



# Stem Cell Pluripotency Genes *Klf4* and *Oct4* Regulate Complex SMC Phenotypic Changes Critical in Late-Stage Atherosclerotic Lesion Pathogenesis

**BACKGROUND:** Rupture and erosion of advanced atherosclerotic lesions with a resultant myocardial infarction or stroke are the leading worldwide cause of death. However, we have a limited understanding of the identity, origin, and function of many cells that make up late-stage atherosclerotic lesions, as well as the mechanisms by which they control plaque stability.

**METHODS:** We conducted a comprehensive single-cell RNA sequencing of advanced human carotid endarterectomy samples and compared these with single-cell RNA sequencing from murine microdissected advanced atherosclerotic lesions with smooth muscle cell (SMC) and endothelial lineage tracing to survey all plaque cell types and rigorously determine their origin. We further used chromatin immunoprecipitation sequencing (ChIP-seq), bulk RNA sequencing, and an innovative dual lineage tracing mouse to understand the mechanism by which SMC phenotypic transitions affect lesion pathogenesis.

**RESULTS:** We provide evidence that SMC-specific *Klf4*- versus *Oct4*-knockout showed virtually opposite genomic signatures, and their putative target genes play an important role regulating SMC phenotypic changes. Single-cell RNA sequencing revealed remarkable similarity of transcriptomic clusters between mouse and human lesions and extensive plasticity of SMC- and endothelial cell-derived cells including 7 distinct clusters, most negative for traditional markers. In particular, SMC contributed to a *Myh11*<sup>+</sup>, *Lgals3*<sup>+</sup> population with a chondrocyte-like gene signature that was markedly reduced with *SMC-Klf4* knockout. We observed that SMCs that activate *Lgals3* compose up to two thirds of all SMC in lesions. However, initial activation of *Lgals3* in these cells does not represent conversion to a terminally differentiated state, but rather represents transition of these cells to a unique stem cell marker gene-positive, extracellular matrix-remodeling, “pioneer” cell phenotype that is the first to invest within lesions and subsequently gives rise to at least 3 other SMC phenotypes within advanced lesions, including *Klf4*-dependent osteogenic phenotypes likely to contribute to plaque calcification and plaque destabilization.

**CONCLUSIONS:** Taken together, these results provide evidence that SMC-derived cells within advanced mouse and human atherosclerotic lesions exhibit far greater phenotypic plasticity than generally believed, with *Klf4* regulating transition to multiple phenotypes including *Lgals3*<sup>+</sup> osteogenic cells likely to be detrimental for late-stage atherosclerosis plaque pathogenesis.

Gabriel F. Alencar, PhD\*  
Katherine M. Owsiany, PhD\*  
Santosh Karnewar, PhD†  
Katyayani Sukhavasi, MSct  
Giuseppe Mocci, PhD  
Anh T. Nguyen, PhD  
Corey M. Williams, BS  
Sohel Shamsuzzaman, PhD  
Michal Mokry<sup>id</sup>, MD  
Christopher A.  
Henderson, BS  
Ryan Haskins, PhD  
Richard A. Baylis, PhD  
Aloke V. Finn, MD, PhD  
Coleen A. McNamara<sup>id</sup>, MD  
Eli R. Zunder, PhD  
Vamsidhar Venkata, PhD  
Gerard Pasterkamp<sup>id</sup>, MD  
Johan Björkegren<sup>id</sup>, MD,  
PhD  
Stefan Bekiranov<sup>id</sup>, PhD  
Gary K. Owens<sup>id</sup>, PhD

\*Drs Alencar and Owsiany contributed equally.

†Dr Karnewar and K. Sukhavasi contributed equally.

**Key Words:** coronary disease  
■ single nucleotide polymorphisms  
■ smooth muscle cell ■ smooth muscle differentiation ■ smooth muscle progenitor cell

Sources of Funding, see page 2058

© 2020 The Authors. *Circulation* is published on behalf of the American Heart Association, Inc., by Wolters Kluwer Health, Inc. This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution, and reproduction in any medium, provided that the original work is properly cited.

<https://www.ahajournals.org/journal/circ>

## Clinical Perspective

### What Is New?

- Smooth muscle cells in late-stage atherosclerotic lesions undergo transition through Lgals3<sup>+</sup> to several states, including inflammatory and Klf4-dependent osteogenic phenotypes.
- Transcription factors Klf4 and Oct4 bind putative target genes associated with genetic risk for coronary artery disease.

### What Are the Clinical Implications?

- Smooth muscle phenotypic switching produces cells that can be beneficial or detrimental to lesion stability and may be an important mechanism controlling risk of unstable atherosclerotic plaque and myocardial infarction or stroke.

Despite decades of research, we have little understanding of the origin and function of many cells that make up late-stage atherosclerotic lesions, as well as the mechanisms by which they control plaque stability and risk for plaque rupture leading to myocardial infarction or stroke.<sup>1,2</sup> Rigorous lineage tracing studies in mouse models,<sup>3–5</sup> and in human lesions using epigenetic marker analyses,<sup>6</sup> have shown that smooth muscle cells (SMCs) are far more abundant and play a much more complex role in lesion pathogenesis than predicted based on marker gene analyses. Shankman et al found that >80% of SMC in atherosclerotic plaque have lost expression of contractile markers like Acta2.<sup>3</sup> Furthermore, up to 30% of these have activated expression of Lgals3/Mac2, and a smaller percentage express markers of stem cells or myofibroblasts like Sca1/Ly6a or the platelet-derived growth factor (PDGF)- $\beta$  receptor. SMC-specific *Klf4* knockout (Myh11-Cre<sup>ERT2</sup> eYFP apoE Klf4<sup>ΔΔ</sup>, SMC<sup>Klf4-KO</sup>) resulted in lesions that were 50% smaller, exhibited evidence for increased plaque stability including a doubling in the Acta2<sup>+</sup> fibrous cap, and had a >60% decrease in SMC-derived Lgals3<sup>+</sup> cells.<sup>3</sup> As such, Klf4-dependent changes in SMC phenotype and subsequent effects appear to exacerbate lesion pathogenesis. In contrast, SMC-specific *Oct4* knockout (Myh11-Cre<sup>ERT2</sup> eYFP apoE Oct4<sup>ΔΔ</sup>, SMC<sup>Oct4-KO</sup>) resulted in opposite effects including increases in lesion size and evidence for reduced plaque stability including the nearly complete absence of an SMC-enriched Acta2<sup>+</sup> fibrous cap, reduced mature collagen content, increased lipid content, and increased intraplaque hemorrhage.<sup>4</sup> Recent work by Wirka et al used single-cell (sc) RNA sequencing (RNA-seq) in combination with lineage tracing to define the transcriptional signature of SMC-derived cells in atherosclerosis, detecting an Lgals3<sup>+</sup> cluster expressing genes for multiple ECM proteins.<sup>7</sup> However, their analyses

were performed on aortic root segments such that the majority of SMC and other cells analyzed were derived from the medial and adventitial layers, not lesions, thus severely limiting their sensitivity in detecting SMC lesion phenotypes. Moreover, their conclusion that SMCs give rise to a single so-called beneficial “fibrocyte” phenotype is incompatible with results of SMC-specific knockout studies clearly establishing that SMCs can play either a detrimental or beneficial role in plaque stability.<sup>3,4</sup> As such, further definition of SMC subsets within lesions is critical, with the hope of identifying factors and mechanisms that promote beneficial SMC phenotypic transitions as novel therapeutic targets.

To better define the cellular origins and phenotypic properties of SMC and non-SMC within atherosclerotic lesions, we used a combination of bulk and scRNA-seq of advanced brachiocephalic artery (BCA) lesions from SMC-specific lineage tracing apoE<sup>-/-</sup> mice with or without SMC specific conditional knockout of Klf4 or Oct4. Given the profound differences in lesion pathogenesis in these 2 knockout models, we hypothesized that studies would provide insights about not only the complexity of phenotypes exhibited by SMC, but also if these changes are likely to be beneficial or detrimental for late-stage plaque pathogenesis. Remarkably, we provide evidence that Klf4 and Oct4 control nearly opposite patterns of gene expression in SMC and based on in vivo ChIP-seq analyses have identified >80 potential Klf4 or Oct4 target genes that may impact SMC phenotypic transitions important in lesion pathogenesis. In addition, scRNA-seq studies on a unique dual recombinase lineage mouse generated by our laboratory and our previously published SMC-Klf4 knockout mice show that several SMC lesion phenotypes are derived from a subset of Lgals3<sup>+</sup> transitional state SMCs that initially exhibit an extracellular matrix remodeling phenotype but ultimately contribute to multiple transcriptomic clusters, including populations of osteogenic and proinflammatory state cells likely to be detrimental for lesion pathogenesis.

## METHODS

Data are available on request from the authors.

### Mice

All experiments followed guidelines of the University of Virginia Animal Care and Use Committee (Protocol 2500).

SMC<sup>Klf4</sup> and SMC<sup>Oct4</sup> mice were described previously.<sup>3,4</sup> Littermate controls were used for all studies. Rosa-tdTomato-eGFP mice were obtained from Jackson Labs (stock No. 026931). Myh11-Dre<sup>ERT2</sup> mice and Lgals3-internal ribosomal entry site-Cre mice were made by Cyagen Labs. The *myosin heavy chain 11* (*Myh11*) promoter was placed upstream of a Dre recombinase fused to an *estrogen receptor 2* (*ERT2*) domain. Mouse embryos were injected with this transgene, which randomly inserted into the genome. For the Lgals3-Cre mice, a herpesvirus internal ribosomal entry site followed by a Cre recombinase

was integrated into the endogenous *Lgals3* locus after the stop codon in exon 6. All mice were injected with 0.1 mL tamoxifen (1 mg/mL, Sigma No. T-5648) dissolved in peanut oil in 10 injections from 6 to 8 weeks of age. Mice were fed a Western diet containing 21% milk fat and 0.15% cholesterol (Tekland) for 18 weeks (10 weeks when noted). Because the *Myh11-Cre<sup>ERT2</sup>* is located on the Y chromosome, only male mice were analyzed in these studies. However, because *Myh11-Dre<sup>ERT2</sup>* does not have this limitation, both male and female mice were analyzed.

### Bulk RNA-seq

Total RNA was isolated using Trizol (Invitrogen) from the BCA and aortic arch region of *SMC-Klf4<sup>WT/WT</sup>*, *SMC-Klf4<sup>Δ/Δ</sup>*, *SMC-Oct4<sup>WT/WT</sup>*, and *SMC-Oct4<sup>Δ/Δ</sup>* (n=4 or 5) mice fed a high-fat Western diet for 18 weeks. The RNA library was prepared according to the Illumina RNA-Seq library kit.

One hundred nucleotide paired-end reads were mapped to the mm10 reference genome using STAR software version 2.4<sup>8</sup> against the mouse genome release M21 (GRCm38.p6) from Gencode. A table of gene counts was generated using FeatureCounts.<sup>9</sup> We used the DESeq2<sup>10</sup> Bioconductor R package to identify differentially expressed genes at a 5% false discovery rate (*P* value adjusted ≤ 0.05) using the Benjamini-Hochberg procedure to adjust *P* values. Gene set enrichment analysis was performed using Ingenuity Pathway Analysis.<sup>11</sup> Significantly enriched pathways were identified using a 5% false discovery rate cutoff, and their enrichment significance was quantified using  $-\log_{10}$  of *P* value adjusted. Data are presented as pathways downregulated or upregulated in knock-out mice compared with wild-type control mice.

