Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis

Robert C. Wirka1,2, Dhananjay Wagh3, David T. Paik4,1,2, Milos Pjanić1,2, Trieu Nguyen1,2, Clint L. Miller4, Ramen Kundu1,2, Manabu Nagao1,2, John Coller3, Tiffany K. Koyano5, Robyn Fong5, Y. Joseph Woo5, Boxiang Liu6, Stephen B. Montgomery6, Joseph C. Wu1,2, Kuixi Zhu7,8, Rui Chang7,8, Melissa Alamprese7,8, Michelle D. Tallquist9, Juyong B. Kim1,2,10 and Thomas Quertermous1,2,10*

In response to various stimuli, vascular smooth muscle cells (SMCs) can de-differentiate, proliferate and migrate in a process known as phenotypic modulation. However, the phenotype of modulated SMCs in vivo during atherosclerosis and the influence of this process on coronary artery disease (CAD) risk have not been clearly established. Using single-cell RNA sequencing, we comprehensively characterized the transcriptomic phenotype of modulated SMCs in vivo in atherosclerotic lesions of both mouse and human arteries and found that these cells transform into unique fibroblast-like cells, termed ‘fibromyocytes’, rather than into a classical macrophage phenotype. SMC-specific knockout of TCF21—a causal CAD gene—markedly inhibited SMC phenotypic modulation in mice, leading to the presence of fewer fibromyocytes within lesions as well as within the protective fibrous cap of the lesions. Moreover, TCF21 expression was strongly associated with SMC phenotypic modulation in diseased human coronary arteries, and higher levels of TCF21 expression were associated with decreased CAD risk in human CAD-relevant tissues. These results establish a protective role for both TCF21 and SMC phenotypic modulation in this disease.

The most significant consequence of coronary artery disease (CAD) occurs when an ‘unstable’ atherosclerotic lesion ruptures and triggers an occlusive thrombus, resulting in a myocardial infarction. Compared with stable coronary lesions, these vulnerable plaques are characterized by a large necrotic lipid core and a thin overlying fibrous cap that is prone to rupture. During atherosclerosis, smooth muscle cells (SMCs) from the vessel wall likely contribute to both the fibrous cap and the underlying necrotic core via a process known as ‘phenotypic modulation’, in which SMCs de-differentiate, proliferate and migrate in response to atherogenic stimuli. The current view is that phenotypically modulated SMCs can develop into one of two distinct phenotypes, depending on environmental cues, with very different potential consequences for plaque stability: (1) pro-inflammatory, dysfunctional macrophage-like cells, characterized in vivo by the upregulation of the macrophage marker Lgals3 (ref. 4), which may serve to destabilize the lesion; or (2) extracellular matrix-producing ‘synthetic’ SMCs that may contribute to the protective fibrous cap, which would serve to prevent plaque rupture and myocardial infarction. Despite the significant uncertainty regarding the phenotype of modulated SMCs, it has become increasingly clear that these cells are an important component of the developing plaque in animal models of atherosclerosis.

A recent lineage-tracing study of SMCs in the mouse aortic root revealed that phenotypically modulated SMCs contribute ~30% of all cells in the atherosclerotic plaque. There is some evidence that SMC phenotypic modulation occurs in human atherosclerosis, but the phenotype of these cells and their contribution to human disease remains to be elucidated. TCF21—a basic helix–loop–helix transcription factor—is the causal gene at the CAD-associated locus at 6q23.2 (refs. 8–10). In murine cardiac development, Tcf21 is expressed in proepicardial cells that give rise to both cardiac fibroblasts and coronary artery SMCs11,12. In this context, Tcf21 is required for cardiac fibroblast development but is downregulated in cells that eventually become coronary artery SMCs11, suggesting that sustained Tcf21 expression shifts these precursor cells away from the SMC lineage. Similarly, in cultured human coronary artery SMCs (HCASMCs), TCF21 knockdown results in upregulation of SMC differentiation markers. In adult mice, Tcf21 is primarily expressed in the adventitia surrounding the coronary arteries and the aortic root, and also sporadically in some cells in the medial layer of the aortic root. During development of atherosclerotic disease in the aortic root of ApoE−/− mice, there is robust expression of Tcf21 in many cells within the lesion. However, several fundamental questions remained: (1) what cell type(s) express Tcf21 during lesion development; (2) how does Tcf21...
Articles
Nature Medicine

Fig. 1 | Transcriptomic characterization of mouse aortic root atherosclerotic plaques and Tcf21 expression. a–c, t-SNE visualization of cell types present in the mouse aortic root at baseline (a; n = 3 mice), after 8 weeks of HFD (b; n = 3 mice) and after 16 weeks of HFD (c; n = 3 mice), illustrating the appearance of a disease-specific cell type, the ‘modulated SMC’ cluster. All cell cluster identities are indicated. d, SMC lineage-traced cells, identified by their expression of the tdT reporter gene via FACS, are labeled in red for all time points. tdT+, cells expressing tdT; tdT−, cells not expressing tdT. e, The top eight genes defining each type of cell cluster in a–c are listed. Gray bars are added to aid interpretation. The size of each circle represents the fraction of cells in each cluster that express at least one detected transcript of each gene. The color scale indicates the expression level (blue, low; red, high). f, Percentage of cells of each cell type that contained detectable (non-zero) Tcf21 levels at baseline, 8 weeks and 16 weeks of disease.

affect the phenotype of these cells; and (3) how does Tcf21 affect disease risk?

Results
Single-cell RNA sequencing (scRNA-Seq) defines cellular composition of mouse atherosclerosis and reveals that Tcf21 expression is upregulated during SMC phenotypic modulation. To lineage trace vascular SMCs, we used bacterial artificial chromosome (BAC) transgenic mice expressing a tamoxifen-inducible Cre recombinase locus (ROSAtdT/+; CreERT2) [15,16] as well as a Cre-responsive reporter gene (tandem dimer Tomato reporter [tdT]) inserted at the ROSA26 locus (ROSAtdT+/+) [17] on the ApoE−/− background (SMC lineage-tracing (SMClin) mice; Extended Data Fig. 1a, top panel). When these mice were administered tamoxifen at 8 weeks of age, before a high-fat diet (HFD) and disease onset, all SMCs (and any progeny resulting from subsequent proliferation) were permanently labeled with tdT fluorescence. To examine single-cell gene expression, we used fluorescence-activated cell sorting (FACS) to sort cells isolated from the aortic root and ascending aorta into two groups: tdT+ (SMC origin) and tdT− (all other cell types; Extended Data Fig. 1c) at baseline and after 8 and 16 weeks of HFD (Extended Data Fig. 1b). We then performed scRNA-Seq on both groups of cells in parallel using the 10x Chromium platform [18], and the estimated recombination efficiency in SMCs was 98.9% (see Methods). Figure 1 illustrates all of the major cell types identified by their gene expression profiles in the aortic root and ascending aorta at baseline before HFD (Fig. 1a), after 8 weeks of HFD (Fig. 1b) and after 16 weeks of HFD (Fig. 1c). The top cell type-specific genes are shown in Fig. 1e. Interestingly, the main SMC population was divided into two groups (SMC1 and SMC2), based on significant differences in gene expression patterns (Supplementary Tables 1–3).

With the development of atherosclerosis, the most notable change was the appearance of a distinct group of cells that were juxtaposed in t-stochastic neighbor embedding (t-SNE) space to the contractile SMC clusters (Fig. 1b,c, red), increasing in prevalence from 8 to 16 weeks of disease. Visualization of all of the cells that had been FACS sorted as tdT+ and were thus of SMC lineage revealed that the vast majority of cells in this disease-associated cell cluster were SMC derived (Fig. 1d) and therefore represented phenotypically modulated SMCs. A total of 11% of cells within the disease-associated cluster were lineage negative, suggesting an origin other than SMC. Interestingly, at the baseline time point, a small proportion of SMCs (1.3%) were already classified as phenotypically modulated SMCs, consistent with recently published data showing a population of Sca1+ SMCs in healthy mice [19].

To characterize Tcf21 expression in the different vascular cell lineages, we measured the percentage of cells from each group that expressed detectable levels of Tcf21 (Fig. 1f). Fibroblasts had the highest percentage of Tcf21+ cells, but exhibited a decrease with disease progression. Other cell populations had only small changes in Tcf21 levels during disease. An exception to this was found in the modulated SMC group—although derived from baseline SMCs that contained only 3.1% Tcf21+ cells, the modulated SMC cluster markedly upregulated Tcf21 (29% Tcf21+ cells) by 8 weeks of disease. The percentage
of Tcf21+ cells subsequently declined to 9.8% by 16 weeks of disease and did not return to the baseline levels seen in the contractile SMC clusters. Thus, the increase in Tcf21 expression within the lesion during disease was the result of a prominent upregulation of Tcf21 specifically in SMCs during phenotypic modulation.

