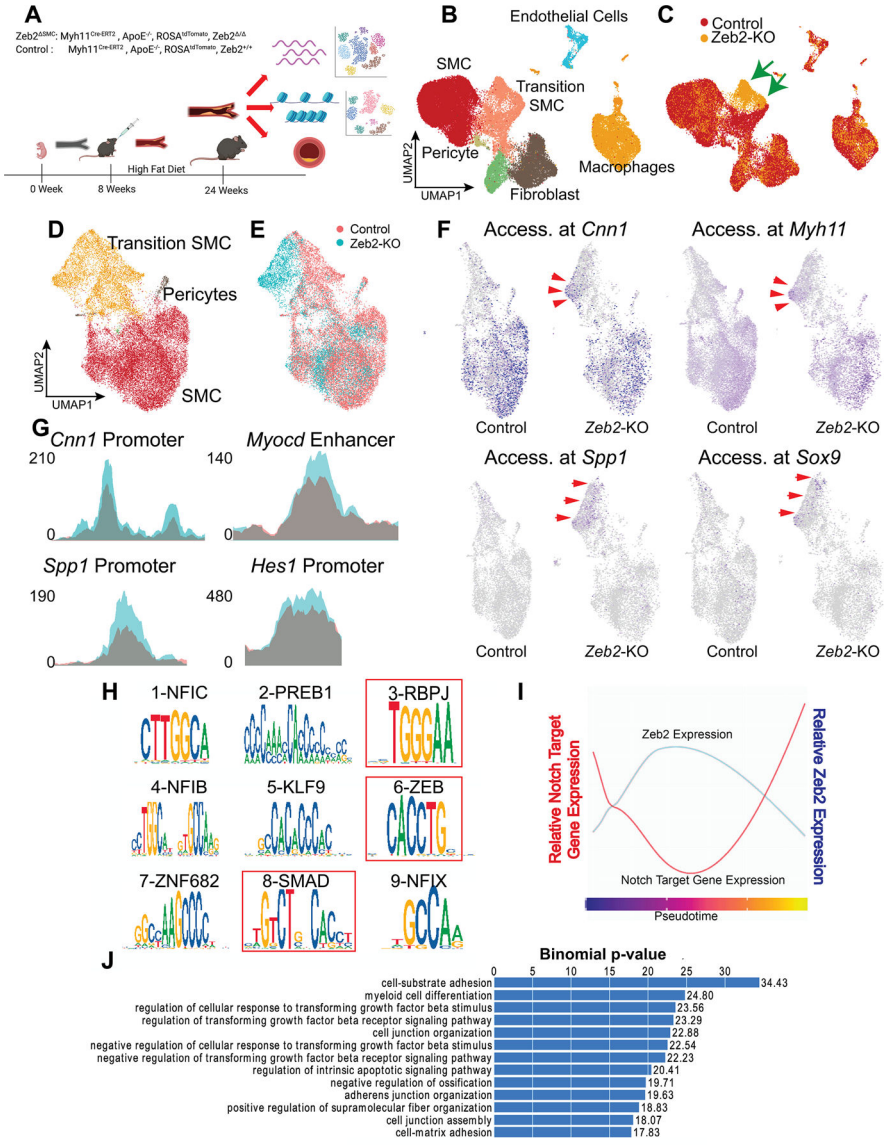

Figure 3: Loss of *Zeb2* alters the epigenetic trajectory of disease related SMC transitions.

(A) Schematic of experimental design. (B) UMAP of scATACseq dataset from all cells in control and *Zeb2*$^{\Delta SMC}$ aortic root after 16 weeks of high-fat diet, demonstrating presence of all expected different cell types. (C) Same UMAP of scATACseq data colored by control vs *Zeb2*$^{\Delta SMC}$ cells. (D) UMAP of only SMC-lineage traced cells, clustered at low resolution into quiescent and modulated SMC as well as pericytes. (E) The same UMAP grouped to demonstrate location of control and *Zeb2*$^{\Delta SMC}$ cells. (F) Featureplots of chromatin accessibility around *Cnn1, Myh11, Spp1*, and *Sox9* in control and *Zeb2*$^{\Delta SMC}$ lineage traced cells. (G) Pseudo-bulk accessibility around *Cnn1, Myocd, Spp1* and *Hes1* in control and transition SMCs. Teal color represents accessibility in *Zeb2*$^{\Delta SMC}$ cells, dark red-grey color represents accessibility in control cells. (H) Top 9 enriched motifs in differentially open peaks in *Zeb2*$^{\Delta SMC}$ transition SMCs compared to control. (I) Correlation of *Zeb2* expression with Notch target gene expression along pseudotime demonstrating a strong