### Klf4 and Oct4 ChIP-seq

Segments of the aorta from the arch to the aortic root and up to the carotid bifurcation isolated from 18-week Western diet-fed *SMC-Klf4<sup>WT/WT</sup>*, *SMC-Klf4<sup>Δ/Δ</sup>*, *SMC-Oct4<sup>WT/WT</sup>*, and *SMC-Oct4<sup>Δ/Δ</sup>* (n=15 each group) mice were snap-frozen in liquid nitrogen and then processed for ChIP assays using the Klf4- or Oct4-specific antibody (Figure 1 in the Data Supplement) and sequenced using Illumina TruSeq Chip Library Kit. Sequencing reads were aligned to the mouse genome release M21 (GRCm38.p6) using the BOWTIE alignment tool.<sup>12</sup> These aligned reads were then processed, involving removal of duplicate reads and format conversions using the SAMtools<sup>13</sup> suite and converted into bam/bai format and loaded in the Integrative Genomics Viewer for visualization. The reads were also converted to BED format, and peaks were identified using MACS2<sup>14</sup> callpeaks function against the mouse genome with a *q* value < 0.05.

### Atherosclerotic Plaque scRNA-seq

For murine studies, individual atherosclerotic plaques from the BCA were removed from underlying media with forceps and deposited into 1% BSA in PBS plus 1 μg/mL Actinomycin D (Gibco, No. 11805017). Two thousand cells in each group were targeted in Chromium 10X genomics libraries, which, after barcoding, were pooled and sequenced on the Illumina NextSeq, 150 cycle high-output.

Human atherosclerotic plaques were obtained with informed consent from 14 male and 4 female patients

undergoing a primary carotid endarterectomy in the Athero-Express Biobank Study (www.atheroexpress.nl) approved by the University Medical Center Utrecht Medical Ethical Committee. Viable cells were sorted 1 cell per well and immediately frozen at  $-80^{\circ}\text{C}$  until further processing using the Sorting and Robot-Assisted Transcriptome Sequencing (SORT-seq) protocol.<sup>15</sup> Cells were then pooled in 1 library, and the aqueous phase was separated from the oil phase, followed by in vitro transcription and library construction using the CEL-Seq2 protocol.<sup>16</sup> Detailed cell isolation protocols for humans and mice are in the Methods in the Data Supplement.

Gene-barcode matrices were analyzed in R using Seurat v3.<sup>17</sup> Cells were filtered for 200 to 5,000 reads per unique molecular identifiers, ≤ 10% mitochondrial and < 5% hemoglobin gene content. Significant principal components of variation were calculated using JackStraw test with 10,000 repetitions, and clusters were calculated with 19 principal components of variation. Raw data are available at GSE150644. Differential expression analysis was done using Model-Based Analysis of Single-Cell Transcriptomics (MAST). Code is available on request.

### Immunofluorescent Staining and Quantification

Paraffin BCA sections were deparaffinized, rehydrated, antigen-retrieved (citrate), and blocked with fish skin gelatin PBS (6 g/L) containing 10% horse serum for 1 hour at room temperature. Frozen sections were permeabilized with acetone at  $-20^{\circ}$  for 10 minutes. Slides were incubated with the antibodies shown in Table 1 in the Data Supplement. Slides were imaged using a Zeiss LSM700 confocal microscope to acquire a series of z-stack images at 1-μm intervals. Maximal intensity projection was used to generate the representative images included in the figures, and Fiji was used to process and format images.

### Statistical Analysis

Statistics were performed using GraphPad Prism 7. For comparison of 2 groups of continuous variables with normal distribution and equal variances, 2-tailed unpaired Student *t* tests (with additional Welch correction for unequal variances) were performed with a significance threshold of *P* ≤ 0.05. For multiple group comparison, we performed 1-way or 2-way ANOVA followed by the Sidak method of multiple pairwise comparisons. The number of mice used for each analysis is indicated in the figure legends, selected by a power analysis to detect a 30% change with 10% error and 95% confidence.

## RESULTS

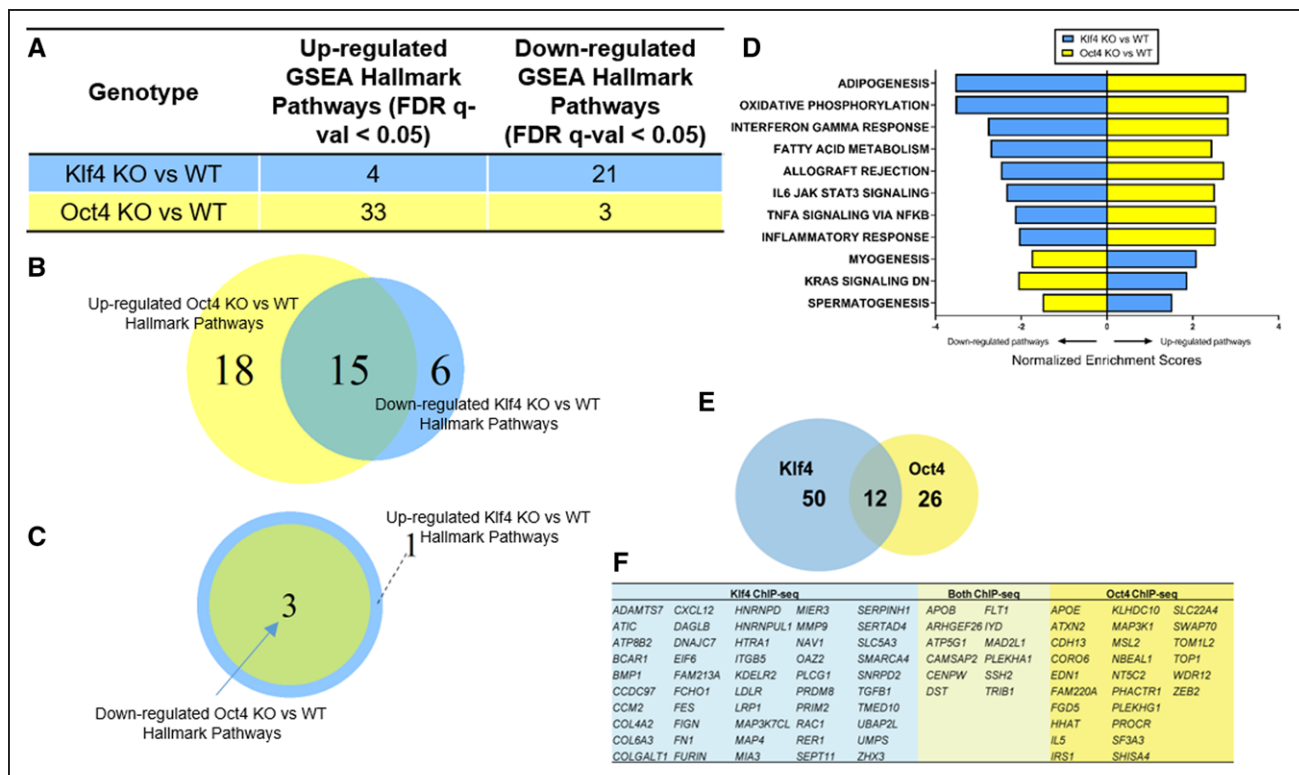
Klf4 and Oct4 control opposite patterns of gene expression in late-stage atherosclerotic lesions and may regulate genes associated with human coronary artery disease (CAD) risk.

To study global changes in gene expression that accompanied our atherosclerotic phenotypes of Klf4 and Oct4 knockout in SMCs, we conducted Klf4 and Oct4 ChIP-seq as well as bulk RNA-seq on the BCA lesion area of *SMC<sup>Klf4</sup>* and *SMC<sup>Oct4</sup>* knockout and wild-type mice after

18 weeks of Western diet (Figure I in the Data Supplement). By comparing ChIP-seq data between wild-type littermate control mice and our knockout mice, we were able to distinguish putative binding targets for each transcription factor in SMC, as well as any secondary changes in other cell types. Consistent with our previous data that  $SMC^{Klf4}$  and  $SMC^{Oct4}$  knockout have opposite phenotypes in terms of lesion morphology and cell composition, gene set enrichment analysis of bulk RNA-seq data using the MSigDb showed >70% of the down-regulated hallmark and Kyoto Encyclopedia of Genes and Genomes pathways in  $SMCKlf4-\Delta/\Delta$  versus  $SMCKlf4^{WT/WT}$  were up-regulated in  $SMC^{Oct4-\Delta/\Delta}$  versus  $SMC^{Oct4^{WT/WT}}$ , and nearly all up-regulated pathways in  $SMCKlf4-\Delta/\Delta$  versus  $SMCKlf4^{WT/WT}$  were down-regulated in the SMC Oct4 knockout model (Figure 1A–1C, Figure II in the Data Supplement). Ingenuity Pathway Analysis of the ChIP-seq data showed that Oct4 putative targets in SMC were enriched for genes involved in pluripotency, cancer, and migration, whereas Klf4 putative targets in SMC were enriched for genes involved in leukocyte recruitment, actin regulation, and

extracellular matrix organization (Figure IIIA–IIID in the Data Supplement, and example of ChIP-seq peaks, Figure III I and IIIJ in the Data Supplement). Of major interest, 88 genes linked with 39% (64/163) of accepted human CAD genome-wide association study (GWAS) loci<sup>18</sup> were identified as the nearest gene binding target in  $SMC^{Klf4}$  or  $SMC^{Oct4}$  murine atherosclerotic lesions (Figure 1E and 1F,  $P$  value of gene overlap <0.01 and almost 2-fold [1.94 fold] higher than expected by chance). Extensive additional studies would be required to determine if Klf4 or Oct4 directly regulate or modulate any of these candidate GWAS loci. However, we have shown that approximately one third of these genes are regulated by Klf4 or Oct4 using an in vitro model of cholesterol loading ( $SMC^{Klf4}$ ) or hypoxia plus POVPC (1-palmitoyl-2-(5-oxo-valeryl)-sn-glycero-3-phosphorylcholine) ( $SMC^{Oct4}$ , Figure IIIE and IIIF in the Data Supplement).

To further investigate the importance of *Klf4* and *Oct4* in human disease, we performed network analysis on the human STARNET database and identified a *Klf4* module of the atherosclerotic wall (Figure IV in the



**Figure 1.** Bulk RNA-seq and  $SMC^{Oct4}$  and  $SMC^{Klf4}$  ChIP-seq analysis of advanced BCA lesion areas from  $SMC$ -specific *Klf4* vs *Oct4* knockout apoE<sup>-/-</sup> mice showed virtually opposite genomic signatures and identify candidate  $SMC$  *Klf4* or *Oct4* target genes that may play an important role in late-stage atherosclerotic lesion pathogenesis.