**SMC phenotypic modulation in vivo during disease results in a specific fibroblast-like phenotype.** Data from SMC<sup>Cre</sup> mice at all time points are combined in Fig. 2a–g. In the phenotypically modulated SMC cluster, markers of SMC differentiation including transglisin (Tagln; Fig. 2b) and calponin (Ctnl1; Fig. 2c) showed a gradient of decreasing expression from the parental SMC lineage, and a gradient of increased expression for Lgals3 (a known marker of SMC phenotypic modulation; Fig. 2d), suggesting that these cells were undergoing SMC phenotypic modulation. There was marked upregulation of many other genes, including fibronectin 1 (Fnn1; Fig. 2e), osteoprotegerin (Tnfrsf11b; Fig. 2f) and collagen 1α1 (Col1a1; Extended Data Fig. 2a) in this cell group. In particular, during phenotypic modulation, there was striking upregulation of small leucine-rich proteoglycans such as lumican (Lum; Fig. 2k), decorin (Den) and biglycan (Bgn)—genes that are otherwise specific to the fibroblast and fibroblast 2 populations. We then used a gene expression score, derived from the top differentially expressed genes distinguishing the modulated SMC cluster from the contractile SMC cluster (see Methods), to quantify the degree of SMC phenotypic modulation for each cell. Although the modulated SMC cluster was readily distinguished from contractile SMCs during clustering, we observed a progressive increase in the gene expression score in those cells more distant from the contractile SMC clusters (Fig. 2g). This transcriptional shift away from contractile SMCs also progressed over time (Fig. 2h). Phenotypically modulated SMCs have previously been noted to express markers associated with myofibroblasts (Acta2), mesenchymal stem cells (Scal) and macrophages (Lgals3), and this information was employed to suggest evidence of transdifferentiation into these cellular lineages. However, we found that the expression of these markers appears instead to reflect progression along a single path of phenotypic modulation, and not separation into distinct lineages (Extended Data Fig. 2b–d). To further investigate this possibility, we asked whether these phenotypically modulated SMCs were becoming more transcriptionally similar to other cell types within the lesion. We calculated the Euclidean distance between the centroids of all cell groups in 20-dimensional principal component space and determined, with contractile SMCs as a reference point, how the phenotypically modulated SMCs had shifted in relation to each cell cluster. This analysis revealed that during SMC phenotypic modulation, the transcriptional signature of these cells clearly shifts towards that of the fibroblast clusters (Fig. 2i). The relationship between all cell types is illustrated in Extended Data Fig. 2m. Molecular pathway analysis (ingenuity pathway analysis (IPA)) performed on differentially regulated genes between the phenotypically modulated SMCs and contractile SMCs also showed strong upregulation of genes expressed by fibroblasts (for example, hepatic stellate cell activation) and downregulation of integrin and integrin-linked kinase signaling pathways (Fig. 2j). Specifically, integrin beta-1 and integrin alpha-8 were downregulated, which comprise an integrin heterodimer restricted to SMCs<sup>Cre</sup>. Given the clear transition to a fibroblast-like phenotype, and their separation from other cell clusters, we termed the modulated SMCs 'fibromyocytes', reflecting their origin from smooth muscle myocytes and their adoption of fibroblast phenotype.

To localize these fibromyocytes within the plaque, we first searched for genes that were specific for the fibromyocyte cluster in the scRNA-Seq dataset. We found that the gene Lum, when coexpressed with the tdT lineage marker, captured the majority of fibromyocytes with excellent (96%) specificity (Fig. 2k). We performed RNAscope in situ hybridization for Lum and tdT genes in sections of the atherosclerotic mouse aortic root. This revealed that (Lum+tdT+) cells were found throughout the lesion and also at the fibrous cap (Fig. 2l–n), consistent with the notion that these cells migrate into the lesion during disease. There were also Lum<sup>-</sup>/tdT<sup>+</sup> cells in the media underlying the lesion, but these cells were not present in the media of non-diseased areas of the vessel (Extended Data Fig. 2g), suggesting that medial SMCs begin to undergo phenotypic modulation before entering the plaque. We also performed RNAscope for osteopontin (Spp1), the expression of which is limited to more 'extreme' modulated SMCs that also express multiple chondrocyte markers. These cells localized only to the lesion and not the media (Extended Data Fig. 2j–l), indicating that the Lum<sup>-</sup>/tdT<sup>+</sup> cells in the media are likely at an earlier stage of phenotypic modulation.

A notable observation from these analyses is that, at a transcriptional level, SMCs undergoing phenotypic modulation do not appear to be shifting towards a monocyte-derived macrophage phenotype within the plaque. Despite their shared expression of the macrophage marker Lgals3, fibromyocytes lack significant expression of virtually all of the other top markers that distinguish the macrophage cell cluster (Figs. 1e and 2o). Indeed, whole-transcriptome analyses revealed that, compared with contractile SMCs, fibromyocytes are actually more distant from macrophages, suggesting that these cells are becoming less similar to monocyte-derived macrophages in the mouse (Fig. 2i). We then sought to determine whether...
modulated SMCs express macrophage markers at the protein level using multiple techniques. We performed immunostaining for the macrophage marker Cd68 in mouse lesions, which did not identify significant Cd68 expression in tdT+ cells (Fig. 2p). We then incubated a single-cell suspension from the atherosclerotic aortic root and ascending aorta of SMClin mice with antibodies
against the macrophage markers Cd16 and Cd32 and performed flow cytometric analysis, which revealed that SMC-derived tdT+ cells do not express significant levels of Cd16 or Cd32 (Extended Data Fig. 2n). Finally, to integrate our transcriptional findings with protein-level data, we incubated a single-cell suspension from the atherosclerotic aortic root and ascending aorta of two SMCin mice with a panel of six DNA-barcoded antibodies against commonly used macrophage markers (Cd16, Cd32, Cd11b, Cd64, Cd86 and F4/80) before performing scRNA-Seq (cellular indexing of transcriptomes and epitopes by sequencing (CITE-Seq))\textsuperscript{21}; Extended Data Fig. 2o–r). Antibody binding was then assessed by recovering the antibody-associated DNA barcodes in the complementary DNA (cDNA) library. This experiment confirmed that these macrophage markers are not upregulated in modulated SMCs compared with quiescent SMCs (Extended Data Fig. 2p–q), consistent with the transcriptomic data. We then assessed lipid uptake by SMC-derived cells within the lesion using the boron-dipyrromethene (BODIPY) neutral lipid stain. Consistent with previous reports\textsuperscript{22}, we found that many modulated SMCs in the lesion do indeed contain lipid droplets (Fig. 2q). However, these were quantitatively and qualitatively distinct from the large macrophage-derived foam cells in the lesion (Fig. 2r). Taken together, these data suggest that although modulated SMCs in the plaque take up lipid, they do so without adopting a macrophage-like transcriptional phenotype.

Loss of Tcf21 in SMCs inhibits phenotypic modulation. To determine the effect of Tcf21 on SMC phenotypic modulation, we performed scRNA-Seq in the aortic root and ascending aorta of SMC-specific conditional Tcf21 knockout mice (SMCin-KO). These mice were identical to the SMCin mice, except at the Tcf21 locus where both Tcf21 alleles were flanked with LoxP sites (Tcf21\textsuperscript{fl/fl};\textsuperscript{cDNA/ADAC}) (Extended Data Fig. 1a, bottom panel). Thus, when tamoxifen was administered to these SMCin-KO mice before HFMD and disease onset, in addition to lineage marker activation, the Tcf21 gene was permanently deleted in all SMCs and any progeny resulting from subsequent proliferation. As a control, we used SMCin mice. We first assessed the efficacy of Tcf21 deletion in the scRNA-Seq data and found that Tcf21 expression was reduced by 95% in SMC-derived cells from SMCin-KO mice relative to controls (Extended Data Fig. 3a). We measured the proportions of contractile SMCs and fibromyocytes in the scRNA-Seq data at the 16-week disease time point and found that, compared with SMCs from SMCin controls (Fig. 3a), SMCs from SMCin-KO mice exhibited a marked reduction in the ability to undergo phenotypic modulation (Fig. 3b). This reduction in SMC modulation was observed at both 8 weeks (8% in the wild type (WT) versus 1% in the knockout (KO)) and 16 weeks of disease (Fig. 3c; 48% in the WT versus 16% in the KO; chi-squared $P = 2.2 \times 10^{-16}$). Plaque characteristics in the aortic root in a large cohort of mice ($n = 17$ WT; 22 KO) supported the scRNA-Seq findings. At the 16-week disease time point, SMCin-KO mice exhibited a decrease in the proportion of lineage-traced tdT+ cells in the lesion relative to controls (Fig. 3d; $P = 0.01$), despite a similar tdT+ area within the whole vessel wall (Fig. 3g). Importantly, in the SMCin-KO mice, there was also a lower proportion of tdT+ SMC lineage-traced cells in the area of the fibrous cap (Fig. 3e,f; $P = 0.003$). The SMCin-KO mice also exhibited a reduction in the tdT+Lgals3+ area within the lesion (Fig. 3h; $P = 0.001$), more specifically showing fewer modulated SMCs within the lesion. The total Lgals3+ area within the lesion was also reduced (Fig. 3i and Extended Data Fig. 3b). In contrast, staining for the contractile SMC marker Tagln was increased (Fig. 3k,l; $P = 0.008$). The increase in Tagln area within the SMCin-KO group corresponded to an increased medial area in these mice (Fig. 3j; $P = 0.01$). As Lgals3 is also expressed in monocye-derived macrophages, we stained for the macrophage-specific marker Cd68 to further exclude the possibility that changes in Lgals3 staining were caused by differences in macrophage content between the two groups. Indeed, there was no difference in Cd68 staining between SMCin and SMCin-KO mice ($P = 0.34$; Extended Data Fig. 3c). There was also no significant difference in lesion area between the two groups (Extended Data Fig. 3d). Together, these findings strongly suggest that loss of Tcf21 results in inhibition of SMC modulation and fewer fibromyocytes in the lesion and fibrous cap.

Identification and characterization of modulated SMCs in human coronary arteries reveals a similar fibromyocyte phenotype. To determine whether our findings in the mouse could be observed in humans, we performed scRNA-Seq in dissociated cells from human atherosclerotic coronary arteries. Diseased segments within the right coronary artery of four cardiac transplant recipients (Supplementary Table 4) were dissociated and subjected to scRNA-Seq. Diseased segments ranged from noncalcified plaques to more advanced lesions with areas of calcification, and stented areas were excluded. Cell types were assigned to each cluster based on the top defining genes in each cluster (Fig. 4a and Supplementary Table 5). Two contiguous clusters, labeled ‘SMC’ and ‘fibromyocyte’ (Fig. 4a), were characterized by diminishing CNN1 expression (Fig. 4b), which appeared to parallel the gradual loss of CNN1 expression in murine fibromyocytes (Fig. 2c). The decrease in expression of CNN1 and other markers of SMC differentiation in these cell groups was accompanied by a corresponding increase in markers of SMC modulation, including FN1 (Fig. 4c) and LUM (Extended Data Fig. 4a), suggesting that these two clusters could represent SMCs undergoing phenotypic modulation to become fibromyocytes. We then sought to more definitively identify fibromyocytes in the human coronary artery based on whole-transcriptome similarity to bona fide lineage-traced fibromyocytes in the mouse. To this end, we combined the mouse and human datasets and, using the aligned canonical correlation analysis feature of the Seurat package, performed joint clustering on the combined mouse and human dataset (Extended Data Fig. 5). We found that this approach accurately clustered together known orthologous cell types from each species (Extended Data Fig. 5e). We then identified the human cells that had clustered with the bona fide, lineage-traced fibromyocytes in the mouse (Extended Data Fig. 5b, red cluster) and highlighted these cells within the context of the human single-cell dataset (Fig. 4d and Extended Data Fig. 5d). We found that these cells mapped back to the same region undergoing loss of SMC markers and upregulation of fibromyocyte markers. We found that 86% of these cells mapped back to the ‘fibromyocyte’ cluster in Fig. 4a and 8% mapped to the immediately adjacent ‘fibroblast 2’ cluster, together accounting for 94% of all cells classified as fibromyocytes via joint clustering. The unbiased, whole-transcriptome similarity to mouse fibromyocytes and their independent compact clustering in the human dataset strongly suggest that these cells are indeed human fibromyocytes.