negative correlation. (J) Pathway analysis of differentially accessible region comparing control and *Zeb2*$^{\Delta SMC}$ transition SMCs.

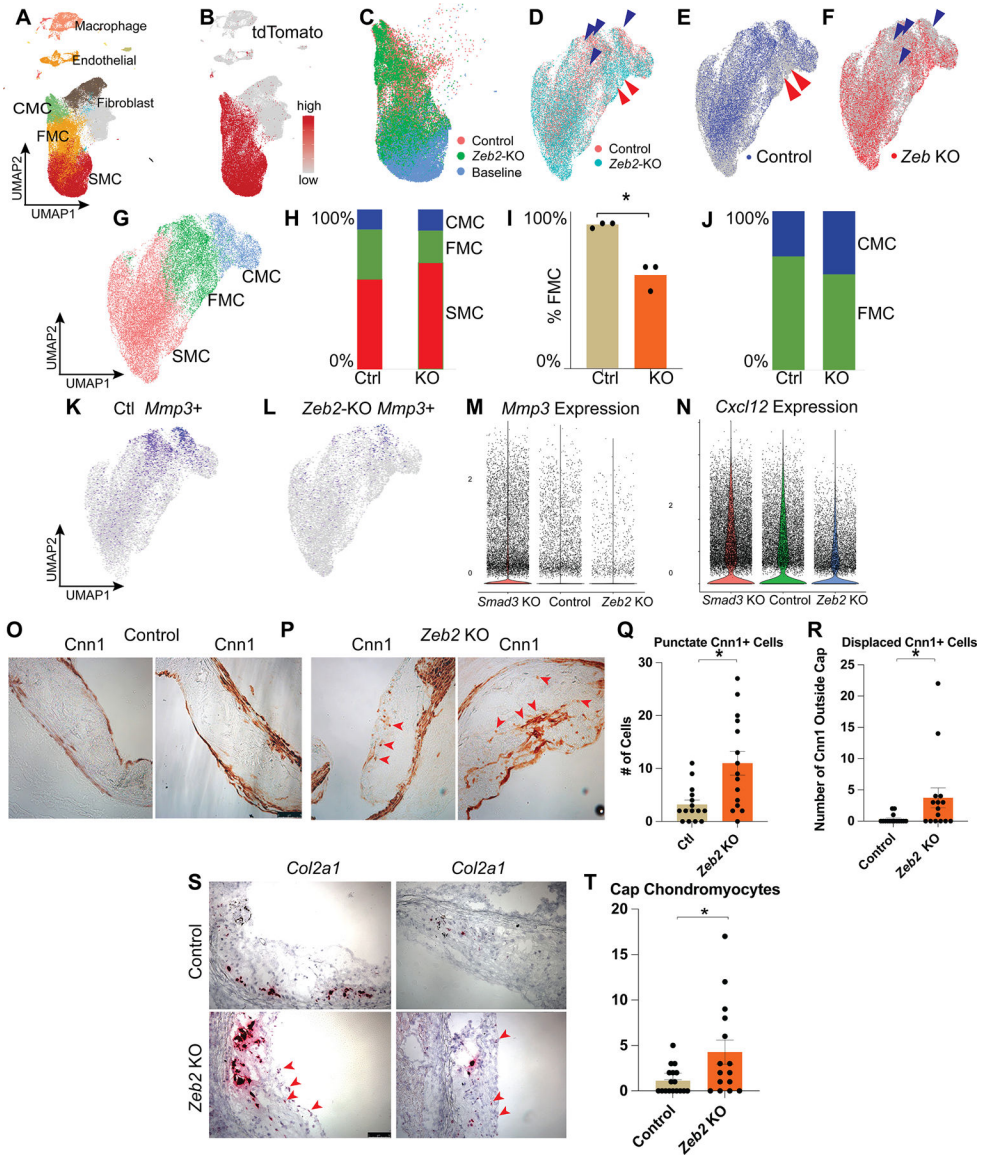

Figure 4: Loss of *Zeb2* impairs SMC phenotypic transition into FMC and accelerates chondromyocyte development.