RNA-seq was performed on the BCA regions from 18-week Western diet-fed  $SMCKlf4-\Delta/\Delta$  vs  $SMCKlf4^{WT/WT}$  and  $SMC^{Oct4-\Delta/\Delta}$  vs  $SMC^{Oct4^{WT/WT}}$  mice. **A**, GSEA analysis showing MSigDB Hallmark pathways from comparing KO vs WT of both lines are shown. **B**, Venn diagram showing the common pathways between downregulated  $SMCKlf4-\Delta/\Delta$  vs  $SMCKlf4^{WT/WT}$  and upregulated in  $SMC^{Oct4-\Delta/\Delta}$  vs  $SMC^{Oct4^{WT/WT}}$ . **C**, Venn diagram showing the common pathways between upregulated  $SMCKlf4-\Delta/\Delta$  vs  $SMCKlf4^{WT/WT}$  and downregulated in  $SMC^{Oct4-\Delta/\Delta}$  vs  $SMC^{Oct4^{WT/WT}}$ . **D**, Bar graph showing the top 8 common pathways from **B** and all common pathways in **C** differentially regulated pathways between  $SMCKlf4-\Delta/\Delta$  vs  $SMCKlf4^{WT/WT}$  and  $SMC^{Oct4-\Delta/\Delta}$  vs  $SMC^{Oct4^{WT/WT}}$  analyses. **E**, ChIP-seq was done on the BCA regions from 18-week Western diet-fed  $SMCKlf4-\Delta/\Delta$  vs  $SMCKlf4^{WT/WT}$  and  $SMC^{Oct4-\Delta/\Delta}$  vs  $SMC^{Oct4^{WT/WT}}$  mice (n=15). The Venn diagram shows overlap of putative “closest gene” binding partners of Klf4 and Oct4 in SMC (ie, those binding partners that are lost in the  $SMCKlf4-\Delta/\Delta$  compared with  $SMCKlf4^{WT/WT}$ ) that are also human CAD GWAS loci.  $P$  value of gene overlap <0.01. **F**, Table showing genes present in the Venn diagram in **E**. BCA indicates brachiocephalic artery; CAD, coronary artery disease; ChIP-seq, chromatin immunoprecipitation sequencing; FDR, false discovery rate; GSEA, gene set enrichment analysis; GWAS, genome-wide association study; KO, knockout; RNA-seq, RNA sequencing; SMC, smooth muscle cell; and WT, wild-type.

Data Supplement). This module was strongly associated with CAD (preoperative angiographically assessed DUKE score) and plasma low-density lipoprotein levels. In addition, genes present in this network module are upregulated in patients with CAD compared with CAD negative controls. Taken together, these results suggest that *Klf4* is not only a GWAS CAD gene<sup>19</sup> but might also be regulating several previously reported CAD-associated gene loci.

### scRNA-seq Analysis of Advanced BCA Lesions Identified 14 Distinct Cell Clusters Including 7 of SMC Origin

To better understand the heterogeneity of phenotypes that SMC can acquire within atherosclerotic lesions, we performed scRNA-seq analyses on advanced BCA lesions of our Myh11-Cre<sup>ERT2</sup> Rosa-eYFP apoE<sup>-/-</sup> SMC lineage tracing mice and Cdh5-Cre<sup>ERT2</sup> Rosa-eYFP apoE<sup>-/-</sup> endothelial cell lineage tracing mice fed a Western diet for 18 weeks. Contrary to previous scRNA-seq analysis in the field,<sup>20–22</sup> instead of analyzing whole atherosclerotic aortas, we isolated late-stage lesions microdissected from the BCA. To efficiently detect transcriptome changes in lineage-traced SMC or EC and to avoid any issues with read depth on eYFP (enhanced yellow fluorescent protein) transcript detection, we also sorted lineage-traced eYFP<sup>+</sup> cells from these late-stage lesions. We also analyzed a healthy aortic sample from chow-fed 8-week-old SMC lineage tracing apoE<sup>-/-</sup> mice directly after inducing SMC lineage tracing to provide a baseline scRNA-seq profile of medial (nonlesion) cells. Clustered together, results showed 14 different uniform manifold approximation and projection transcriptomic clusters including clusters corresponding to macrophages, T cells, SMCs, and endothelial cells identified by traditional markers (Figure 2A–2C). Using data from Myh11-lineage traced sorted cells, we observed that SMCs were the major source of at least 7 different clusters (Figure 2F). Interestingly, most of the clusters containing lineage-traced SMCs do not express *Myh11* (Figure 2C) or other SMC contractile protein markers, but instead express markers such as *Vcam1*, *Dcn/Lum*, *Lgals3*, *Spp1*, and *Sox9* (Figure V in the Data Supplement). Using Reactome and gene ontology pathway analysis of the top 100 genes in each cluster (Figure VI in the Data Supplement, Excel File I in the Data Supplement), we broadly characterized these clusters as inflammatory (clusters 4 to 5, “immunoregulatory interactions”), ECM-rich (cluster 6, “extracellular matrix organization” and “degradation of extracellular matrix”), and osteogenic (cluster 7, “chondrocyte differentiation” and “bone mineralization”).

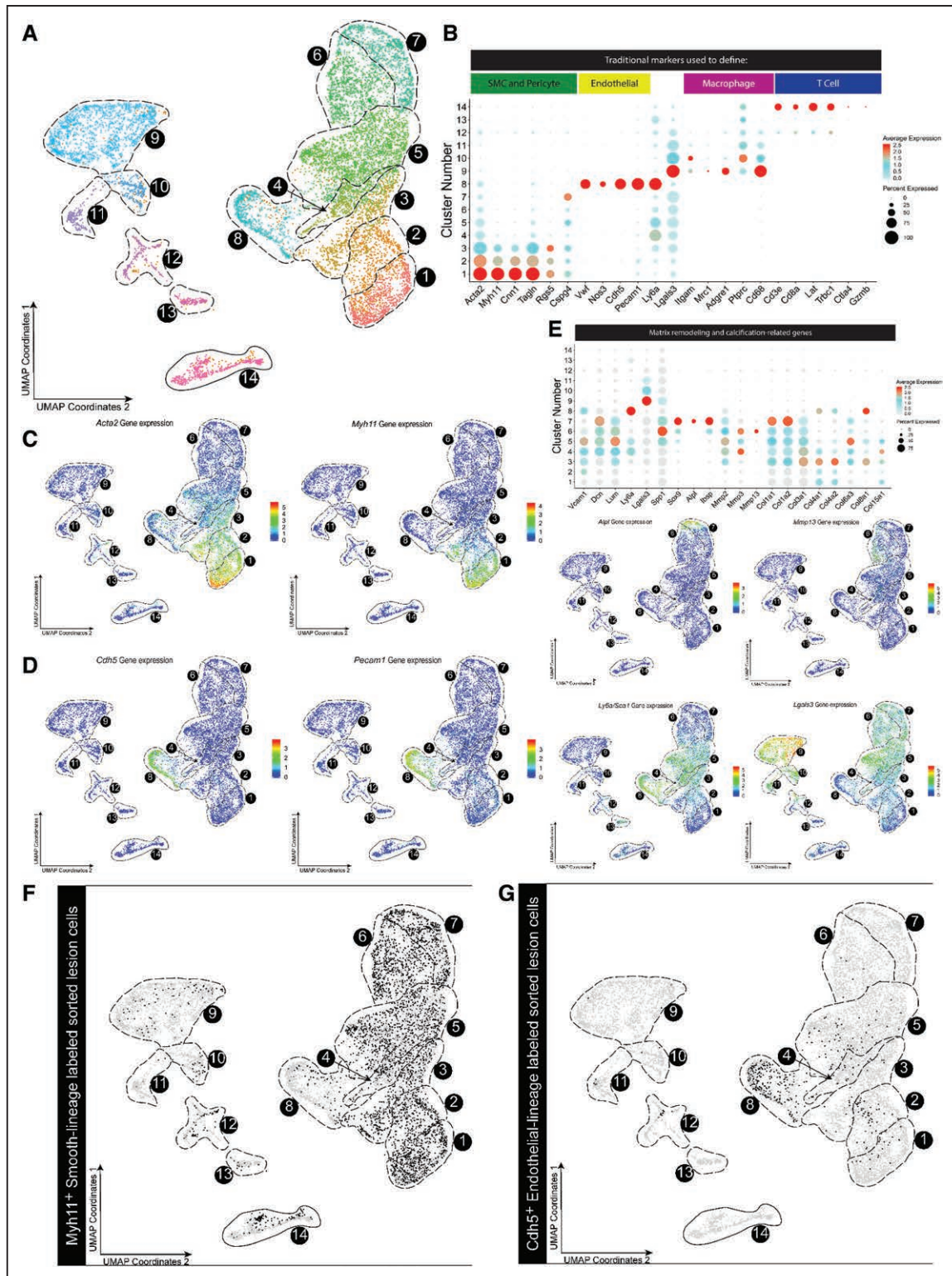
Similarly, we observed endothelial cells expressing SMC marker genes in clusters 1 to 3 that are likely related to endothelial-mesenchymal transition, including activation of *Acta2* and *Vimentin*. Interestingly, there

were also EC-derived cells in clusters 2 to 5, consistent with phenotypes related to phenotypically modified SMC (Figure 2B and 2G, Figure VII in the Data Supplement). We did observe some SMC-derived cells in clusters defined by canonical macrophage markers; however, it has been reported that single-cell methodologies undersample macrophages in general by up to 65%, which may explain the low abundance of these cells.<sup>20</sup> We also observed some SMC-derived cells in the T cell island (cluster 14), consistent with studies by Hansson et al<sup>23</sup> showing that SMCs can take on antigen presentation properties within lesions. Taken together, these results suggest that cells from multiple origins exhibit extraordinary plasticity during lesion development, with cells acquiring alternative phenotypes to adapt to the extreme lesion environment.

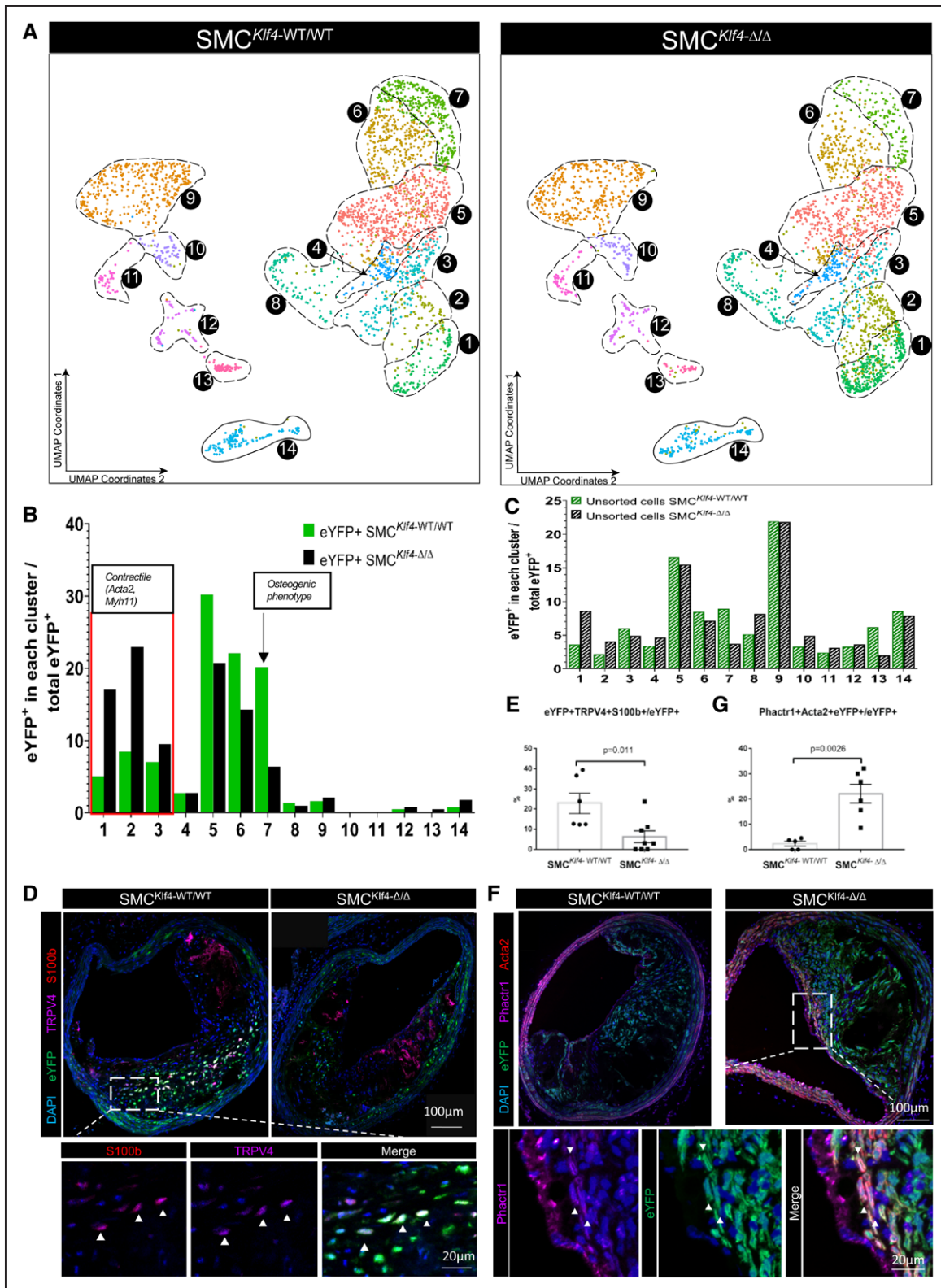
### scRNA-seq Analysis of Advanced BCA Lesions From SMCKlf4<sup>-Δ/Δ</sup> Versus SMCKlf4<sup>-WT/WT</sup> apoE<sup>-/-</sup> Mice Identified New Klf4-Dependent SMC Phenotypes