To localize these fibromyocytes in their anatomical context within the human coronary artery lesion, we first searched for markers that were highly specific for human fibromyocytes (Fig. 4d, brown) in the human scRNA-Seq dataset. We found that the gene TNFRSF11B, encoding osteoprotegerin, exhibited 97% specificity and 53% sensitivity for human fibromyocytes in the scRNA-Seq dataset (Fig. 4e). We then performed RNAseq in situ hybridization to visualize the distribution of TNFRSF11B within the human lesions. This revealed strong TNFRSF11B staining primarily in the fibrous neointima (Fig. 4f and Extended Data Fig. 4b), with few cells strongly positive for TNFRSF11B in the media or the adventitia, consistent with the expected location of fibromyocytes. No staining was observed using a negative control probe (Extended Data Fig. 4c).

Interestingly, using t-SNE visualization, the ‘fibromyocyte’ cluster appeared to be continuous with the fibroblast 2 population. Calculating Euclidean distance in 20-dimensional principal
component space between cell cluster centroids in the human dataset confirmed that the ‘fibromyocyte’ cluster develops striking similarity to the fibroblast 2 cluster (Extended Data Fig. 4d). Both the ‘fibromyocyte’ and the ‘SMC’ cluster remained significantly dissimilar to monocyte-derived macrophages (Extended Data Fig. 4d). Thus, consistent with the mouse data, SMCs undergoing phenotypic modulation in the human artery also appear to be acquiring a specific fibroblast-like ‘fibromyocyte’ phenotype, characterized by strong upregulation of collagen, fibronectin 1, lumican and secreted proteoglycans. Of note, consistent with previous reports6,23,

**Fig. 3 | SMC-specific Tcf21 knockout markedly inhibits SMC phenotypic modulation in mice.**  
*a,b*, Prevalence of contractile SMCs (blue) and fibromyocytes (red) at 16 weeks of disease in SMC<sup>tn</sup> (<a>n</a> = 3 mice) and SMC<sup>tn-KO</sup> (<b>n</b> = 3 mice).  
*c*, Proportions of contractile (blue) and modulated (red) SMCs after 16 weeks of disease in SMC<sup>tn</sup> and SMC<sup>tn-KO</sup> mice (<i>n</i> = 3 mice for each genotype; chi-squared <i>P</i> = 2.2 × 10<sup>−16</sup>).  
*d,e*, Percentage of tdT-positive staining area in the lesion (<d>) and in the fibrous cap (<e>), defined as the area of the lesion within 30 μm of the luminal surface.  
*f*, Representative images of tdT<sup>+</sup> cells in SMC<sup>tn</sup> and SMC<sup>tn-KO</sup> mice. FCA, fibrous cap area.  
*g*, Total tdT content of the vessel.  
*h*, tdT<sup>+</sup>/Lgals3<sup>+</sup> area in the lesion.  
<i>, Representative images of Lgals3 staining in the lesions of SMC<sup>tn</sup> and SMC<sup>tn-KO</sup> mice.  
*j–l*, Medial size (<j>) and Tagln content (<k>) in SMC<sup>tn</sup> and SMC<sup>tn-KO</sup> mice, with representative images of Tagln staining shown in <l>. All data in <d–l> were at 16 weeks of disease. Scale bars in <f>, i and <l> represent 100 μm. Data in <d>, e, g, h, j and k were analyzed by two-sided Student’s <i>t</i>-test. Error bars denote s.e.
we observed upregulation of the macrophage marker CD68 in human fibromyocytes relative to contractile SMCs (Extended Data Fig. 4e), but CD68 was also expressed in multiple non-macrophage cell types within the lesion, suggesting that it does not reflect the acquisition of macrophage-like properties.

**TCF21 is associated with SMC phenotypic modulation in human coronary arteries.** Given the marked effect of Tcf21 on SMC modulation in the mouse model of atherosclerosis, we sought to determine whether TCF21 is also associated with SMC phenotypic modulation in human atherosclerotic coronary arteries. Across the ‘SMC’ and ‘fibromyocyte’ clusters, we performed pairwise Spearman correlation between the expression of TCF21 and every other gene expressed in these cells. We found that TCF21 was highly anticorrelated with markers of differentiated SMCs, demonstrating that increased TCF21 expression in these cell clusters was associated with SMC de-differentiation (Fig. 5a). In addition, we found that TCF21 was highly correlated with many markers of fibromyocytes in both the human and the mouse. As Tcf21 was expressed at low levels (Extended Data Fig. 5f), we visualized the behavior of the TCF21-associated gene program within the cell populations by creating a ‘TCF21 score’ for each cell (see Methods), which reflected the averaged expression of top TCF21-correlated genes across the ‘SMC’ and ‘fibromyocyte’ clusters. Importantly, to establish a causal link with TCF21, all genes included in the score were required to also display robust TCF21 binding within the gene locus as assessed by TCF21 chromatin immunoprecipitation sequencing (ChIP-Seq; example peaks are shown in Extended Data Fig. 5f). This analysis clearly revealed a graded increase in this TCF21-associated gene expression program that correlated with SMC phenotypic switching in vivo, suggesting that TCF21 plays a causal role in SMC phenotypic modulation in human coronaries.
Fig. 5 | TCF21 is associated with SMC modulation in human coronary arteries. a, Pairwise Pearson correlation of TCF21 with every other gene in cells across the ‘SMC’ and ‘fibrocyte’ clusters from Fig. 4a. Selected examples of genes regulated during SMC modulation are labeled. b, t-SNE visualization of cell clusters of the human coronary samples (n = 4 patients). The 20 most highly correlated and anticorrelated genes from a were used to calculate a TCF21-associated human cell expression score, which ranges from highly anticorrelated (blue) to highly correlated (red). c, Expression of SMC modulation marker genes with TCF21 overexpression in HCASMCs. The results are from six independent experiments, each with three technical replicates. Statistical significance was determined by comparing fold-change values using a two-sided Mann-Whitney U-test. NC, empty vector negative control; OE, overexpression. Error bars denote s.d.

CAD risk alleles are associated with decreased TCF21 expression. To further understand the impact of TCF21 expression and phenotypic modulation on human disease risk, we investigated the relationship between genome-wide association study (GWAS) single nucleotide polymorphism (SNP) genotypes at the 6q23.2 locus and TCF21 expression (cis-expression quantitative trait loci (cis-eQTL) analysis). We first assessed seven SNPs associated with CAD at genome-wide significance. Figure 6a illustrates these SNPs and their linkage disequilibrium relationships. In a group of 52 HCASMC lines, we found that, without exception, the CAD risk allele of each SNP was associated with decreased TCF21 expression (Extended Data Fig. 6). We also observed the same finding for these SNPs in highly CAD-relevant tissues in the Genotype-Tissue Expression (GTEx) database (Extended Data Fig. 6). In our HCASMC lines, these risk alleles showed an additive effect; haplotypes accumulating greater numbers of risk alleles were associated with progressively lower TCF21 expression (P = 0.114; R = −0.41; Fig. 6b). We then assessed a larger number of SNPs in the 6q23.2 locus that were associated with CAD risk at a false discovery rate of 1 × 10−5. By performing eQTL analysis in aortic tissue from the STARNET

Fig. 6 | Reduced TCF21 expression is associated with increased coronary disease risk. a, Linkage disequilibrium relationships (R² values) of all genome-wide significant CAD-associated SNPs at the 6q23.2 locus (bottom), relative to the position of TCF21 and long noncoding RNAs (LINC01312 and TARID) within the locus (top). The R² color indicates the degree of linkage disequilibrium between each pair of SNPs, and ranges from 0 (gray) to 1 (red). The correspond ing R² values are also shown in each box. b, Relationship between the number of genome-wide significant CAD risk alleles in each haplotype (x axis) and TCF21 expression (y axis) in 52 primary HCASMC lines. RMP, reads per million mapped reads. c, Correlation between the magnitude of CAD risk imparted by each risk allele (x axis) and the relative TCF21 expression from that allele (y axis) in 36 CAD-associated SNPs at the 6q23.2 locus in aortic tissue from the STARNET database. The P value was calculated using Pearson’s moment correlation coefficient. Gray shaded areas indicating 95% confidence intervals are based on Fisher’s Z-transform.
Articles

Nature Medicine

These eQTL findings from multiple independent SNPs in multicAD-relevant tissues strongly argue that TCF21 expression is protective against disease.

Discussion

The phenomenon of SMC phenotypic modulation has been studied primarily by exposing cultured SMCs to lipids and various growth factors. These in vitro studies have consistently reported downregulation of SMC markers, increased migration, proliferation, extracellular matrix secretion, upregulation of certain inflammatory cytokines, macrophage markers, and increased levels of phagocytic activity. However, SMC phenotypic modulation has been very difficult to study in vivo in mice and humans due to both reduced expression of canonical SMC markers and expression of some SMC markers by other cell types. In a recent landmark paper, Shankman et al. used SMC lineage tracing to definitively identify phenotypically modulated cells of SMC origin in atherosclerotic lesions. However, assessment of modulated SMC phenotype with in situ studies was necessarily limited to a small number of markers. The identification and characterization of modulated SMCs in human plaques has been even more challenging.