(A) UMAP visualization of scRNAseq data of cells in the aortic root pooled from 4 control mice prior to initiation of HFD, as well as cells from 8 control and 8 $Zeb^{\Delta SMC}$ mice after 16 weeks of HFD. (B) Featureplot of *tdTomato* demonstrating populations of cells derived from smooth muscle cells. (C) SMC-derived cells grouped by time points: blue, baseline mice prior to initiation of HFD; red, control animals on HFD; green, $Zeb2^{\Delta SMC}$ animals on HFD. (D, E, F) UMAP of scRNAseq data from smooth muscle derived cells in control and $Zeb2^{\Delta SMC}$ aortic root, with principal components generated from dominant variances in this population. (D) Red; control, blue; $Zeb2^{\Delta SMC}$ cells. (E) Highlighting only control, and (F) highlighting only $Zeb^{\Delta SMC}$ cells. Red arrows indicate cell groups exclusively seen in $Zeb2^{\Delta SMC}$ cells; blue arrows indicate cell groups missing in $Zeb2^{\Delta SMC}$ mice. (G) Clustering of lineage traced cells showing SMC, FMC, and CMC subsets. (H, I) Fraction of SMC, FMC, CMC in all SMC lineage traced cells in control and $Zeb2^{\Delta SMC}$ aortic root. Each dot

represents data from separate 10X captures. (J) Fraction of FMC and CMC in all transition SMC in control and *Zeb2*$^{\Delta SMC}$ lineage-traced cells in the aortic root. (K) FeaturePlot of *Mmp3* expression in control and (L) *Zeb2*$^{\Delta SMC}$ lineage-traced cells in the aortic root. (M) Distribution of *Mmp3* and (N) *Cxcl12* expression in *Zeb2*$^{\Delta SMC}$, control, and *Zeb2*$^{\Delta SMC}$ SMC-lineage traced cells. (O) Cnn1 staining of representative lesions in the aortic root in control and (P) *Zeb2*$^{\Delta SMC}$ atherosclerotic mice. (Q, R) Quantification of ectopic and isolated *Cnn1* expressing cells in control and *Zeb2*$^{\Delta SMC}$ lesions. Each dot represents quantification from 40X images of lesions from identical level sections from individual animals, with up to three cusps quantified per section. (S) Representative *Col2a1* RNA-scope of atherosclerotic lesions in control and *Zeb2*$^{\Delta SMC}$ mice, demonstrating ectopic *Col2a1* expression near the cap, with (T) the number of ectopic cells quantified. Each dot represents quantification from 40X images of lesions from identical level sections from individual animals, with up to three cusps quantified per section. Error bars are standard error of mean, * $p<0.05$.