To determine mechanisms by which SMC conditional knockout of *Klf4* induced marked reductions in lesion size and increased indices of plaque stability, we performed scRNA-seq analyses on advanced BCA lesions from our SMCKlf4<sup>-Δ/Δ</sup> versus SMCKlf4<sup>-WT/WT</sup> apoE<sup>-/-</sup> mice in the same manner as in Figure 2. Results showed that advanced BCA lesions from SMCKlf4<sup>-Δ/Δ</sup> mice had a near doubling in the proportion of SMC-derived lesion cells expressing traditional markers in clusters 1 and 2 (Figure 3A and 3B). Furthermore, SMC<sup>Klf4</sup> knockout resulted in a 4-fold decrease in the proportion of SMC-derived cells transitioning to an Lgals3<sup>+</sup> osteogenic phenotype in cluster 7 (Figure 3A and 3B), characterized by expression of *Sox9*, *Runx2*, *Cyt11*, *Ibsp*,<sup>24</sup> and *Alpl*, which is a key enzyme driving an osteogenic response in cells.<sup>25–27</sup> In addition, differential gene expression analysis showed increased expression of *Acta2*, *Myh11*, and other SMC contractile markers, and a decrease in expression of *Spp1*, *Lgals3*, and *Sox9* in the SMCKlf4<sup>-Δ/Δ</sup> versus SMCKlf4<sup>-WT/WT</sup> advanced BCA lesions (Figure VIII in the Data Supplement, Excel File II in the Data Supplement with CAD loci-associated genes highlighted). There were also changes in non-SMC populations with SMCKlf4<sup>-Δ/Δ</sup> knockout, including a 3-fold increase in cluster 13, marked by histocompatibility genes associated with activation of B cells and T cells.

To determine if transcriptomic clusters observed by scRNA-seq can also be identified at the protein level, we performed mass cytometry (CyTOF) analyses on cells isolated from advanced BCA lesions from SMCKlf4<sup>-Δ/Δ</sup> versus SMCKlf4<sup>-WT/WT</sup> mice using an antibody panel of 20 conventional markers supplemented with 11 novel markers identified in our scRNA-seq analyses (Figure IX and Table II in the Data Supplement). Further, to



**Figure 2.** scRNA-seq analysis of healthy aorta and advanced BCA lesions identified 14 distinct cell clusters including 7 predominately of SMC origin. **A**, Cells were harvested for scRNA-seq from microdissected advanced BCA lesions from SMC lineage-traced (*Myh11-Cre<sup>ERT2</sup> Rosa-eYFP apoE<sup>-/-</sup>*) or endothelial lineage-traced (*Cdh5-Cre<sup>ERT2</sup> Rosa-eYFP apoE<sup>-/-</sup>*) 18-week Western diet-fed mice, as well as healthy 8-week-old *apoE<sup>-/-</sup>* SMC lineage-traced mouse aorta as a control. In aggregate, these data are the results of 7 experiments, 19 Chromium 10X libraries, and 23 mice. Results are shown as a UMAP colored by cluster. **B**, Dot plot of traditional markers used for cell identity. **C** and **D**, Feature plots showing gene expression of *Myh11*, *Acta2* (**C**) and *Cdh5* and *Pecam1* (**D**), as representative genes of classical markers for SMC and endothelial cells, respectively. *Myh11* was detected only in clusters 1 to 2, whereas *Cdh5* was detected exclusively in cluster 8. **E**, Dot plots and feature plots of selected genes representing 3 SMC phenotypic groups, stem/inflammatory, ECM remodeling, and osteogenic. Further plots and pathway analysis justifying the groups can be found in Figures V and VI in the Data Supplement. **F** and **G**, UMAP of mouse cells showing (F) *Myh11-Cre<sup>ERT2</sup> Rosa-eYFP apoE<sup>-/-</sup>* (black dots) or (G) *Cdh5-Cre<sup>ERT2</sup> Rosa-eYFP apoE<sup>-/-</sup>* lineage-traced cells from BCA plaques of mice fed a Western diet for 18 weeks (black dots). BCA indicates brachiocephalic artery; scRNA-seq, single-cell RNA sequencing; SMC, smooth muscle cell; and UMAP, uniform manifold approximation and projection.



**Figure 3.** scRNA-seq of advanced BCA lesions from SMC-specific *Klf4* knockout vs wildtype littermate control mice identified multiple *Klf4*-dependent clusters.

**A**, Cells from advanced BCA lesions were isolated from  $SMCKlf4^{WT/WT}$  and  $SMCKlf4^{\Delta/\Delta}$  mice after 18 weeks of Western diet and submitted for scRNA-seq. UMAP analysis shows cells from  $SMCKlf4^{WT/WT}$  on the left and from  $SMCKlf4^{\Delta/\Delta}$  cells on the right. The analysis includes 5469 cells from 2 different experiments (Chromium 10x runs) and 11 animals. **B**, Percentage of SMC-lineage traced cells (eYFP<sup>+</sup> cells sorted from BCA plaque) in each UMAP cluster. **C**, Percentage of non-SMC lineage-tracing cells (unsorted cells) in each UMAP cluster from each respective group described in **A**. **D–G**, BCA lesions from  $SMCKlf4^{WT/WT}$  and  $SMCKlf4^{\Delta/\Delta}$  mice were stained for selected immunofluorescent markers from scRNA-seq clusters. **D**, Costaining for TRPV4 and S100b (markers present in cluster 7) with eYFP in  $SMCKlf4^{WT/WT}$  and  $SMCKlf4^{\Delta/\Delta}$  animals. Images are based on confocal microscopy z-stacks of 1  $\mu$ m thickness on 10- $\mu$ m sections of BCA lesions. (Continued)

**Figure 3 Continued.** The top left panel shows a maximum intensity projection  $\times 20$  zoom. A close-up view of eYFP<sup>+</sup>TRPV4<sup>+</sup>S100b<sup>+</sup> cells is shown in a 1- $\mu$ m individual z-stack in a WT animal. **E**, Quantification of the frequency of triple-positive (eYFP<sup>+</sup> TRPV4<sup>+</sup> S100b<sup>+</sup>) cells as a percent of total eYFP<sup>+</sup> SMC-derived cells in SMCKlf4<sup>WT/WT</sup> and SMCKlf4 <sup>$\Delta\Delta$</sup>  (n=6–8, error bars show SEM). **F**, Results of costaining for eYFP and Phactr1 (a marker of cluster 3) in advanced BCA lesions from SMCKlf4<sup>WT/WT</sup> and SMCKlf4 <sup>$\Delta\Delta$</sup>  animals. The top panels show  $\times 20$  images. The bottom panels show a close-up image of a BCA lesion from SMCKlf4 <sup>$\Delta\Delta$</sup>  animal. **G**, Quantification of the frequency of double-positive (eYFP<sup>+</sup> Phactr1<sup>+</sup>) cells as a percent of total eYFP<sup>+</sup> SMCs in advanced BCA lesions from SMCKlf4<sup>WT/WT</sup> and SMCKlf4 <sup>$\Delta\Delta$</sup>  mice (n=6, error bars show SEM). BCA indicates brachiocephalic artery; scRNA-seq, single-cell RNA sequencing; SMC, smooth muscle cell; UMAP, uniform manifold approximation and projection; and WT, wild-type.

determine the location of these clusters within intact lesions, we stained sections of advanced mouse BCA lesions with antibodies to TRPV4, S100b, Sox9 (cluster 7), and Phactr1 (cluster 2). High-resolution confocal microscopy identified eYFP<sup>+</sup>Trpv4<sup>+</sup>S100b<sup>+</sup> cells (Figure 3D and 3E) in the body of the lesions whose frequency was significantly decreased by >50% in SMCKlf4 <sup>$\Delta\Delta$</sup>  versus SMCKlf4<sup>WT/WT</sup>. In contrast, eYFP<sup>+</sup>Acta2<sup>+</sup>Phactr1<sup>+</sup> cells (Figure 3F and 3G) were increased 10-fold in SMCKlf4 <sup>$\Delta\Delta$</sup>  versus SMCKlf4<sup>WT/WT</sup>.

Taken together, our results indicate that loss of Klf4 in SMC is associated with marked decreases in lesion SMCs that exhibit an Lgals3<sup>+</sup> osteogenic phenotype, which is of major interest given our previous studies in SMC Klf4 knockout mice that show loss of Lgals3-expressing SMC as well as decreased lesion size and increased fibrous cap thickness.<sup>3</sup> Together with our finding that SMC<sup>Klf4</sup> may regulate genes involved in genetic risk for CAD, this implies that Lgal3<sup>+</sup> SMCs have negative implications for atherosclerosis pathogenesis. In addition, our scRNA-seq data indicate that *Lgals3* activation is not a specific marker of transition of SMC to a macrophage-like state but rather is first observed in transition clusters (Figure 2E) along with *Ly6a/Sca1* and *Vcam1*. Further, knockdown of *Lgals3* in cultured SMCs prevented PDGF-induced expression of *Sca1* (Figure X in the Data Supplement), suggesting *Lgals3* marks a key transitional state.