On the basis of these studies, the current approach is that modulated SMCs can adopt either: (1) a pro-inflammatory macrophage-like phenotype characterized by Lgals3 expression, which could contribute to the protective fibrous cap; or possibly (3) a mesenchymal stem cell-like population of unclear significance. Because of this uncertainty regarding the phenotype of modulated SMCs within the lesion, it is also unclear whether the process of SMC phenotypic modulation leads to a more stable or less stable atherosclerotic plaque, and thus whether SMC modulation is protective or increases the risk for CAD and myocardial infarction.

In this study, we found that instead of assuming multiple distinct cell phenotypes, SMCs undergoing phenotypic modulation appear to exhibit a shift in gene expression along a continuous trajectory from a contractile SMC towards a fibroblast-like cell, which we term a ‘fibromyocyte’. This name was created to emphasize the opposite phenotypic trajectory of these cells compared with a fibroblast-derived ‘myofibroblast’ that acquires properties of SMCs. Although fibromyocytes display a decreasing gradient of SMC gene expression, they are a highly distinct population and cluster independently, even at very low clustering resolutions. The transcriptional profile of SMC lineage-traced fibromyocytes in the mouse was employed to identify an orthogonal human fibromyocyte population.

There has been much interest in the possibility that phenotypically modulated SMCs may adopt a detrimental macrophage-like phenotype during atherosclerosis, engulfing oxidized low-density lipoprotein and dying cells, and eventually becoming plaque destabilizing foam cells. SMCs subjected to cholesterol loading in culture accumulate intracellular lipid reminiscent of foam cells and display modest upregulation of the macrophage markers LGALS3 and CD68 (ref. ), as well as a modest increase in phagocytic behavior.

In mice, Mhyh11 lineage-traced SMCs that migrate into the lesion express Lgals3 (ref. ). In human coronaries, CD68 expression was observed in cells within the lesion that also expressed an SMC-specific epigenetic mark. Our data, combining SMC lineage tracing with scRNA-Seq measuring thousands of genes simultaneously, suggest that fibromyocytes in vivo do not acquire a macrophage-like transcriptional phenotype. We confirmed these findings using multiple methods at the protein level. We note that our data do not exclude the possibility that SMC-derived cells could be participating in effectorcytosis as ‘non-professional’ phagocytes.

Interestingly, we found a small number of modulated SMCs at the baseline time point, which is consistent with the recent finding of a rare Sca1+ SMC population in healthy mice. While the authors of this recent study did not demonstrate that these Sca1+ SMCs are an exclusive contributor to lesional SMC-derived cells, many modulated SMCs within the lesion do express Sca1 (Extended Data Fig. 2c).

The oligoclonal nature of SMC contribution to atherosclerosis has already been established by multiple groups using multicolored lineage-tracing methods. When interpreted in the context of our findings, these studies suggest that a small number of SMCs undergo phenotypic modulation and expand to create the fibromyocyte population. However, it is not currently possible to determine whether some or all fibromyocytes arise from a distinct population of modulated SMCs that exist homeostatically within the healthy vessel wall.

SMC-specific Tcf21 knockout revealed that Tcf21 promotes phenotypic modulation in vivo. It is interesting that, during phenotypic modulation, SMCs upregulate Tcf21 and transform into a fibroblast-like phenotype, and that the loss of Tcf21 inhibits this phenotypic transition. In embryogenesis, Tcf21 regulates fundamental cell fate decisions in the developing epicardium, serving as a fate-determining factor for the divergence of coronary vascular SMC and cardiac fibroblast lineages. In this setting, Tcf21 is downregulated in cells that are fated to become differentiated coronary SMCs, and expression remains in the interstitial and adventitial fibroblast lineage. Strikingly, Tcf21−/− mice show a near-complete absence of cardiac fibroblasts, indicating that Tcf21 expression is required for fibroblast development. Thus, during atherosclerosis, Tcf21 appears to be recapitulating its developmental role by directing cell fate decisions away from SMCs and towards a fibroblast gene expression program.

This study shows that a gene that is causally associated with CAD at genome-wide significance can fundamentally alter the process of SMC phenotypic modulation in vivo. Given our finding that Tcf21 expression promotes SMC phenotypic modulation in vivo, we had a unique opportunity to assess whether TCF21 and SMC phenotypic modulation are protective or deleterious during CAD. We identified multiple independent lines of evidence suggesting that TCF21 expression is causally associated with a reduced risk of CAD. Taken together, these data suggest that both TCF21 expression and SMC phenotypic modulation are beneficial during the disease process, although likely in a time and context-dependent manner. On the basis of the finding that a loss of Tcf21 results in fewer fibromyocytes in the lesion and at the protective fibrous cap, it is quite plausible that Tcf21 exerts its protective effect by promoting the infiltration of fibromyocytes into the lesion and fibrous cap.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41591-019-0512-5.

Received: 18 September 2018; Accepted: 5 June 2019;
Published online: 29 July 2019

References

1. Davies, M. J., Richardson, P. D., Woolf, N., Katz, D. R. & Mann, J. Risk of thrombosis in human atherosclerotic plaques: role of extracellular lipid, macrophage, and smooth muscle cell content. Br. Heart J. 69, 377–381 (1993).
2. Libby, P. & Aikawa, M. Stabilization of atherosclerotic plaques: new mechanisms and clinical targets. Nat. Med. 8, 1257–1262 (2002).
3. Ross, R. & Glomset, J. A. Atherosclerosis and the arterial smooth muscle cell: proliferation of smooth muscle is a key event in the genesis of the lesions of atherosclerosis. Science 180, 1332–1339 (1973).
4. Bennett, M. R., Sinha, S. & Owens, G. K. Vascular smooth muscle cells in atherosclerosis. Circ. Res. 118, 692–702 (2016).
5. Gomez, D. & Owens, G. K. Smooth muscle cell phenotypic switching in atherosclerosis. Cardiovasc. Res. 95, 156–164 (2012).
6. Shankman, L. S. et al. KL4-dependent phenotypic modulation of smooth muscle cells has a key role in atherosclerotic plaque pathogenesis. Nat. Med. 21, 628–637 (2015).
7. Gomez, D., Shankman, L. S., Nguyen, A. T. & Owens, G. K. Detection of histone modifications at specific gene loci in single cells in histological sections. Nat. Methods 10, 171–177 (2013).
8. Schunkert, H. et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat. Genet. 43, 333–338 (2011).
9. Miller, C. L. et al. Disease-related growth factor and embryonic signaling pathways modulate an enhancer of TCF21 expression at the 6q23.2 coronary heart disease locus. PLoS Genet. 9, e1003652 (2013).
10. Nurnberg, S. T. et al. Coronary artery disease-associated variation in the myocardin axis to convert aortic smooth muscle cells to a dysfunctional phenotype. Arterioscler. Thromb. Vasc. Biol. 31, 13531–13536 (2013).
11. Dettman, R. W., Denetclaw, W. Jr, Ordahl, C. P. & Bristow, J. Common epicardial origin of coronary vascular smooth muscle, perivascular fibroblasts, and intermyocardial fibroblasts in the avian heart. Dev. Biol. 193, 169–181 (1998).
12. Winter, E. M. & Gittenberger-de Groot, A. C. Epicardium-derived cells in cardiogenesis and cardiac regeneration. Cell. Mol. Life Sci. 64, 692–703 (2007).
13. Acharya, A. et al. The βH1L1 transcription factor Tcf21 is required for lineage-specific EMT of cardiac fibroblast progenitors. Development 139, 2139–2149 (2012).
14. Nurnberg, S. T. et al. Coronary artery disease associated transcription factor TCF21 regulates smooth muscle precursor cells that contribute to the fibrous cap. PLoS Genet. 11, e1005155 (2015).
15. Herring, B. P., Hoggatt, A. M., Burlak, C. & Offermanns, S. Previously identified a transcriptional network linking multiple independent coronary artery disease loci. PLoS Genet. 11, e1005202 (2015).
16. Wirth, A. et al. G12-G13-LARG-mediated signaling in vascular smooth muscle is required for salt-induced hypertension. Nat. Med. 14, 64–68 (2008).
17. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13, 133–140 (2010).
18. Zheng, G. X. et al. Massively parallel digital transcriptional profiling of single cells. Nat. Commun. 8, 14049 (2017).
19. Dohukuri, L. et al. Disease-relevant transcriptional signatures identified in individual smooth muscle cells from healthy mouse vessels. Nat. Commun. 9, 4567 (2018).
20. Kitchen, C. M., Cowan, S. L., Long, X. & Miano, J. M. Expression and promoter analysis of a highly restricted integrin alpha gene in vascular smooth muscle. Gene 513, 82–89 (2013).
21. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. Nat. Methods 14, 865–868 (2017).
22. Jacobsen, K. et al. Diverse cellular architecture of atherosclerotic plaque derives from clonal expansion of a few medial SMCs. Nat. Commun. 8, 12260 (2016).
23. Vengrenyuk, Y. et al. Cholesterol loading reprograms the microRNA-143/145–myocardin axis to convert aortic smooth muscle cells to a dysfunctional macrophage-like phenotype. Arterioscler. Thromb. Vasc. Biol. 35, 535–546 (2015).
24. Sazonova, O. et al. Characterization of TCF21 downstream target regions identifies a transcriptional network linking multiple independent coronary artery disease loci. Genes 10, e1005202 (2015).
25. Franzen, O. et al. Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. Science 353, 827–830 (2016).
26. Clement, N. et al. Notch3 and IL-1β exert opposing effects on a vascular smooth muscle cell inflammatory pathway in which NF-κB drives crosstalk. J. Cell Sci. 120, 3352–3361 (2007).
27. Pidkovka, N. A. et al. Oxidized phospholipids induce phenotypic switching of vascular smooth muscle cells in vivo and vitro. Circ. Res. 101, 792–801 (2007).
28. Dandre, F. & Owens, G. K. Platelet-derived growth factor-BB and Ets-1 transcription factor negatively regulate transcription of multiple smooth muscle cell differentiation marker genes. Am. J. Physiol. Heart Circ. Physiol. 286, H2042–H2051 (2004).
29. Rong, J. X., Shapiro, M., Trogan, E. & Fisher, E. A. Transdifferentiation of mouse aortic smooth muscle cells to a macrophage-like state after cholesterol loading. Proc. Natl Acad. Sci. USA 100, 13351–13356 (2003).
30. Rong, J. X., Berman, J. W., Taubman, M. B. & Fisher, E. A. Lysophosphatidylcholine stimulates monocyte chemoattractant protein-1 gene expression in rat aortic smooth muscle cells. Arterioscler. Thromb. Vasc. Biol. 22, 1617–1623 (2002).
31. Allahverdian, S., Chehroudi, A. C., McManus, B. M., Abraham, T. & Francis, G. A. Contribution of intimal smooth muscle cells to cholesterol accumulation and macrophage-like cells in human atherosclerosis. Circulation 129, 1551–1559 (2014).
32. Kanisicak, O. et al. Genetic lineage tracing defines myofibroblast origin and function in the injured heart. Nat. Commun. 7, 12260 (2016).
33. Chappell, J. et al. 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. 119, 1313–1323 (2016).
METHODS