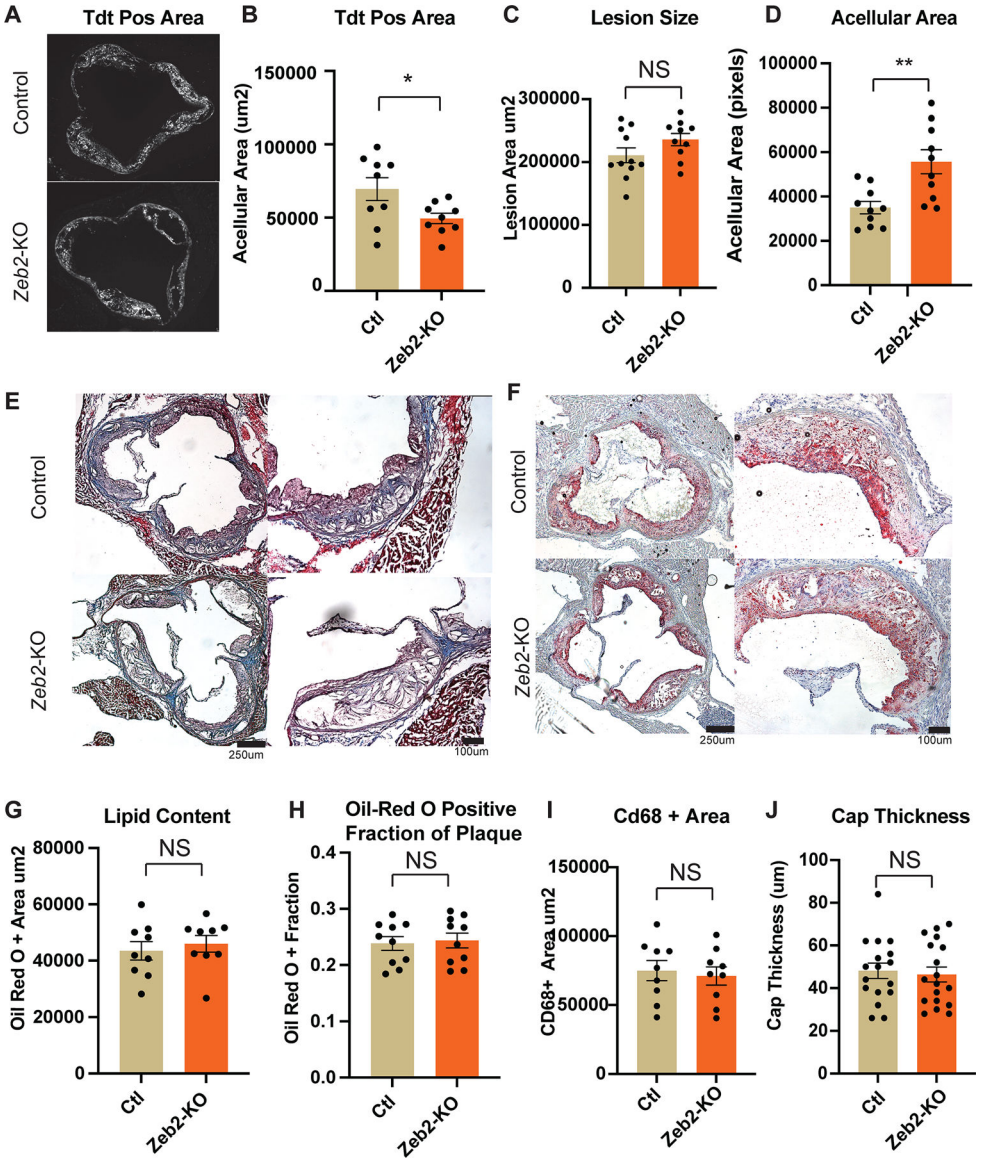

Figure 5: Loss of Zeb2 Decreases SMC contribution to lesion resulting in lesions with larger acellular area

(A) Representative image and (B) tdTomato quantified positive area in control and *Zeb2*$^{\Delta SMC}$ animals. Each dot represents identical level sections from individual animals. (C) Lesion size comparison between control and *Zeb2*$^{\Delta SMC}$ animals, as defined by total area encompassed by the internal elastic lamina. (D) Acellular plaque area comparison between control and *Zeb2*$^{\Delta SMC}$ animals. (E) Representative trichrome and (F) Oil Red O stained images of control and *Zeb2*$^{\Delta SMC}$ animals, low power (left) and high power magnification (right). (G) Oil Red O positive area comparison between control and *Zeb2*$^{\Delta SMC}$ animals. (H) Oil Red O positive fractional area relative to total plaque size in control and *Zeb2*$^{\Delta SMC}$ animals. (I) CD68 positive stained area comparison between control and *Zeb2*$^{\Delta SMC}$ animals. (J) Cap thickness comparison between control and *Zeb2*$^{\Delta SMC}$ animals, as defined by the thickest part of Tagln staining on each cusp. NS, Non-significant, ** p<0.01. Each dot

represents individual animal replicates, except for cap thickness where each dot represents a measurement from each individual cusp.