### Studies With a Novel SMC<sup>Dual Lineage</sup> Tracing Mouse Show That >60% of Lesion SMCs Activate *Lgals3* and Give Rise to Multiple Downstream Nonmacrophage SMC Clusters

To rigorously define the phenotype and potential functional properties of SMC-derived Lgals3<sup>+</sup> lesion cells, we generated a novel dual lineage tracing model that uses sequential activation of Dre and Cre recombinases to track the transitions of SMC from Myh11<sup>+</sup> to Lgals3<sup>+</sup> states in adult mice. In brief, the Myh11 promoter drives a tamoxifen-inducible Dre recombinase, which removes a roxed stop cassette in the *Rosa* locus in front of a tdTomato-STOP-eGFP (enhanced green fluorescent protein) fluorescent reporter and a stop cassette on the *Lgals3* locus in front of an internal ribosomal entry site-Cre. If this cell later goes on to express *Lgals3*, it will express Cre and remove the *tdTomato* flanked by loxP sites, allowing expression of eGFP. Figures XI and

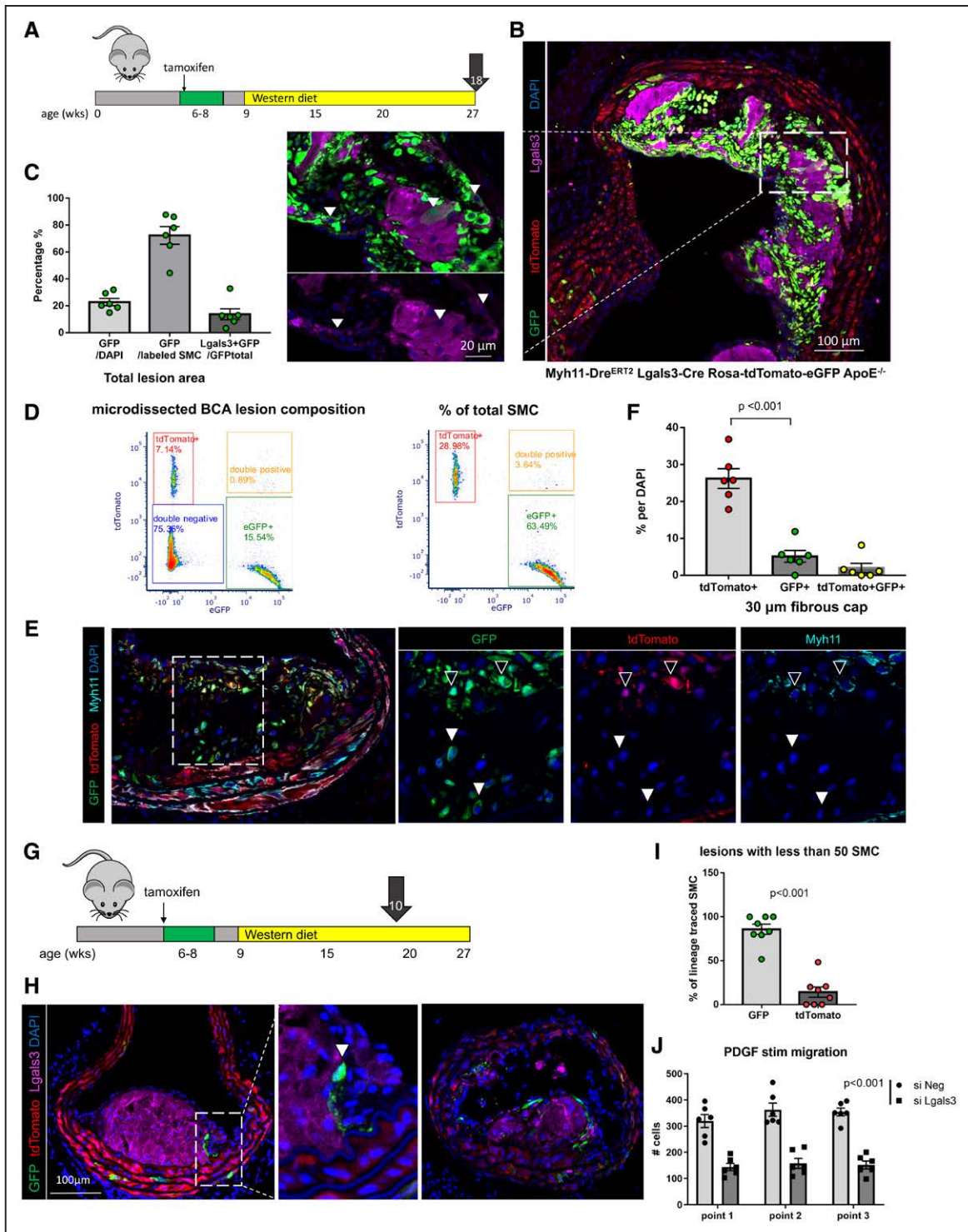
XII in the Data Supplement provide an overview and extensive validation of this mouse model. At baseline, medial SMCs of the aorta are tdTomato<sup>+</sup> only. However, after 18 weeks of Western diet, eGFP<sup>+</sup> cells accumulate in atherosclerotic plaques of the BCA, aortic arch, and renal arteries (Figures 4A, 4B, and 5E). These eGFP<sup>+</sup> cells make up 60% to 80% of all advanced lesion SMCs measured by both flow cytometry and high-resolution confocal microscopy (Figure 4C and 4D). Original SMC lineage tracing placed the number of Lgals3<sup>+</sup> SMC-derived cells at 30%<sup>3</sup>; however, because our dual lineage tracing system is cumulative, SMCs that expressed Lgals3 at any time will remain eGFP<sup>+</sup>. Indeed, only 25% of eGFP<sup>+</sup> SMCs in lesions continued to show surface Lgals3 immunostaining (Figure 4C).

To test the hypothesis that SMCs transitioned through Lgals3 take on multiple phenotypes in lesions, we performed scRNA-seq analyses on advanced BCA plaques from our Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup> (SMC<sup>Dual Lineage</sup>) mice fed a Western diet for 18 weeks, along with GFP<sup>+</sup> cells sorted from lesions and tdTomato<sup>+</sup> sorted from whole aortas with plaque to identify clusters containing SMC-derived cells (Figure 5A). Consistent with Lgals3 expression in Figure 2, the lower clusters contain tdTomato<sup>+</sup> cells expressing multiple contractile SMC markers, whereas more than half of all eGFP<sup>+</sup> cells localize in clusters 5 to 7 expressing *Lgals3*, inflammatory genes (*Ccl2*, *Vcam1*), ECM genes (*Dcn*, *Lum*, *Mmp3*), or *Sox9*, *Runx2*, and *Spp1*, the signature of the osteogenic phenotype cells decreased with SMC<sup>Klf4</sup> knockout (Figures 3B and 5C, Figure V in the Data Supplement). We further validated the existence of these cells using high-resolution confocal analysis, showing that tdTomato<sup>+</sup> cells dominate Acta2<sup>+</sup> SMCs but GFP<sup>+</sup> cells dominate Sox9<sup>+</sup> SMCs (Figure 5D and 5E). Taken together, our results suggest that instead of a terminal marker of SMC switched to a macrophage-like state as originally postulated,<sup>3,5</sup> *Lgals3* activation in these cells appears to be a marker of an earlier transitional state from which cells may subsequently exhibit multiple different phenotypes.

### SMCs Transitioned Through *Lgals3* Dominate Initial Investment of SMCs Into Lesions

In our early analyses of SMC<sup>Dual Lineage</sup> mice, there were a small number of tdTomato<sup>+</sup>eGFP<sup>+</sup> double-positive cells observed by flow cytometry (2%) and confocal





**Figure 4.** A large fraction of Myh11-tdTomato<sup>+</sup> SMCs phenotypically transition through an Lgals3 activation step and accumulate within advanced BCA lesions of Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup> mice.

**A**, Experimental design; SMC<sup>Dual Lineage</sup> (Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup>) mice were injected with tamoxifen at 6 to 8 weeks of age and subsequently placed on Western diet for 18 weeks to induce advanced atherosclerosis. Video of mouse model is available at <https://www.cvr.virginia.edu/Owens/educationaltopics.html>. **B**, Sections of the BCA at the origin from the aorta in an SMC<sup>Dual Lineage</sup> mouse, immunostained for GFP, tdTomato, and Lgals3, were imaged with confocal microscopy z-stacks of 1  $\mu$ m thickness. Picture at right shows the maximum-intensity projection image of the 10- $\mu$ m tissue thickness, and the close-up view shows eGFP<sup>+</sup> Lgals3<sup>+</sup> cells within a 1- $\mu$ m individual z-stack. **C**, Quantification of cells positive for eGFP or Lgals3 by single-cell counting in 1- $\mu$ m z-stacks of advanced BCA lesions from Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup> mice (n=6, error bars show SEM). **D**, Flow-cytometric evaluation of cells freshly isolated from microdissected BCA lesions from Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup> mice (n=4). These analyses included gating on live, single cells with fluorescence-minus-1 controls. Results show clear resolution of tdTomato<sup>+</sup>, eGFP<sup>+</sup>, and dual tdTomato<sup>+</sup> eGFP<sup>+</sup> (transition state) cells with normalization to the total number of lesion cells in the left panel or lineage-traced SMC in the right panel. **E**, Immunostaining of BCA lesions from Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup> 18-week Western diet-fed mice for GFP, tdTomato, and Myh11 reveals rare tdTomato<sup>+</sup>GFP<sup>+</sup> cells in (Continued)

**Figure 4 Continued.** the fibrous cap (open arrows), overlying GFP<sup>+</sup>-only cells (filled arrows). **F**, Quantification of the proportion of cells in the fibrous cap of advanced BCA lesions that were tdTomato<sup>+</sup>, GFP<sup>+</sup>, or double-positive as a percent of total DAPI-stained fibrous cap cells. The fibrous cap was defined as the 30- $\mu$ m lesion area underlying the lumen (n=6, error bars show SEM; the *P* value refers to unpaired *t* test between tdTomato<sup>+</sup> and GFP<sup>+</sup> cells). **G**, Experimental design; Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup> mice were injected with tamoxifen between 6 and 8 weeks of age and put on Western diet for 10 weeks to induce early atherosclerotic lesions. Ten animals were sampled at 4 locations in the BCA and imaged to capture early-stage lesions with <50 SMCs based on counting of lineage-traced SMC-derived cells in the entire lesion. **H**, Immunostaining for GFP, tdTomato, and Lgals3 by confocal microscopy as in **A**. **I**, Quantification of GFP<sup>+</sup> and tdTomato<sup>+</sup> cells in lesions of <50 lineage-traced SMCs total (n=8, error bars show SEM, *P*<0.01 by unpaired *t* test). **J**, SMCs isolated from Myh11 lineage tracing mice were treated with siLgals3 or a scrambled control siNeg, serum starved, and scratched with a pipet tip in 10% serum media + 10  $\mu$ g/mL PDGF-BB. Light microscopic images of 3 points along the scratch were each taken at 0, 24, and 48 hours, and cells within the scratch area defined at time 0 hours were quantified (n=6, error bars show SEM). BCA indicates brachiocephalic artery; GFP, green fluorescent protein; PDGF, platelet-derived growth factor; siLgals3, small interfering RNA to Lgals3; and SMC, smooth muscle cell.

microscopy within plaques (Figure 4D and 4E). Given that the half-life of tdTomato is 3 and a half days,<sup>28</sup> we reasoned that these cells represented SMC currently in the process of transition (ie, cells have excised the *tdTomato* sequence and activated eGFP but not yet fully degraded the tdTomato protein). We also were able to observe tdTomato<sup>+</sup>eGFP<sup>+</sup> cells in explants of aortas that lose tdTomato<sup>+</sup> status over time (Figure XIII in the Data Supplement, Movies I and II in the Data Supplement) and detected double-positive cells by CyTOF (Figure XIV in the Data Supplement). We also made sure that tdTomato<sup>+</sup>eGFP<sup>+</sup> cells did not represent cells that were poorly recombined at only 1 Rosa locus by observing this phenomenon in animals with only 1 Rosa-tdTomato-eGFP allele (Figure XV in the Data Supplement). Using confocal microscopy, we observed tdTomato<sup>+</sup>eGFP<sup>+</sup> transitional SMC in the fibrous cap (Figure 4E) as well as the underlying media, often expressing VCAM1 (Figure 5E). Because these cells exist at the time of harvest after 18 weeks of Western diet, their double-positive status suggests that the transition of Myh11<sup>+</sup> cells to a Lgals3<sup>+</sup> state is an ongoing process throughout plaque pathogenesis.