Mouse strains. To enact SMC-specific lineage tracing and Tcf21 knockout, we used mice containing a well-characterized BAC transgene that expresses a tamoxifen-inducible Cre recombinase driven by the SMC-specific Myh11 promoter (Tcf21xGt(ROSA)26Sor tm14(CAGtdTomato)Hze/J; 007979; JAX). These were bred with a flexed tdT fluorescent reporter line (B6.Cg-Gt(Rosa)26Sortm1(CAGtdTomato)Sor/J; 007914; JAX) to allow SMC-specific lineage tracing. A Tcf21x allele was constructed involving placing LoxP sites flanking the 5′ promoter region and first exon of the Tcf21 gene. All mice were bred onto the C56BL/6 background. The final genotypes of SMC lineage-tracing (SMC+/-) mice were: Tcf21Gt(Rosa)26Sortm14l(CAGtdTomato)Hze+/-; Myh11-Cre+/-; Rosa26tdTomato+/-; and ApoE+/-.

Induction of lineage marker and Tcf21 knockout by Cre recombinase. For all scRNA-Seq, CITE-Seq, RNASEQ and immunohistochemistry experiments involving BODIPY, the tamoxifen gavage schedule was as follows: two doses of tamoxifen, at 0.2 mg g−1 bodyweight, were administered by oral gavage at 7 weeks of age, with each dose separated by 48 h. Two doses were used to ensure complete activation of the Cre-ERT2. The recombinase efficiency was estimated using the FACS-sorted scRNA-Seq data at the baseline time point. This revealed that 3,538 out of 3,577 cells (99.8%) classified as SMCs by scRNA-Seq in the baseline mouse were also FACS-positive for the tdT lineage marker. Approximately 48 h after the second dose of tamoxifen, HFD was started (101511; Dyets; 21% anhydrous milk fat, 19% casein and 0.15% cholesterol). For the quantitative immunohistochemistry and immunofluorescence (in situ) experiments, in addition to the two baseline doses of tamoxifen, a single additional dose of tamoxifen was administered by oral gavage after 8 weeks of HFD and then again at 16 weeks HFD, approximately 48 h before sacrifice (Extended Data Fig. 1b). For these quantitative in situ experiments, we assessed the likelihood that the additional doses of tamoxifen at the 8 and 16-week HFD time points could result in spurious recombination in lineage-negative cell types. We analyzed Myh11 expression in these lineage-negative cell types at 8 and 16 weeks of disease in the scRNA-Seq data and found that none of these cell types had upregulated Myh11 during disease and were therefore extremely unlikely to undergo recombination in response to the additional tamoxifen doses (data not shown).

Mouse aortic root/ascending aorta cell dissociation. Immediately after sacrifice, mice were perfused with phosphate buffered saline (PBS). The aortic root and ascending aorta were excised, up to the level of the brachiocephalic artery. Tissue was washed three times in PBS, placed into an enzymatic dissociation cocktail (2 µM -Liberase (5401127001; Sigma–Aldrich) and 2 µM elastase (LS002279; Worthington) in Hank’s Balanced Salt Solution (HBSS)) and minced. After incubation at 37°C for 1 h, the cell suspension was strained and then pelleted by centrifugation at 500 × g for 5 min. The enzyme solution was then discarded and cells were resuspended in fresh HBSS. To increase biological replication, multiple mice were used to obtain single-cell suspensions at each time point. For the SMC- genotype, three mice were used at baseline, and three mice were used at both 8 and 16 weeks of disease. For the SMC- genotype, one mouse was used at 8 weeks and three mice were used at 16 weeks.

Human coronary artery cell dissociation. 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. The basic clinical and pathological criteria of the four patients included in this study are presented in Supplementary Table 5. The proximal to mid right coronary artery was identified, excised, cleaned of peri-arterial fat, and then rinsed three times in PBS. After excluding stented areas, atherosclerotic lesions were identified, ranging from mild, noncalcified plaques to more advanced lesions with areas of calcification. These areas were cut into approximately 50 mg sections and each section broke into an enzymatic dissociation cocktail (10.4 µM -Liberase and 8 µM elastase (E7885; Sigma–Aldrich) in 1 ml HBSS) and minced. A total of approximately 120–240 mg (~3–6 atherosclerotic sections) were used per patient. After incubation at 37°C for 1 h with periodic agitation, the cell suspension was pipetted up and down to break up any remaining tissue. The cell suspension was strained and then pelleted by centrifugation at 500 × g for 5 min. The enzyme solution was then discarded and cells were resuspended in fresh HBSS.

FACS of mouse aortic root/ascending aorta cells. Cells were sorted on a BD Arista II instrument. An overview of the cell-sorting process is illustrated in Extended Data Fig. 1c. Cells were gated on forward/side scatter parameters to exclude small debris and then gated on forward scatter height versus forward scatter area to exclude obvious doublet events. Events passing these criteria were then sorted into one of two 1.5 ml Eppendorf tubes based on tdT fluorescence levels. tdT cells (considered to be of SMC lineage) and tdT- cells were then captured on separate but parallel runs of the same scRNA-Seq workflow, and datasets were later combined for all subsequent analyses.

FACS of human coronary artery cells. Cells were incubated with the calcine green viability reagent (C34852; Thermo Fisher Scientific) for 30 min at 4°C before sorting, and a small portion of cells were left unstained as a negative control to determine gate placement. Cells were sorted on a Sony SH800S instrument. Cells were gated on forward/side scatter parameters to exclude small debris and then gated on forward forward scatter height to forward scatter area to exclude various doublet events. Calcine− cells were then positively selected and sorted into a 1.5-ml Eppendorf tube for further processing in the scRNA-Seq workflow.

Single-cell capture and library preparation. All single-cell capture and library preparation was performed at the Stanford Functional Genomics Facility. Cells were loaded into a 10x Genomics microfluidics chip and encapsulated with barcode conjugated gel beads using the 10x Genomics Chromium controller according to the manufacturer’s instructions. Single-cell libraries were then constructed according to the manufacturer’s instructions. Libraries from individual samples were multiplexed into one lane before sequencing on an Illumina HiSeq 4000 instrument.

CITE-Seq. Cells were obtained from the atherosclerotic aortic root and ascending aorta of two SMC+ mice on 16-week HFD as already described. After 1 h of enzymatic dissociation, fetal bovine serum (FBS; final concentration 10%) was added to the enzyme mixture, cells were centrifuged and supernatant was discarded. Cells were resuspended in 100 µl cell staining buffer (2% bovine serum albumin and 0.01% Tween in PBS). A pool of TotalSeq antibodies were added at 1 µg each (Cd68/Cd32 (101319; BioLegend), Cd11b (101265; BioLegend), Cd64 (139325; Bio-Legend), Cd86 (105047; BioLegend) and F4/80 (123135; BioLegend)) and cells were incubated for 30 min at 4°C, followed by washing three times in staining buffer. Cells were isolated by FACS as previously described and captured on the 10x Chromium controller. Libraries were constructed according to the manufacturer’s protocol, with the following modifications according to the CITE-Seq protocol (https://cite-seq.com/protocol): (1) during the cdDNA amplification step, an additional primer (5′-CCTGGCCCGAGAATT*C*C) was added to increase the yield of the antibody-derived tag (ADT) products; and (2) during library preparation, ADT-derived and messenger RNA-derived cdDNAs were separated by SPR1 selection. The messenger RNA-derived cdDNA fraction was used to construct 10x libraries according to the manufacturer’s instructions. The ADT-derived cdDNA fraction was then purified using Solid Phase Reversible Immobilization (SPRI) beads and then amplified using a 10x Genomics SI-PCR primer (5′-AATGATACGG CGACACGAGATCTACCTTCTCTCATCACGACGCT*C*C) and an Illumina Small RNA RPI primer (5′-CAAAGGAAAGGCAGATCACGATCGATGTGACTGGATTCTCCTGCGACGGAGATCT*C*C). PCR products were purified with SPRI and pooled with standard 10x libraries for sequencing on a HiSeq 4000 instrument.

FACS analysis for macrophage markers. Cells were obtained from the atherosclerotic aortic root and ascending aorta of one SMC+ mouse (16 weeks HFD) as already described. FBS (final concentration 10%) was added to the enzyme mixture, cells were centrifuged and the supernatant discarded. Cells were resuspended in 100 µl cell staining buffer (2% bovine serum albumin and 0.01% Tween in PBS). Cells were incubated with 1 µg Rat anti-CD16/32 (115191; Bio-Legend) for 10 min at 4°C, washed twice and then incubated with 0.4 µg BV421 goat anti-rat secondary (565013; BD Biosciences) at 4°C for 30 min. Cells were then washed three times before FACS analysis.

Preparation of mouse aortic root sections. Immediately after sacrifice, mice were perfused with 0.4% paraformaldehyde (PFA). The mouse aortic root and proximal ascending aorta, along with the base of the heart, was excised and immersed in 4% PFA at 4°C for 12 h (for immunohistochemistry) to 24 h (for RNAscope). After passing through a sucrose gradient, tissue was frozen in optimal cutting temperature compound (OCT) to make blocks. Blocks were cut into 7-µm-thick sections for further analysis.