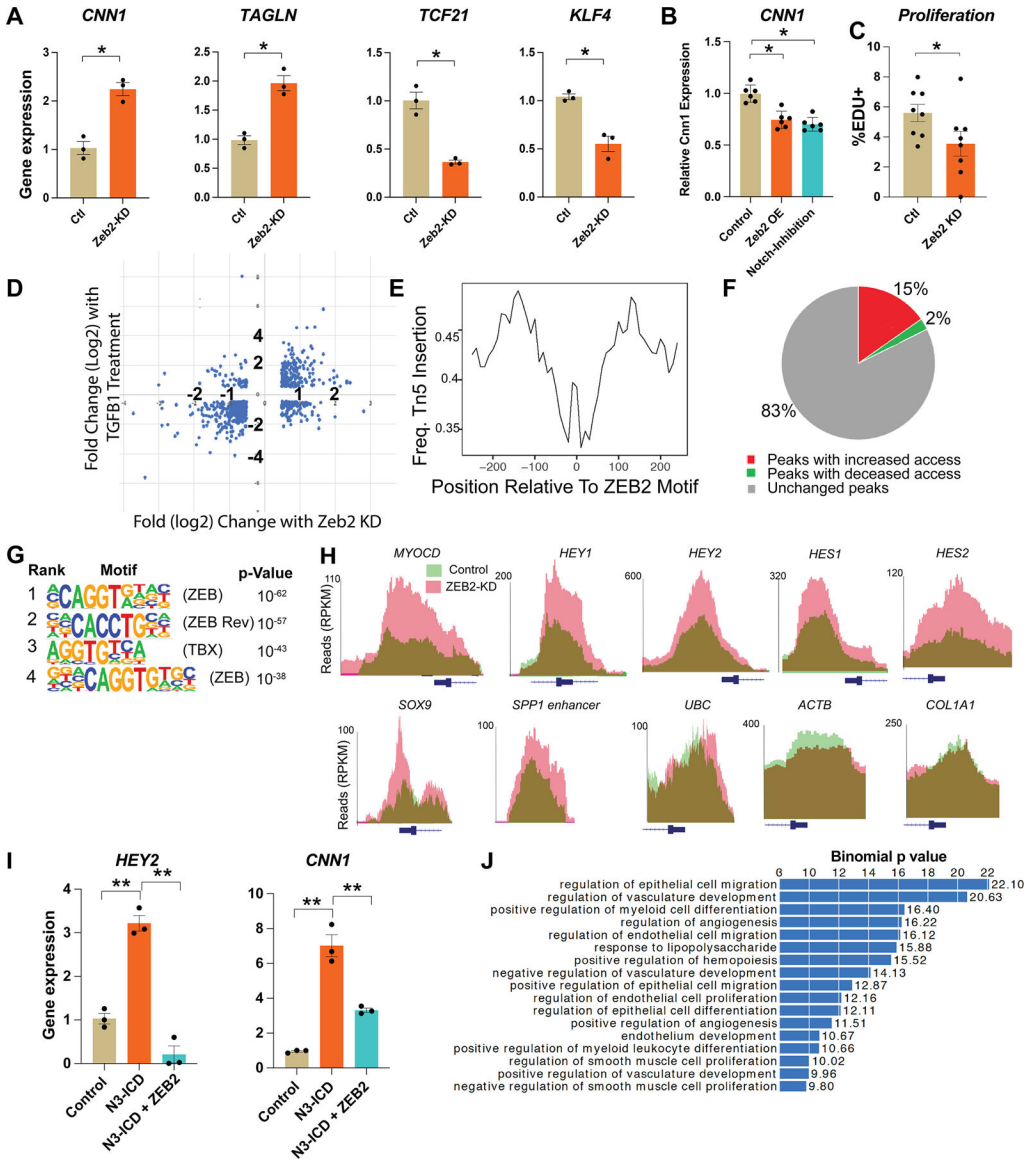

Figure 6: ZEB2 Regulates SMC cell fate and modulates NOTCH and TGFβ Signaling in HCASMC.