Because this transition happened in the late-stage cap, we postulated that early-stage lesions would be dominated by tdTomato<sup>+</sup> SMC. We fed SMC<sup>Dual Lineage</sup> mice a Western diet for 10 weeks and examined lesions at the beginnings of SMC investment, defined as <50 lineage-traced SMC per plaque. Contrary to our expectations, we observed that almost all the SMC involved in early-stage lesions were eGFP<sup>+</sup> (Figure 4G through 4I). These eGFP<sup>+</sup> cells surrounded non-SMC-derived macrophage foam cells, consistent with a recent report suggesting that SMCs first populate lesions via the fibrous cap.<sup>29</sup> However, the mature fibrous cap is dominated by tdTomato<sup>+</sup> cells (Figure 4E and 4F) and features occasional transition cells directly overlying cells that were GFP<sup>+</sup> only. Of note, Sox9<sup>+</sup> GFP<sup>+</sup> SMCs were observed in advanced but not early-stage lesions. Even if eGFP<sup>+</sup> “pioneers” could revert to a Myh11<sup>+</sup>Acta2<sup>+</sup>Lgals3<sup>-</sup> state, they cannot return to being tdTomato<sup>+</sup>. Thus, our data suggest that eGFP<sup>+</sup>Lgals3<sup>+</sup> cells have privileged capacity to migrate into early-stage lesions but are later overtaken by tdTomato<sup>+</sup> cells that form the stable fibrous cap (Figure XVI in the Data Supplement). Although we originally selected Lgals3 purely as a marker of phenotypically modulated SMC, in vitro studies suggest Lgals3 may have a functional role, with small interfering RNA knockdown

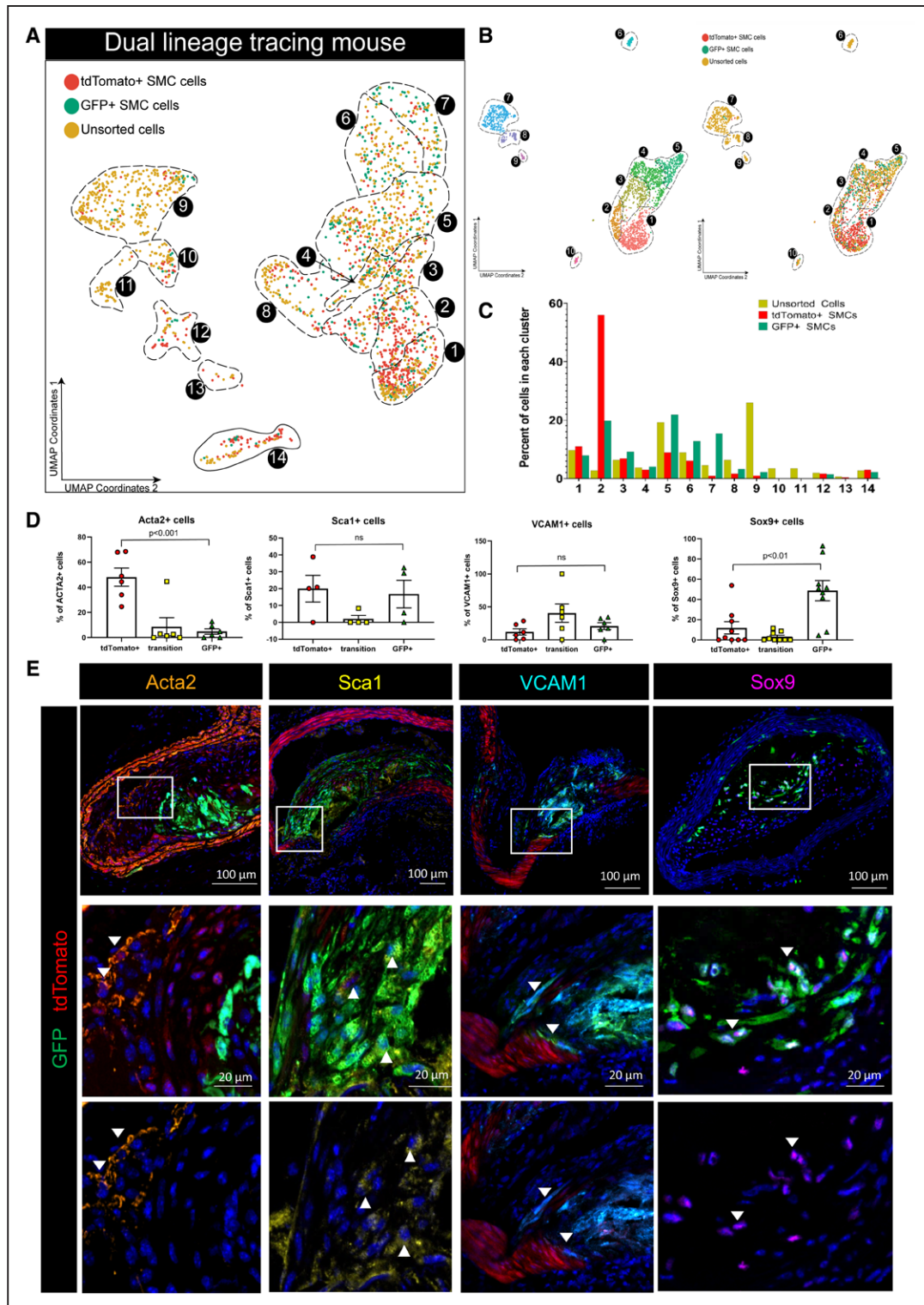
of Lgals3 reducing SMC migration and transcriptional activation of several collagens in response to PDGF-BB (Figure 4J, Figures X and XVII in the Data Supplement).

### scRNA-seq and Confocal Analysis of Human Atherosclerotic Lesions Show Similar Transcriptomic Cell Subsets Between Human and Mouse Lesions

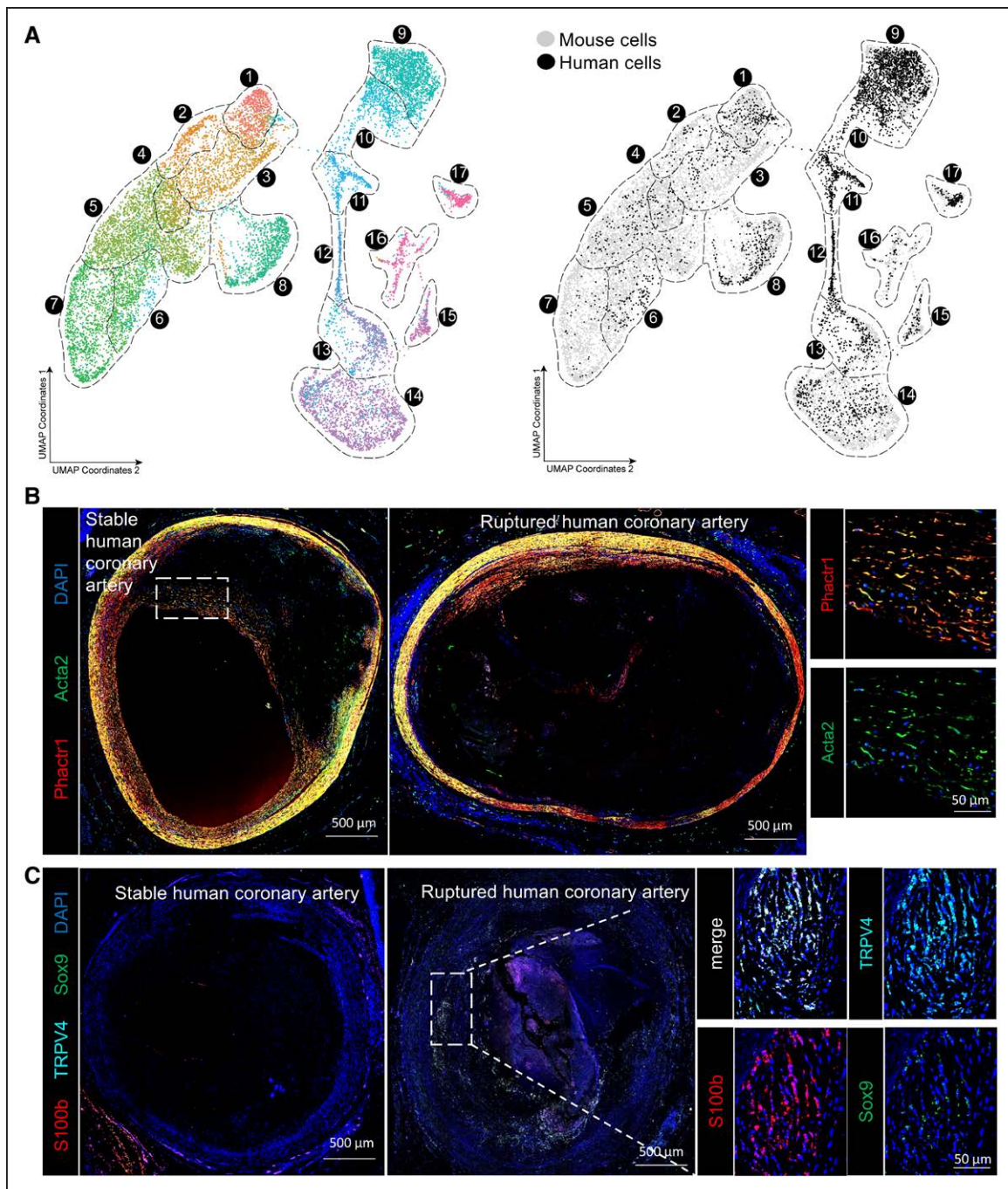
One of the strengths of scRNA-seq analysis is the ability to compare cell types resolved using different technologies and even species.<sup>17</sup> Therefore, we tested if the populations observed in our mouse models were present in human lesions via scRNA-seq. Of major significance, all clusters present in mouse advanced lesions were present in the human lesions, including cells in the osteogenic phenotype that was reduced with SMC<sup>Klf4</sup> knockout (Figure 6A, Figure XVIII in the Data Supplement). In addition, we validated these findings using confocal microscopy showing that specific markers of the osteogenic clusters found in our mouse analyses, including TRPV4, S100B, and Sox9, are present in ruptured human coronary lesions but not stable plaques (Figure 6C). Conversely, coronary lesions graded as stable were enriched for Phactr1 in the Acta2<sup>+</sup> cells in the fibrous cap (Figure 6B). Taken together, our results suggest that most SMC phenotypes in advanced atherosclerotic lesions are conserved between humans and mice, and, taken with previous knowledge that SMC<sup>Klf4</sup> is atheropromoting in mice and contains a human GWAS CAD risk locus, help us to infer that Klf4-dependent changes in SMC phenotype to an Lgals3<sup>+</sup> state are likely detrimental for atherosclerosis pathogenesis.