Immunohistochemistry. Slides were air dried and OCT was removed with two washes in deionized water. Slides were immersed in 4% PFA for 2 min, followed by four washes in deionized water. Slides were dried and sections were encircled with a liquid blocking pen, followed by Peroxidized (P9968; Biocare Medical) treatment for 5 min. Sections were washed three times with deionized water, and then incubated with Rodent Block M reagent (RBM961; Biocare Medical) for 30 min. Sections were washed twice in Tris-buffered saline (TBS), then incubated overnight at 4°C with an anti-SM22alpha rabbit polyclonal primary antibody (1:300 dilution; ab14106; Abcam), a Lgals3 rat monoclonal antibody (1:1000 dilution; Cl842AP; Cederlane Labs) or a CD68 rabbit polyclonal antibody (1:300 dilution; ab125212; Abcam). Sections were washed for 5 min twice with TBS and then incubated with the Rabbit-on-Rodent HRP Polymer (RMR622; Biocare Medical) or Rat Probe followed by Rat-on-Mouse HRP Polymer (RT517; Biocare Medical) for 30 min at room temperature. Sections were washed twice with TBS.
and then incubated with the Becton DAKO chromogen reagents (BDB2004; Biocare Medical) for 4 min at room temperature. Sections were washed twice in deionized water and air dried, followed by mounting with EcoMount medium (EM897L; Biocare Medical). The processed sections were visualized using a 2-Leica DM5000 microscope under 5x and 10x (for lesion cap analysis) objective magnifications, and images were obtained using Leica Application Suite X software. Sections obtained at equal distance measured from the superior margin of the aortic sinus were used for comparison. Areas of interest were quantified using ImageJ (National Institutes of Health) software, and compared using a two-sided t-test. The lesion cap was defined as a 30-μm segment from the luminal surface as previously described. Medial size was calculated by measuring the areas within the outer circumference and inner circumference of the medial layer. The difference between the two areas was used as the medial size/area.

The Tagln immunostained images, which include the representative images shown in Fig. 3i, were used to derive the murine medial layer and delineated the medial layer clearly. Researchers were blinded to the genotype.

The Cd68/BODIPY staining.

delineated the medial layer clearly. Researchers were blinded to the genotype.

The dataset was trimmed of cells expressing fewer than 500 genes, and an aggregated dataset was then analyzed using the R package Seurat version 2.3.4 using the CellRanger aggr command without subsampling normalization. The SMC modulation score was then calculated using the top 20 genes that distinguished modulated SMCs and contractile SMCs (SMC1, SMC2, pericytes and a small number of phenotypically modulated SMCs).

The transcriptional ‘shift’ of fibromyocytes towards or away from a given cell cluster ‘X’ in Fig. 2i was calculated as: (distance between the combined quiescent SMCs centroid and cluster ‘X’ centroid) – (distance between the fibromyocyte centroid and cluster ‘X’ centroid).

HCASMC culture. A total of 65 primary HCASMC lines were purchased from PromoCell (catalog number: C-12511), Cell Applications (catalog number: 350-05a), Lonza (catalog number: CC-2583), Lifeline Cell Technology (catalog number: FC-090G), and ATCC (catalog number: PCS-100-021). Cells were cultured in smooth muscle growth medium (Lonza; catalog number: CC-3182) supplemented with human epidermal growth factor, insulin, human basic fibroblast growth factor and 5% FBS, according to the manufacturer’s instructions. All HCASMC lines were used at passages 4–8.

Pooled TCF21 chromatin immunoprecipitation sequencing (ChIP Seq).

Approximately 300,000 HCASMCs from each cell line were cross linked in 1% formaldehyde for 10 min and then washed with PBS. Cell pellets were frozen at −80 °C. All cell pellets were thawed on ice, combined for a total of 19.5 million cells, and resuspended in cold PBS. PBS was removed and replaced with hypotonic buffer (20 mM Heps (pH 7.9), 10 mM KCl, 1 mM EDTA (pH 8) and 10% glycerol) and the cells were incubated on ice for 6 min. Cells were dounce homogenized with 20 strokes on ice using a 7-ml glass homogenizer. Nuclear lysates were sonicated using a Branson 250 Sonifier (power setting 4; constant duty for 12 rounds of 20-s pulses), resulting in chromatin fragments of 250–400 base pairs. DNA at 37 °C was treated overnight with 5 U of Superscript II (Invitrogen) (HAPA01389; Sigma–Aldrich). Protein–DNA complexes were captured on Protein A agarose beads (16-266; Millipore Sigma) and eluted in 1% SDS TE buffer at 65 °C. After reverse cross-linking, and RNase A and proteinase K digestion, chromatin was purified using a QiaGen PCR purification kit (catalog number: 28106). ChIP DNA sequencing libraries were generated as previously described36 and sequenced on an Illumina HiSeq X instrument (150-base pair paired-end reads). Fastq files were mapped to the Hg19 genome with BWA-MEM aligner version 0.7.12 (ref. 30). ChIP-Seq peaks were then called with MACS version 2 (ref. 30) using default parameters. From this output, ‘robust’ peaks were selected by specifying a minimum fold-enrichment of 10 and a minimum log2(p-value) of 60. To determine which gene regions contained these robust TCF21 ChIP-Seq peaks, we expanded each National Center for Biotechnology Information ReSeq-annotated gene region by 5 kilobases in each direction and then determined the overlap of these gene regions with the TCF21 ChIP-Seq data.

TCF21 overexpression in HCASMCs. TCF21 cDNA was cloned into a second-generation lentiviral vector (pWP1; 12254; Addgene) and packaged under the Stanford Gene Vector and Virus Core. HCASMCs were treated at 60% confluence with lentivirus at a multiplicity of infection of 5 for 24 h. The virus was removed and cells were collected 48 h later. Gene expression was assessed using TCF21 (in brackets) quantification probes (Thermo Fisher Scientific) for TCF21 (Hs00162646_m1), DNA (Hs00754870_s1), LUM (Hs00928660_m1) and matrix-Gla protein (MGP; Hs09694990_m1) according to the manufacturer’s instructions on a ViaA 7 Real-Time PCR system (Applied Biosystems). A total of six independent experiments were performed, each with three technical replicates. Fold-change values from control and overexpression conditions were compared with a two-sided Mann–Whitney U-test in Prism version 8 (GraphPad).

HCASMC genome and transcriptome sequencing. HCASMC genomic DNA was isolated using a Qiagen DNeasy Blood & Tissue Kit (catalog number: 69506). Libraries were prepared with Illumina’s TruSeq DNA PCR-Free Library Preparation Kit and sequenced on a HiSeq X Ten System. We sequenced using a Qiagen miRNAeasy Mini Prep Kit (catalog number: 74106). Libraries were made using an Illumina TruSeq Stranded Total RNA Library Prep Kit (catalog number: 20020597) and sequenced on a HiSeq 2500 Platform. Whole-genome sequencing data were processed with the GATK best practices pipeline with hg19 as the reference genome. VCF records were phased with Beagle. Demultiplexed FASTQ files were mapped with STAR version 2.4.0 in two-pass mode with hg19.

TCF21 eQTL analyses. We first examined all SNPs in the 6q23.2 locus that were associated with CAD at genome-wide significance (P < 5 × 10−8; Fig. 6b). We queried individual eQTL relationships from the GTex database and in our cohort of 52 HCASMC lines. To test for an additive effect of these cis CAD risk alleles on TCF21 expression in the HCASMC lines, we determined the relationship between the number of risk SNPs inherited and the TCF21 expression level in the HCASMC lines. Total read counts were determined using RNA-Seq and transcript-per-million values were then calculated for the TCF21 gene. To reduce noise, TCF21 transcript-per-million data from cell lines with identical haplotypes
were averaged before analysis. We then examined a larger set of SNPs at the 6q23.2 locus associated with CAD at a \(P\) value of \(<1 \times 10^{-5}\), extracted from CAD GWAS summary data (www.cardiogramplusc4d.org). We also obtained cis-eQTL summary data from the STARNET database (dbGaP), derived from human aortic tissue, for these CAD-associated SNPs. We filtered for SNPs with an absolute value of beta coefficients (\(\log[\text{odds ratio}]\)) greater than 0.3 for both GWAS and eQTL. A linear mixed model was then used to compute a smooth local regression between the CAD GWAS beta and the eQTL \(\log[\text{odds ratio}]\). Pearson’s correlation coefficient \(R\) and \(P\) values of significance were calculated using cor.test in R.

**Statistical methods.** Differentially expressed genes in the scRNA-Seq data were identified using a Wilcoxon rank-sum test, as implemented in the Seurat package version 2.3.4. In the scRNA-Seq data, the distribution of fibromyocytes and quiescent SMCs across the SMC\(^{\text{WT}}\) and SMC\(^{\text{KO}}\) genotypes (\(n=3\) mice in each genotype) was calculated using Pearson’s chi-squared test (chiq.test) in R. For the quantitative immunohistochemistry experiments, cohorts included 17 SMC\(^{\text{WT}}\) (WT) and 22 SMC\(^{\text{KO}}\) (KO) mice. Due to occasional tissue block, section damage and folding, some animals had to be eliminated for some comparisons. In Fig. 3d, WT = 16 and KO = 17. In Fig. 3e, WT = 15 and KO = 18. In Fig. 3f, WT = 16 and KO = 17. In Fig. 3g, WT = 14 and KO = 17. In Fig. 3j, WT = 16 and KO = 22. In Extended Data Fig. 3k, WT = 15 and KO = 16. In Extended Data Fig. 3e, WT = 16 and KO = 20. In Extended Data Fig. 3f, WT = 16 and KO = 22. Comparisons in the mouse aortic root were made with a Student’s \(t\)-test (two sided). For qPCR analysis comparing TCF21 overexpression versus control in HCASMCs, six independent experiments were performed, each with three technical replicates. Fold-change values (\(n=6\) in each group) were analyzed using a two-sided Mann–Whitney \(U\)-test in Prism version 8 (GraphPad). To test the relationship between the number of TCF21 risk SNPs and TCF21 expression in HCASMCs, and to test the relationship between the CAD GWAS beta value and the magnitude of the eQTL in the STARNET dataset, linear mixed models were used; the \(P\) value of significance was based on Pearson’s product moment correlation coefficient, and 95% confidence intervals were based on Fisher’s \(Z\)-transform, as computed using cor.test in R. For HCASMCs, \(n=52\) cell lines with 16 distinct haplotypes were analyzed. For STARNET data, \(n=36\) SNPs were analyzed.