(A) qPCR of mature SMC and transition SMC markers in HCASMC in control vs *ZEB2*-KD. (B) *CNN1* expression with *ZEB2* overexpression and Notch inhibition. (C) Percent of EDU labeled cells in control vs *ZEB2*-KD HCASMC. (D) Plot of fold-changes in highly differentially regulated genes with TGFβ stimulation vs *ZEB2*-KD (E) Relative probability of Tn5 insertion relative to ZEB2 motif locations genome wide in HCASMC. (F) Proportion of peaks identified by scATACseq that were more open (red), closed (green), or unchanged (grey) upon knockdown of *ZEB2*. (G) Top 4 motifs enriched in differentially open peaks. ZEB Rev indicates Reverse complement of the ZEB motif on opposite DNA strand. (H) Relative chromatin accessibility around promoters and enhancers of SMC lineage determining factor *MYOCD*, NOTCH responsive genes *HES1, HES2, HEY1, HEY2,* and chondromyocyte regulators *SPP1* and *SOX9*, or unrelated control genes *UBC, ACTB, COL1A1* in control (green) or *ZEB2*-KD (salmon), overlap appears brown. (I) Relative

expression of *HEY2* and *CNN1* in HCASMC infected with viruses over-expressing *N3ICD* and *ZEB2*. (J) Biological processes enriched in differentially accessible peaks comparing *ZEB2*-KD vs control, as analyzed by GREAT.

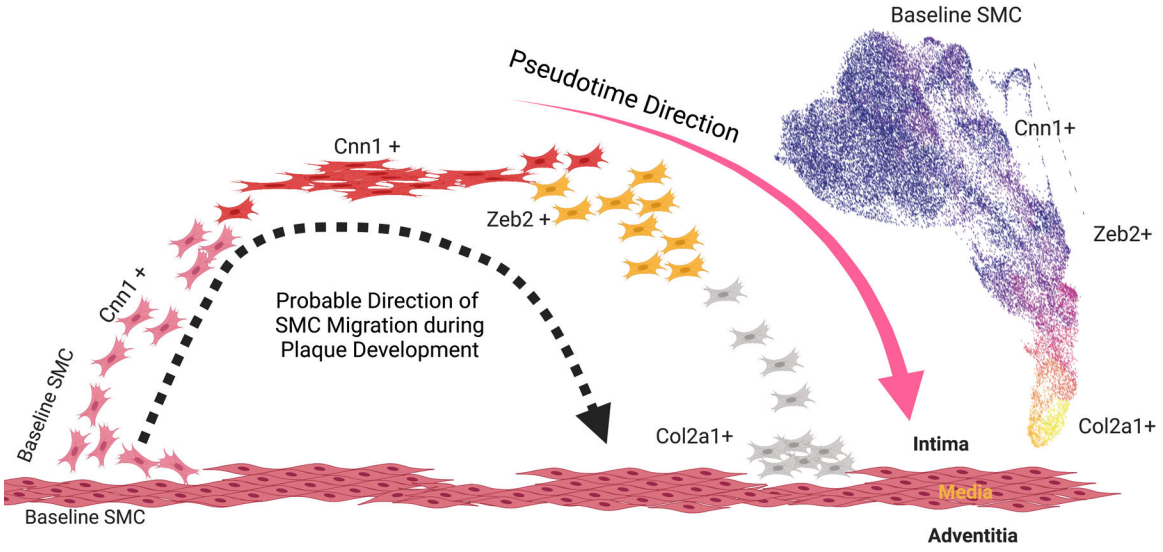

Figure 7. Model for SMC transition along pseudo-time and spatial trajectories in atherosclerotic lesions.

Overlay of plaque anatomy, *Zeb2* expression, and SMC pseudotime trajectory suggest cap cells may be the source of plaque SMC-derived cells. The *Cnn1* positive cap cells that exhibit *Zeb2* expression undergo a migratory event into the plaque, and eventually turn off *Zeb2* expression and initiate expression of chondrocyte genes including *Col2a1* at the base of the plaque. These events are consistent with the order of expression of *Cnn1*, *Zeb2*, and *Col2a1* from the cap to the base of the lesion, and along the SMC pseudotime as shown in experiments reported here.