## DISCUSSION

Different SMC phenotypes have been hypothesized to have distinct roles in atherosclerotic lesion pathogenesis, some beneficial and some detrimental. In the present study, direct comparison of gene programs showed that SMC-specific knockout of Klf4 versus Oct4 resulted in nearly opposite genomic signatures consistent with the profoundly different plaque phenotypes exhibited by these mice (ie, smaller lesions with a thickened SMC-enriched fibrous cap for SMC<sup>Klf4</sup> mice and larger lesions nearly lacking a SMC-enriched



**Figure 5. Myh11-expressing differentiated SMC that transition through an *Lgals3* activation state take on multiple phenotypes.** **A**, scRNA-seq was performed on BCA plaque cells isolated from Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup> mice fed a Western diet for 18 weeks, as well as eGFP<sup>+</sup> and tdTomato<sup>+</sup> cells sorted from plaque or whole atherosclerotic aorta. The results of UMAP analyses are shown colored by the library of origin. Results are from 2143 cells from 3 separate Chromium 10x runs and 6 animals. **B**, UMAP of analysis based on only cells from Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup> experiments showing clear separation of tdTomato<sup>+</sup> and GFP<sup>+</sup> (right side). **C**, Quantification of unsorted, GFP<sup>+</sup>, and tdTomato<sup>+</sup> cells in each cluster in **A**. **D** and **E**, Atherosclerotic lesions from Myh11-Dre<sup>ERT2</sup> Lgals3-Cre Rosa-tdTomato-eGFP apoE<sup>-/-</sup> were immunostained for major markers of GFP<sup>+</sup> clusters from our scRNA-seq analyses and Acta2. Frozen sections (10  $\mu$ m) of renal arteries were used for VCAM1 and Sca1 antibody staining, along with tdTomato signal (endogenous) and a GFP antibody. Paraffin sections (10  $\mu$ m) of BCA were used for Acta2 and Sox9. Because of limited available antibodies, Sox9 required staining for tdTomato and GFP on sequential sections, which were used for quantification in **D**. Sox9 and GFP are shown as the representative image in **E**. Sox9 positivity was not found on SMCs outside the most advanced BCA lesions, including renal and aortic root sections. BCA indicates brachiocephalic artery; GFP, green fluorescent protein; scRNA-seq, single-cell RNA sequencing; SMC, smooth muscle cell; and UMAP, uniform manifold approximation and projection.



**Figure 6. Human scRNA-seq shows similar groups of fibrous cap and osteogenic SMC in advanced lesions.**

**A**, UMAP analysis of scRNA-seq results from human carotid endarterectomy samples, colored by cluster (left) or by origin (right) against mouse scRNA-seq data from Figure 2. **B**, Human coronary artery lesions graded as stable or ruptured were immunostained for Acta2 and Phactr1, a marker of cluster 1 to 3 SMCs identified in our scRNA-seq analyses of advanced BCA lesions. Images of whole immunostained human coronary lesions were obtained by taking a tile scan of  $\times 20$  zoom confocal microscopy z-stacks of full 10- $\mu\text{m}$  thickness. **C**, Human coronary artery lesions graded as stable or ruptured were immunostained for Sox9, TRPV4, and S100b, which were markers of osteogenic SMCs identified in our scRNA-seq analyses of mouse advanced BCA lesions, imaged as described in **B**. BCA indicates brachiocephalic artery; scRNA-seq, single-cell RNA sequencing; SMC, smooth muscle cell; and UMAP, uniform manifold approximation and projection.

fibrous cap in  $\text{SMC}^{\text{Oct4}}$  mice).<sup>3,4</sup> The results are of major interest, because they indicate that just 2 transcription factors that act cooperatively in regulating stem cell pluripotency<sup>30</sup> paradoxically result in virtually opposite lesion phenotypes by altering the transcriptome of SMCs.

Our scRNA-seq experiments, the first scRNA-seq focused exclusively on atherosclerotic lesions, identified 14 clusters, of which 7 are derived from SMCs based on rigorous lineage tracing. Only 3 of these 7 would have been identified as SMCs using traditional marker gene expression. These results are consistent with our previous

findings showing that >80% of SMC-derived cells in advanced lesions have lost expression of traditional markers.<sup>3</sup> However, our results greatly extend those studies by defining the complexity of SMCs in late-stage lesions by the genes these cells express, including clusters defined by inflammatory markers (*Vcam1*, *Ccl2*), stem-like markers (*Ly6a*), production of ECM (*Lum*, *Dcn*, *Col1a1*), and osteogenic markers (*Runx2*, *Sox9*, *Cyt11*). Our analysis showing how SMC knockout of the CAD GWAS gene *Klf4* impacts these clusters provides several key insights about mechanisms whereby *Klf4*-dependent processes might exacerbate lesion pathogenesis, including not only SMC-macrophage transitions but also SMC-osteogenic transitions and a pioneer state marked by *Lgals3*. Altogether, expression of *Klf4* is associated with negative effects on lesion pathogenesis that may be highly influential. Indeed, the *Klf4* network module in STARNET explains variance for CAD with a score >2, as large an effect size as low-density lipoprotein cholesterol.

Further, our studies allow us to reinterpret the meaning of *Lgals3* in SMC phenotypic transitions in the context of atherosclerosis. In our original studies of mice with SMC-specific *Klf4* knockout, we showed that approximately one third of *Lgals3*<sup>+</sup> cells within mouse advanced BCA lesions were of SMC, not myeloid, origin, and that SMC-derived cells also express many additional markers of macrophages including CD11b and F4/80.<sup>3</sup> Similar observations have been reported by many groups.<sup>5,29,31–33</sup> The general assumption has been that a subset of SMCs within lesions have phenotypically modulated into a macrophage-like state that is negatively affecting lesion pathogenesis. Similar to Dobnikar et al,<sup>33</sup> we did see some SMC-derived cells that expressed traditional macrophage markers CD11b and F4/80, but surprisingly, these were not the major population to change with *Klf4* knockout in SMCs, although these cells may be undersampled by scRNA-seq.<sup>20</sup> However, *Lgals3* was expressed in osteogenic clusters lost in SMC<sup>Klf4</sup> mice, consistent with our original finding that *Lgals3*<sup>+</sup> SMCs decrease in these mice. In addition to its role in macrophage differentiation and activation,<sup>34,35</sup> *Lgals3* has been shown to play a role in osteoblast differentiation<sup>36</sup> and osteoclast function/recruitment,<sup>37</sup> as well as SMC transition to an osteogenic phenotype in vitro.<sup>38</sup> Interestingly, *Lgals3* is necessary for transformation and maintaining stemness in some cancer cell lines<sup>39–41</sup> and the migration of cancer cells and epithelial cells.<sup>42,43</sup> Our scRNA-seq analysis also detected *Lgals3*<sup>+</sup> activation in clusters with rich expression of matrix components consistent with observations by Wirka et al<sup>7</sup> describing an *Lgals3*<sup>+</sup> fibrocyte that they postulate plays a role in plaque stability. However, the ECM-matrix population decreases with SMC<sup>Klf4</sup> knockout, and instead we see a 10-fold increase in SMC-derived Acta2<sup>+</sup> Phactr2<sup>+</sup> cells in the fibrous cap. In the natural course of bone formation, cartilaginous deposition of

ECM components is a necessary predecessor to calcification,<sup>44</sup> and it follows that matrix formation important for plaque stability might be continuous with the calcification process. Therefore, distinguishing SMC subsets that affect this process and their regulation is crucial for development of future therapies.

Our analysis also identified novel markers of SMC phenotypic clusters that can be identified at the transcript and protein level in both mouse models and human tissues. In particular, the congruence of our mouse and human scRNA-seq data offers hope that specific transcriptional signatures can be used to identify SMC-derived cells in humans without lineage tracing. We also found an intriguing overlap between our putative target genes in SMC<sup>Klf4</sup> or SMC<sup>Oct4</sup> murine atherosclerotic lesions and genes linked with nearly 40% of human CAD GWAS loci.<sup>18</sup> For the majority of CAD GWAS loci, the mechanisms of action influencing the disease phenotype are completely or mostly unknown<sup>18,19,45,46</sup> and reveal an auspicious depth of untapped potential for us to understand how CAD pathogenesis is influenced by SMC phenotype. In this work, we have shown that at least one third of the putative CAD target genes from murine lesions are altered by loss of *Klf4* or *Oct4* in cultured SMCs. However, we fully recognize that extensive further studies of each of these individual putative *Klf4* and *Oct4* target genes will be needed to show that they have a causal role in regulating SMC phenotype and CAD pathogenesis, including determining if CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9-induced human GWAS risk alleles are associated with functional changes in SMCs that exacerbate lesion pathogenesis.<sup>47</sup>

Taken together, our results show that *Klf4* regulates transition of SMCs to an *Lgals3*<sup>+</sup> osteogenic phenotype, likely detrimental given our previous atheroprotective phenotype in SMC<sup>Klf4</sup> mice. In addition, we provide novel evidence that diverse SMC lesion phenotypes are derived from a subset of *Lgals3*<sup>+</sup> transitional state SMCs and show marked similarity to human plaque cells, suggesting this may be a rich area for development of novel therapeutics that promote beneficial versus detrimental changes in SMC phenotype.

## ARTICLE INFORMATION

Received March 5, 2020; accepted June 16, 2020.

The Data Supplement is available with this article at <https://www.ahajournals.org/doi/suppl/10.1161/CIRCULATIONAHA.120.046672>.

## Correspondence

Gary K. Owens, PhD, University of Virginia School of Medicine, Robert M. Berne Cardiovascular Research Center, PO Box 801394, MR5 Building, Charlottesville, VA 22908-1394. Email [gko@virginia.edu](mailto:gko@virginia.edu)

## Affiliations

Robert M. Berne Cardiovascular Research Center (G.F.A., K.M.O., S.K., A.N., C.M.W., S.S., C.A.H., R.H., R.A.B., C.A.M., E.R.Z., G.K.O.), Department of

Biochemistry and Molecular Genetics (G.F.A., K.M.O., C.A.H., R.A.B., S.B.), Department of Biomedical Engineering (C.M.W., E.R.Z.), School of Medicine, Division of Cardiovascular Medicine, Department of Medicine (C.A.M.), University of Virginia, Charlottesville. Department of Cardiac Surgery, Tartu University Hospital, Estonia (K.S.). Integrated Cardio Metabolic Centre, Department of Medicine, Karolinska Institutet, Karolinska Universitetssjukhuset, Huddinge, Sweden (G.M., V.V., J.B.). Laboratory of Clinical Chemistry and Hematology, Division Laboratories and Pharmacy (M.M., G.P.), Department of Cardiology (M.M.), University Medical Center Utrecht, University Utrecht, The Netherlands. CVPath Institute, Inc, Gaithersburg, MD (A.V.F.). Department of Genetics and Genomic Sciences (J.B.), Icahn Institute of Genomics and Multiscale Biology (J.B.), Icahn School of Medicine at Mount Sinai, New York.

## Acknowledgments

The authors thank the University of Virginia genomics core facility, including Dr Katia Sol-Church, Alyson Prorock, and Yongde Bao for assistance with single cell sequencing; the University of Virginia flow cytometry core facility, especially Michel Solga and Claude Chew for cell sorting assistance; and the University of Virginia advanced microscopy core, especially Stacey Guillot for imaging and Karen Soohoo for assistance with immunofluorescent imaging quantification.

## Sources of Funding

This work was supported by National Institutes of Health grants R01 HL132904, HL136314, and HL141425 to Dr Owens.

## Disclosures

None.