**Data availability**

High-throughput sequencing data (FASTQ) files for all scRNA-Seq, CITE-Seq and ChiP-Seq, as well as cell–gene count matrices for all scRNA-Seq and CITE-Seq experiments, have been deposited in the Gene Expression Omnibus database with the SuperSeries reference number GSE131780. These data were used to generate the images in Figs. 1–5 and Extended Data Figs. 2–5. FASTQ files and processed data are also available from the corresponding author upon request.

**References**

34. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 411–420 (2018).
35. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. 33, 495–502 (2015).
36. Kasowski, M. et al. Variation in transcription factor binding among humans. Science 328, 232–235 (2010).
37. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009).
38. Zhang, Y. et al. Model-based analysis of ChiP-Seq (MACS). Genome Biol. 9, R137 (2008).
Extended Data Fig. 1 | Design of mouse experiments. **a.** Alleles present in SMC^lin^ and SMC^KO^ mice. KO (knockout) refers to Tcf21^lin^, lineage tracing; Tg, transgene; ΔSMC, SMC cell-specific KO. **b.** Mice were maintained on a chow diet from birth until 7 weeks of age, then underwent gavage and HFD treatment. For the scRNA-Seq, RNAscope, CITE-Aeq, histology involving BODIPY and FACS staining experiments (upper timeline), mice were gavaged only at 7 weeks of age, before the onset of HFD, as denoted by red arrows. For the scRNA-Seq experiments, mice were sacrificed at baseline (72 h after initial tamoxifen gavage), or after 8 or 16 weeks of HFD. For the RNAscope experiments, mice were sacrificed after either 8 or 16 weeks of HFD. For the CITE-Seq experiment, mice were sacrificed after 16 weeks of HFD. For the BODIPY studies, mice were sacrificed after 16 weeks of HFD. For the FACS staining experiment, two mice (one after 12 weeks HFD and another after 15 weeks HFD) were used. For quantitative histology experiments (lower timeline), mice were gavaged at 7 weeks of age, after 8 weeks of HFD and after 16 weeks of HFD (48 h before sacrifice), as denoted by red arrows. For these quantitative histology experiments, all mice were sacrificed after 16 weeks of HFD. **c.** FACS workflow for isolating single cells from the mouse aortic root.
Extended Data Fig. 2 | see figure caption on next page.
Extended Data Fig. 2 | SMC phenotypic modulation in the mouse aortic root. a–d, t-SNE visualization of cell types present in the WT mouse aortic root from all time points overlaid with the expression of Col1a1, Acta2, Sca1 and Lgals3 (n = 9 mice). e–g, RNAscope staining for Lum (green) and tdT (red) in a plaque after 8 weeks of HFD (e), the non-diseased media of a mouse on 16 weeks HFD (f) and a baseline healthy aorta (g). h, RNAscope negative control. Images in e–h are representative of two experiments and scale bars indicate 25 μm. i, t-SNE visualization of cell types present in the WT mouse aortic root from all time points overlaid with osteopontin (Spp1) expression (n = 9 mice). j,k, RNAscope co-localization of Spp1 (green) and tdT (red) in a plaque after 16 weeks of HFD. Yellow arrows indicate co-localization of Spp1 and tdT. l, RNAscope negative control. Images from j–l are representative of four experiments, and scale bars indicate 50 μm. m, Heat map representation of the Euclidean distance between cell cluster centroids in 20-dimensional principal component space with the smallest distances in yellow and the largest distances in black. Data are after 16 weeks of HFD. n, Staining of a single-cell suspension from the atherosclerotic mouse aortic root and ascending aorta with antibodies against the macrophage markers Cdl6 and Cdl32, and analysis of co-expression with the tdT SMC lineage marker. Data are from one experiment and n = 2 mice (after 12 and 15 weeks of HFD). o–t, Single cells from the atherosclerotic mouse aortic root and ascending aorta at 16 weeks HFD were incubated with DNA-barcode antibodies against the macrophage markers Cdl6, Cdl32, Cdl11b, Cdl64, Cdl86 and F4/80 before undergoing scRNA-Seq (CITE-Seq), yielding simultaneous transcriptomic and antibody binding data within each individual cell. o, Cell type assignments were determined with scRNA-Seq as described previously. p–t, Quantitative antibody binding within each cell type. Results are from one experiment and n = 2 mice.
Extended Data Fig. 3 | Additional characteristics of SMC∞ versus SMC∞-ko mice. a, b, Tcf21 expression in SMC lineage-labeled cells from SMC∞ (WT) and SMC∞-ko (KO) mice from all time points combined (n = 13 mice). a, Tcf21 expression for all WT cells (left; minimum = 0; maximum = 2.55; mean = 0.071) and all KO cells (right; minimum = 0; maximum = 1.97; mean = 0.004). b, Mean Tcf21 expression visualized for all SMC lineage-labeled WT and KO cells. c, Total Lgals3⁺ area in the lesion is reduced in SMC∞-ko mice. d, Cd68 immunohistochemistry quantification (left) and representative images (right). Scale bars represent 100 μm. e, Lesion area, normalized to the total vessel area. Data from c–e are after 16 weeks HFD, and analyzed using a two-sided Student’s t-test. Error bars indicate standard error.
Extended Data Fig. 4 | Joint clustering approach identifies human phenotypically modulated SMCs. a, Joint clustering of mouse and human datasets using canonical correlation analysis (CCA) as per the Seurat package. b, The shared mouse/human cluster containing bona fide SMC lineage-traced, phenotypically modulated SMCs (fibromyocytes) from the mouse is highlighted in red. c, Mouse cells in the shared mouse/human fibromyocyte cluster in b are highlighted in the independently clustered mouse dataset, confirming their location within the known fibromyocyte cell cluster. d, Human cells in the shared mouse/human fibromyocyte cluster in b are highlighted in the independently clustered human dataset, illustrating their location predominantly in the ‘fibromyocyte’ cluster (also shown in brown in Fig. 4d). e, All joint mouse/human clusters in a were mapped back to the human dataset. Agreement is identified in cell type assignment between the joint clustering approach and the independently clustered human dataset.
Extended Data Fig. 5 | Human phenotypically modulated SMCs. a, t-SNE visualization of cell types in the right coronary artery of four patients, overlaid with LUM expression. Expression levels are indicated by scales in the lower right. b, TNFRSF11B RNAscope staining in a human coronary artery section. Hybridization events are seen as red dots. e, A negative control RNAscope probe shows no staining. The images in b and c are representative of four experiments, and the scale bars represent 50 μm. d, Heat map representation of the Euclidean distance between cell cluster centroids in 20-dimensional principal component space, with the smallest distances in yellow and the largest distances in black. The relationship between ‘fibromyocyte’ and ‘fibroblast 2’ clusters is highlighted with white asterisks. The ‘fibromyocyte’, ‘SMC’ and the main ‘macrophage’ clusters are denoted by black asterisks. e, f, t-SNE visualization of cell types in the right coronary artery of four patients overlaid with CD68 expression (e) and TCF21 expression (f). g, UCSC Genome Browser shots of representative TCF21 ChIP-Seq peaks within the PRELP and MYH11 genes, which are highly correlated and anticorrelated, respectively, with TCF21 and the fibromyocyte phenotype. Images are from one ChIP-Seq experiment.
Extended Data Fig. 6 | Association of genome-wide significant CAD risk SNPs at the 6q23.2 locus with TCF21 expression. Seven SNPs in the 6q23.2 locus were associated with CAD at genome-wide significance. The association between risk and protective genotypes and TCF21 expression for each of these SNPs was determined using the GTEx database in CAD-relevant tissues and a cohort of 52 HCASMC lines. The number of independent tissue samples included for each SNP is indicated in the GTEx data (‘N’), and n = 52 cell lines for the HCASMC data. In each box plot, the middle line represents the median, the box represents the first to third quartile range, and whiskers represent 1.5× the interquartile range.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

☐ □ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐ □ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐ □ The statistical test(s) used AND whether they are one- or two-sided
☐ □ Only common tests should be described solely by name; describe more complex techniques in the Methods section.
☐ □ A description of all covariates tested
☐ □ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐ □ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐ □ For null hypothesis testing, the test statistic (e.g. \( F \), \( t \), \( r \)) with confidence intervals, effect sizes, degrees of freedom and \( P \) value noted
☐ □ Give \( P \) values as exact values whenever suitable.
☐ □ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐ □ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐ □ Estimates of effect sizes (e.g. Cohen's \( d \), Pearson's \( r \)), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

| Data collection | Leica Application Suite X software version 3.3.3 was used to acquire all immunohistochemistry, immunofluorescence and RNAscope images. |
|-----------------|---------------------------------------------------------------------------------------------------------------|
| Data analysis   | FASTQ files were processed with the CellRanger Pipeline (10X Genomics). Single-cell RNAseq data were analyzed using Seurat version 2.3.4 and R version 3.4.4. Immunohistochemistry data were quantified using ImageJ software (NIH). Immunohistochemistry and quantitative PCR data were analyzed using Prism version 8 (Graphpad software). Human coronary artery smooth muscle cell (HCASMC) line genome and transcriptome sequencing data were mapped using STAR version 2.4.0i. TCF21 ChIPseq sequencing data were mapped using BWA-MEM version 0.7.12, and ChIPseq peaks were called using MACS version 2. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

High throughput sequencing data (FASTQ) files for all scRNA-seq, CITE-seq and ChIP-seq, as well as cell-gene count matrices for all scRNAseq and CITE-seq experiments, have been deposited at Gene Expression Omnibus (GEO) with SuperSeries reference number GSE131780. These data were used to generate images in Figs. 1-5 and Extended Data Figs. 2-5. FASTQ files and processed data are also available from the corresponding author upon request.
**Field-specific reporting**

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

---

**Life sciences study design**

All studies must disclose on these points even when the disclosure is negative.

| Sample size | The number of mice used for immunohistochemistry experiments was based upon power calculations that take into account effect size and variance from previous publications investigating plaque vulnerability measures and fibrous cap phenotype (references available upon request). |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data exclusions | No data were excluded from the analyses, but on rare occasion some histology slides had folded/torn sections and could not be processed further. |
| Replication | We verified the reproducibility of the effect of Tcf21 knockout on smooth muscle cell phenotypic modulation using single cell RNAseq in a different conditional Tcf21 mouse model. This confirmed the effect we observed in the primary mouse model presented in the manuscript. In the expression quantitative trait locus analysis, we used multiple sets of coronary artery disease-associated polymorphisms and multiple eQTLs datasets, which yielded similar results. |
| Randomization | Mice were allocated into experimental groups based upon their genotypes. Genotypes in the two groups were identical except for the presence of an conditional Tcf21 null allele, and all mice were male, so no other co- variates were used. |
| Blinding | Researchers were blinded to group allocation (mouse genotypes) during analyses of immunohistochemistry data. Blinding was not possible during harvesting of mouse tissues for single-cell RNAseq or immunohistochemistry because it was necessary to preserve genotype information for each sample. |

---

**Reporting for specific materials, systems and methods**

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| Involved in the study |
|-----------------------|
| Antibodies |
| Eukaryotic cell lines |
| Palaeontology |
| Animals and other organisms |
| Human research participants |
| Clinical data |

### Methods

| Involved in the study |
|-----------------------|
| ChIP-seq |
| Flow cytometry |
| MRI-based neuroimaging |

---

**Antibodies**

**Antibodies used**

- Anti-sm22alpha (Rabbit polyclonal): Abcam #ab14106. Lot #GR194511-1. 1:300 dilution.
- Anti-Lgals3 (Rat monoclonal): Cedarlane Labs #CL8942AP, clone M3/38. Lot #1642220. 1:100 dilution.
- Anti-Cd68 (Rabbit polyclonal): Abcam #ab25212. Lot #GR300628-10. 1:400 dilution.
- Anti-Cd16/Cd32 (Rat monoclonal): BioLegend #101319. Lot #B269182. 1:100 dilution.
- Anti-Cd11b (Rat monoclonal): BioLegend #101265. Lot #B260350. 1:100 dilution.
- Anti-Cd64 (Mouse monoclonal): BioLegend #139325. Lot #B269325. 1:100 dilution.
- Anti-Cd86 (Rat monoclonal): BioLegend #105047. Lot #B269338. 1:100 dilution.
- Anti-F4/80 (Rat monoclonal): BioLegend #123153. Lot #B264480. 1:100 dilution.
- Anti-Cd68 (Rat monoclonal): BioLegend #137031. Lot #B273997. 1:100 dilution.
- Anti-Cd11c (Hamster monoclonal): BioLegend #117355. Lot #B278844. 1:100 dilution.
- Anti-TCF21(polyclonal): Sigma HPA013189. Lot #G113991. 14ug used.
- Anti-rabbit IgG secondary Ab (goat): Thermo Fisher #A21244. Lot #1386544. 1:500 dilution.

**Validation**

For anti-Lgals3 (#CL8942AP), validation was performed by the manufacturer using FACS staining with antibody versus isotype control in mouse macrophages, and was also independently validated for immunohistochemistry in mouse tissue. Per product data sheet is referenced in 15 publications. Anti-sm22alpha #ab14106 was validated for immunohistochemistry in mouse samples by the manufacturer and is referenced in 137 publications. Anti-Cd68 #ab25212 was validated in mouse samples for
**Eukaryotic cell lines**

| Cell line source(s) | Human coronary artery smooth muscle cell (HCASMC) lines were purchased from PromoCell (catalog #C-12511), Cell Applications (catalog # 350-05a), Lonza (catalog #CC-2583), Lifeline Cell Technology (catalog #FC-0031) and ATCC (catalog #PCS-100-021). |
| Authentication | Promocell: tested for cell morphology, adherence rate, and cell viability. Flow cytometric analyses for smooth muscle cell markers (smooth muscle alpha-actin) are carried out for each lot. Cell Applications: tested for smooth muscle specific alpha-actin positivity and for attachment, spreading, proliferation in growth medium. Lonza: Cells stain positive for smooth muscle alpha-actin and negative for von Willebrand (factor VIII) antigen after differentiation. Lifeline Cell Technology: von Willebrand Factor negative, smooth muscle alpha-actin positive after differentiation. ATCC: Von Willebrand factor positive and smooth muscle alpha-actin negative. |
| Mycoplasma contamination | Promocell: All cells have been tested for the absence of HIV-1, HIV-2, HBV, HCV, HTLV-1, HTLV-2 and microbial contaminants (fungi, bacteria, mycoplasma). Cell Applications: No bacteria, yeast, fungi, mycoplasma, virus. Lonza: HIV-1, hepatitis B and hepatitis C are not detected for all donors and/or cell lots. All cells are performance assayed and tested negative for mycoplasma, bacteria, yeast and fungi. Lifeline Cell Technology: negative for mycoplasma. Negative for bacterial and fungal growth. Negative for HIV-1, HIV-2, HBV and HCV by PCR. ATCC: mycobacterium negative. Bacterium and yeast negative. HBV, HCV, HIV-1 negative. |
| Commonly misidentified lines (See ICLAC register) | None |

**Animals and other organisms**

| Laboratory animals | All mice were on the C57BL/6, ApoE(-/-) background and analyzed at 8 weeks, 16 weeks and 24 weeks of age. Due to integration of the BAC transgene on the Y-chromosome, all mice used in this study were male. |
| Wild animals | This study did not involve wild animals. |
| Field-collected samples | This study did not involve samples collected from the field. |
| Ethics oversight | Stanford University Administrative Panel on Laboratory Animal Care (APLAC) |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

**Human research participants**

| Population characteristics | Age range of human subjects was 54-68 years, 3 male and one female. All patients had either ischemic or non-ischemic cardiomyopathy and were undergoing heart transplantation. |
| Recruitment | Consent was obtained from pre-operative patients by study coordinators from the Human Biorepository Tissue Research Bank. All patients had severe cardiomyopathy at the time of participation, but had varying degrees of atherosclerosis. |
| Ethics oversight | Stanford University Institutional Review Board (IRB) |

Note that full information on the approval of the study protocol must also be provided in the manuscript.
ChIP-seq

Data deposition
- Confirm that both raw and final processed data have been deposited in a public database such as GEO.
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
GEO #GSE131780

Files in database submission
Pooled_HCASM_CTCf21_ChIP_1.fq.gz, Pooled_HCASM_CTCf21_ChIP_2.fq.gz,
Pooled_HCASM_CTCf21_ChIP_1.fq.gz.Pooled_HCASM_CTCf21_ChIP_2.fq.gz.MACS2_peaks.bed

Genome browser session
https://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=tqlab&hgS_otherUserSessionName=tcf21_pooled

Methodology

Replicates
The pooled Tcf21 ChIPseq was compared to our previous Tcf21 ChIPseq data, performed with the same antibody, and exhibited excellent concordance.

Sequencing depth
Total 516,440,244 paired-end reads,Mapped 512,644,923 paired-end reads. After removal of duplicates, 43,727,785 paired-end reads were kept for analysis. All reads were 150bp paired-end reads.

Antibodies
A rabbit polyclonal anti-TCF21 antibody (Sigma #HPA013189, lot #G113991) was used for the ChIP-seq experiments.

Peak calling parameters
ChIPseq peaks were called with MACSv2 using default parameters.

Data quality
After initial peak calling by MACS, peaks were further selected by specifying a minimum fold-enrichment of 10 and a minimum log10 q-value of 60. Thus, all peaks are above 5-fold enrichment.

Software
Model-based analysis of ChIP-Seq (MACS) version 2 was used to analyze the ChIP-seq data.

Flow Cytometry

Plots
- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation
Immediately after sacrifice, mice were perfused with PBS. The aortic root and ascending aorta were excised, up to the level of the brachiocephalic artery. Tissue was washed x3 in PBS, placed into an enzymatic dissociation cocktail (2 units/mL Liberase TM (Sigma #5401127001), 2 units/mL elastase (Worthington #LS002279 ) in HBSS), and minced. After incubation at 37C for 1 hour, the cell suspension was strained and then pelleted by centrifugation at 500g for 5 minutes. The enzyme solution was then discarded and cells were resuspended in fresh HBSS.

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. The proximal to mid right coronary artery (RCA) was identified, excised, cleaned of peri-arterial fat, and then rinsed x3 in PBS. The tissue was cut into approximately 50mg sections, and each section was placed into an enzymatic dissociation cocktail (10.4 units/mL Liberase TM, 8 units/mL elastase (Sigma #E7885) in HBSS), and minced. After incubation at 37C for 1 hour with periodic agitation, the cell suspension was pipetted up and down to break up any remaining tissue. The cell suspension was strained and then pelleted by centrifugation at 500g for 5 minutes. The enzyme solution was then discarded and cells were resuspended in fresh HBSS.

Instrument
Cells were sorted on a BD Aria II instrument and a SONY SH800 instrument.

Software
FACS data were analyzed with FlowJo v10.

Cell population abundance
In the mouse experiments, smooth muscle cell lineage-traced tdTomato+ cells comprised approximately ~20-40% of the total cell population, depending on the sample. An example is shown in Supplemental Figure 1c and Supplemental Figure 2o-q.

Gating strategy
Cells were gated on forward/side scatter parameters to exclude small debris and then gated on forward scatter height vs.
Gating strategy

forward scatter area to exclude obvious doublet events. Events passing these criteria were then sorted into one of two 1.5mL Eppendorf tubes based upon tdTomato fluorescence levels. The gating strategy is also illustrated in Supplemental Figure 1c.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.