## Supplemental Materials

Data Supplement Methods  
Data Supplement Tables I and II  
Data Supplement Figures I–XVIII  
References 48–52  
Data Supplement Excel Files I and II  
Data Supplement Movies I and II

## REFERENCES

- Bentzon JF, Otsuka F, Virmani R, Falk E. Mechanisms of plaque formation and rupture. *Circ Res*. 2014;114:1852–1866. doi: 10.1161/CIRCRESAHA.114.302721
- Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. *Nature*. 2011;473:317–325. doi: 10.1038/nature10146
- 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. doi: 10.1038/nm.3866
- Cherepanova OA, Gomez D, Shankman LS, Swiatlowska P, Williams J, Sarmiento OF, Alencar GF, Hess DL, Bevard MH, Greene ES, et al. Activation of the pluripotency factor OCT4 in smooth muscle cells is atheroprotective. *Nat Med*. 2016;22:657–665. doi: 10.1038/nm.4109
- Feil S, Fehrenbacher B, Lukowski R, Essmann F, Schulze-Osthoff K, Schaller M, Feil R. Transdifferentiation of vascular smooth muscle cells to macrophage-like cells during atherogenesis. *Circ Res*. 2014;115:662–667. doi: 10.1161/CIRCRESAHA.115.304634
- Gomez D, Shankman LS, Nguyen AT, Owens GK. Detection of histone modifications at specific gene loci in single cells in histological sections. *Nat Methods*. 2013;10:171–177. doi: 10.1038/nmeth.2332
- Wirka RC, Wagh D, Paik DT, Pjanic M, Nguyen T, Miller CL, Kundu R, Nagao M, Collier J, Koyano TK, et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. *Nat Med*. 2019;25:1280–1289. doi: 10.1038/s41591-019-0512-5
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21. doi: 10.1093/bioinformatics/bts635
- Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923–930. doi: 10.1093/bioinformatics/btt656
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. doi: 10.1186/s13059-014-0550-8
- Krämer A, Green J, Pollard J Jr, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics*. 2014;30:523–530. doi: 10.1093/bioinformatics/btt703
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–359. doi: 10.1038/nmeth.1923
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25:2078–2079. doi: 10.1093/bioinformatics/btp352
- Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9:R137. doi: 10.1186/gb-2008-9-9-r137
- Muraro MJ, Dharmadhikari G, Grün D, Gronen N, Dielen T, Jansen E, van Gurp L, Engelse MA, Carlotti F, de Koning EJ, et al. A single-cell transcriptome atlas of the human pancreas. *Cell Syst*. 2016;3:385–394.e3. doi: 10.1016/j.cels.2016.09.002
- Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L, Gennert D, Li S, Livak KJ, Rozenblatt-Rosen O, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol*. 2016;17:77. doi: 10.1186/s13059-016-0938-8
- Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36:411–420. doi: 10.1038/nbt.4096
- Erdmann J, Kessler T, Munoz Venegas L, Schunkert H. A decade of genome-wide association studies for coronary artery disease: the challenges ahead. *Cardiovasc Res*. 2018;114:1241–1257. doi: 10.1093/cvr/cvy084
- van der Harst P, 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. doi: 10.1161/CIRCRESAHA.117.312086
- Winkels H, Ehinger E, Vassallo M, Buscher K, Dinh HQ, Kobiyama K, Hamers AAJ, Cochain C, Vafadarnejad E, Saliba AE, et al. Atlas of the immune cell repertoire in mouse atherosclerosis defined by single-cell RNA-sequencing and mass cytometry. *Circ Res*. 2018;122:1675–1688. doi: 10.1161/CIRCRESAHA.117.312513
- Cochain C, Vafadarnejad E, Arampatzis P, Pelisek J, Winkels H, Ley K, Wolf D, Saliba AE, Zernecke A. Single-cell RNA-seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. *Circ Res*. 2018;122:1661–1674. doi: 10.1161/CIRCRESAHA.117.312509
- Kim K, Shim D, Lee JS, Zaitsev K, Williams JW, Kim KW, Jang MY, Seok Jang H, Yun TJ, Lee SH, et al. Transcriptome analysis reveals non-foamy rather than foamy plaque macrophages are proinflammatory in atherosclerotic murine models. *Circ Res*. 2018;123:1127–1142. doi: 10.1161/CIRCRESAHA.118.312804
- Hansson GK, Jonasson L, Holm J, Claesson-Welsh L. Class II MHC antigen expression in the atherosclerotic plaque: smooth muscle cells express HLA-DR, HLA-DQ and the invariant gamma chain. *Clin Exp Immunol*. 1986;64:261–268.
- Komori T. Runx2, an inducer of osteoblast and chondrocyte differentiation. *Histochem Cell Biol*. 2018;149:313–323. doi: 10.1007/s00418-018-1640-6
- Shanahan CM, Crouthamel MH, Kapustin A, Giachelli CM. Arterial calcification in chronic kidney disease: key roles for calcium and phosphate. *Circ Res*. 2011;109:697–711. doi: 10.1161/CIRCRESAHA.110.234914
- St Hilaire C, Ziegler SG, Markello TC, Brusco A, Groden C, Gill F, Carlson-Donohoe H, Lederman RJ, Chen MY, Yang D, et al. NT5E mutations and arterial calcifications. *N Engl J Med*. 2011;364:432–442. doi: 10.1056/NEJMoa0912923
- Durham AL, Speer MY, Scatena M, Giachelli CM, Shanahan CM. Role of smooth muscle cells in vascular calcification: implications in atherosclerosis and arterial stiffness. *Cardiovasc Res*. 2018;114:590–600. doi: 10.1093/cvr/cvy010
- Muzumdar, Tasic, Miyamichi, Li L. A global double-fluorescence Cre reporter mouse. *Genesis*. 2007;45:593–605. doi: 10.1002/dvg.20335
- Misra A, Feng Z, Chandran RR, Kabir I, Rotllan N, Aryal B, Sheikh AQ, Ding L, Qin L, Fernández-Hernando C, et al. Integrin beta3 regulates clonality and fate of smooth muscle-derived atherosclerotic plaque cells. *Nat Commun*. 2018;9:1905–1918. doi: 10.1038/s41467-018-04447-7

30. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006;126:663–676. doi: 10.1016/j.cell.2006.07.024
31. Allahverdian S, Chehroudi AC, McManus BM, Abraham T, Francis GA. Contribution of intimal smooth muscle cells to cholesterol accumulation and macrophage-like cells in human atherosclerosis. *Circulation*. 2014;129:1551–1559. doi: 10.1161/CIRCULATIONAHA.113.005015
32. Wang Y, Dubland JA, Allahverdian S, Asonye E, Sahin B, Jaw JE, Sin DD, Seidman MA, Leeper NJ, Francis GA. Smooth muscle cells contribute the majority of foam cells in apoE (apolipoprotein E)-deficient mouse atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2019;39:876–887. doi: 10.1161/ATVBAHA.119.312434
33. Dobnikar L, Taylor AL, Chappell J, Oldach P, Harman JL, Oerton E, Dzierzak E, Bennett MR, Spivakov M, Jørgensen HF. Disease-relevant transcriptional signatures identified in individual smooth muscle cells from healthy mouse vessels. *Nat Commun*. 2018;9:4567. doi: 10.1038/s41467-018-06891-x
34. MacKinnon AC, Farnworth SL, Hodkinson PS, Henderson NC, Atkinson KM, Leffler H, Nilsson UJ, Haslett C, Forbes SJ, Sethi T. Regulation of alternative macrophage activation by galectin-3. *J Immunol*. 2008;180:2650–2658. doi: 10.4049/jimmunol.180.4.2650
35. Dumic J, Dabelic S, Flögel M. Galectin-3: an open-ended story. *Biochim Biophys Acta - Gen Subj*. 2006;1760:616–635. doi: 10.1016/j.bbagen.2005.12.020
36. Nakajima K, Kho DH, Yanagawa T, Harazono Y, Gao X, Hogan V, Raz A. Galectin-3 inhibits osteoblast differentiation through notch signaling. *Neoplasia*. 2014;16:939–949. doi: 10.1016/j.neo.2014.09.005
37. Li YJ, Kukita A, Teramachi J, Nagata K, Wu Z, Akamine A, Kukita T. A possible suppressive role of galectin-3 in upregulated osteoclastogenesis accompanying adjuvant-induced arthritis in rats. *Lab Invest*. 2009;89:26–37. doi: 10.1038/labinvest.2008.111
38. Menini S, Iacobini C, Ricci C, Blasetti Fantauzzi C, Salvi L, Pesce CM, Relucenti M, Familiari G, Taurino M, Pugliese G. The galectin-3/RAGE dyad modulates vascular osteogenesis in atherosclerosis. *Cardiovasc Res*. 2013;100:472–480. doi: 10.1093/cvr/cvt206
39. Ruvolo PP. Galectin 3 as a guardian of the tumor microenvironment. *Biochim Biophys Acta*. 2016;1863:427–437. doi: 10.1016/j.bbamcr.2015.08.008
40. Kang HG, Kim DH, Kim SJ, Cho Y, Jung J, Jang W, Chun KH. Galectin-3 supports stemness in ovarian cancer stem cells by activation of the Notch1 intracellular domain. *Oncotarget*. 2016;7:68229–68241. doi: 10.18632/oncotarget.11920
41. Nangia-Makker P, Hogan V, Raz A. Galectin-3 and cancer stemness. *Glycobiology*. 2018;28:172–181. doi: 10.1093/glycob/cwy001
42. Zhang D, Chen ZG, Liu SH, Dong ZQ, Dalin M, Bao SS, Hu YW, Wei FC. Galectin-3 gene silencing inhibits migration and invasion of human tongue cancer cells *in vitro* via downregulating  $\beta$ -catenin. *Acta Pharmacol Sin*. 2013;34:176–184. doi: 10.1038/aps.2012.150
43. Margadant C, van den Bout I, van Boxtel AL, Thijssen VL, Sonnenberg A. Epigenetic regulation of galectin-3 expression by  $\beta$ 1 integrins promotes cell adhesion and migration. *J Biol Chem*. 2012;287:44684–44693. doi: 10.1074/jbc.M112.426445
44. Hinton RJ, Jing Y, Jing J, Feng JQ. Roles of chondrocytes in endochondral bone formation and fracture repair. *J Dent Res*. 2017;96:23–30. doi: 10.1177/0022034516668321
45. Schunkert H, von Scheidt M, Kessler T, Stiller B, Zeng L, Vilne B. Genetics of coronary artery disease in the light of genome-wide association studies. *Clin Res Cardiol*. 2018;107(suppl 2):2–9. doi: 10.1007/s00392-018-1324-1
46. Clarke SL, Assimes TL. Genome-wide association studies of coronary artery disease: recent progress and challenges ahead. *Curr Atheroscler Rep*. 2018;20:47. doi: 10.1007/s11883-018-0748-4
47. Lo Sardo V, Chubukov P, Ferguson W, Kumar A, Teng EL, Duran M, Zhang L, Cost G, Engler AJ, Urnov F, et al. Unveiling the role of the most impactful cardiovascular risk locus through haplotype editing. *Cell*. 2018;175:1796–1810.e20. doi: 10.1016/j.cell.2018.11.014
48. Zunder ER, Finck R, Behbehani GK, Amir DE, Krishnaswamy S, Gonzalez VD, Lorang CG, Bjornson Z, Spitzer MH, Bodenmiller B, et al. Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. *Nat Protoc*. 2015;10:316–333. doi: 10.1038/nprot.2015.020
49. Finck R, Simonds EF, Jager A, Krishnaswamy S, Sachs K, Fantl W, Pe'er D, Nolan GP, Bendall SC. Normalization of mass cytometry data with bead standards. *Cytometry A*. 2013;83:483–494. doi: 10.1002/cyto.a.22271
50. Fread KI, Strickland WD, Nolan GP, Zunder ER. An updated debarcoding tool for mass cytometry with cell type-specific and cell sample-specific stringency adjustment. *Pac Symp Biocomput*. 2017;22:588–598. doi: 10.1142/9789813207813\_0054
51. Bendall SC, Simonds EF, Qiu P, Amir DE, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science*. 2011;332:687–696. doi: 10.1126/science.1198704
52. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv*. 2018;1802.03426.