Disease-relevant transcriptional signatures identified in individual smooth muscle cells from healthy mouse vessels

Lina Dobnikar1,2, Annabel L. Taylor2, Joel Chappell2, Phoebe Oldach1,6, Jennifer L. Harman2, Erin Oerton1,7, Elaine Dzierzak3, Martin R. Bennett2, Mikhail Spivakov1,4,5 & Helle F. Jørgensen2

Vascular smooth muscle cells (VSMCs) show pronounced heterogeneity across and within vascular beds, with direct implications for their function in injury response and atherosclerosis. Here we combine single-cell transcriptomics with lineage tracing to examine VSMC heterogeneity in healthy mouse vessels. The transcriptional profiles of single VSMCs consistently reflect their region-specific developmental history and show heterogeneous expression of vascular disease-associated genes involved in inflammation, adhesion and migration. We detect a rare population of VSMC-lineage cells that express the multipotent progenitor marker Sca1, progressively downregulate contractile VSMC genes and upregulate genes associated with VSMC response to inflammation and growth factors. We find that Sca1 upregulation is a hallmark of VSMCs undergoing phenotypic switching in vitro and in vivo, and reveal an equivalent population of Sca1-positive VSMC-lineage cells in atherosclerotic plaques. Together, our analyses identify disease-relevant transcriptional signatures in VSMC-lineage cells in healthy blood vessels, with implications for disease susceptibility, diagnosis and prevention.
vascular smooth muscle cell (VSMC) accumulation is a hallmark of cardiovascular diseases such as atherosclerosis that causes heart attack and stroke. VSMCs are found within the medial layer of large blood vessels, provide mechanical strength to the vessel and regulate vascular tone to control blood flow and blood pressure. VSMCs within healthy vessels are quiescent and characterised by the expression of contractile proteins such as αSMA (also known as ACTA2), Myocardin (MYOCD) and SM-MHC (also known as MYH11). However, VSMCs display remarkable phenotypic plasticity. When stimulated by injury or inflammation, VSMCs downregulate expression of the genes responsible for contractility and acquire a phenotype characterised by increased extracellular matrix production, migration and proliferation.

VSMC heterogeneity within and between different vascular regions with regard to morphology, growth characteristics and expression of specific candidate genes has been identified previously. The observed cell-to-cell variation might result from different vascular structure and blood flow, as well as from the distinct developmental origin of VSMCs in different vascular beds. It has therefore been hypothesised that VSMCs displaying different levels of plasticity co-exist within the healthy vessel wall and might contribute to the non-random disease susceptibility of individual parts of the vasculature. We and others recently demonstrated that VSMC accumulation in atherosclerosis and after injury results from extensive clonal expansion of a small number of VSMCs. This suggests that cells undergoing expansion were originally different from the general VSMC population in the healthy vessel wall, highlighting a possible functional significance of VSMC heterogeneity.

Single-cell RNA-sequencing (scRNA-seq) enables genome-wide profiling of individual cells and is therefore an ideal methodology to detect cellular heterogeneity in an unbiased manner. Here we combine different scRNA-seq methodologies to delineate VSMC heterogeneity in healthy arteries and provide global insight into the nature of distinct cell subsets. We show that while the contractile VSMC signature is expressed relatively uniformly across most cells, there are pronounced differences in single-VSMC expression profiles between and within vascular beds for genes involved in cell adhesion, migration and inflammation. Combining scRNA-seq with VSMC lineage tracing, we reveal a rare subset of VSMC-lineage cells expressing Stem Cell Antigen 1 (Sca1, encoded by Ly6a, referred to below as Ly6a/Sca1). Sca1-positive cells show progressive downregulation of contractile VSMC genes and increased expression of genes associated with wound healing, migration and activation of growth factor signalling. We provide experimental evidence indicating that upregulation of Sca1 is a hallmark of VSMCs undergoing phenotypic switching in vivo and detect an equivalent Sca1-positive cell population within atherosclerotic plaques. These results suggest that Sca1 marks a plastic VSMC subpopulation that gives rise to functionally distinct cells in disease.

**Results**

Single-cell transcriptomes of vascular smooth muscle cells. To delineate VSMC heterogeneity in healthy tissue, we initially generated transcriptional profiles of individual cells using Fluidigm C1 technology. Single cells were isolated from the medial layer of mouse aortas by enzymatic digestion (Fig. 1a) and processed for sequencing. The scRNA-seq profiles of 143 cells from four experiments passed stringent quality control (83% of the analysed cells; Supplementary Fig. 1; see Methods for details). On average, 6704 genes were detected per cell.

All analysed cells showed consistently high expression of the contractile VSMC markers Myh11, Acta2 and Tagln. Lower-expressed marker genes Myocd, Smtn, Vcl and Cnn1 were detected in most cells, similar to what was observed for housekeeping genes (Fig. 1b). This indicates that all medial cells analysed express a contractile VSMC signature. Consistent with this conclusion, principal component analysis (PCA) demonstrated that all analysed single cells clustered tightly with VSMC control samples and away from adventitial control samples, both generated using the tube control protocol (Fig. 1c). Furthermore, the pooled single-cell VSMC expression profiles correlated with bulk RNA-seq data ($R^2=0.51$; Fig. 1d). As expected, the scRNA-seq profiles of ex vivo cells were clearly distinct from published profiles of cultured VSMCs (Fig. 1e), which are known to recapitulate some aspects of VSMC phenotypic switching. In particular, ex vivo cells expressed higher levels of contractile genes compared with the cultured cells (Supplementary Fig. 2). Taken together, these results confirm that the single-cell datasets we have generated are of sufficient coverage, specificity and overall quality to unambiguously identify the profiled cells as contractile VSMCs.

**Regional differences in VSMC gene expression.** We applied single-cell transcriptomics to gain insight into the previously observed VSMC heterogeneity between the athero-prone aortic arch (AA) and the more disease-resistant descending thoracic aorta (DT) (Fig. 1f). AA and DT VSMCs differ in expression of a number of genes, including posterior Hox genes and other transcription factors. These differences may arise from the distinct embryonic origins of VSMCs in the AA (neural crest) and DT (mesoderm) regions. Alternatively, VSMCs in both regions could be heterogeneous with respect to these genes, with specific cell subsets represented in different proportions in the AA compared with the DT. These scenarios can be addressed directly with scRNA-seq. Prior to analysing regional expression differences at the single-cell level, we generated bulk RNA-seq profiles of VSMCs isolated from the medial layer of AA and DT to define robust gene expression signatures associated with VSMCs from these regions at the population level, which informed single-cell analysis (Fig. 2a). In total, 442 genes showed significant differential expression ($p<0.01$, log fold change $>1$), of which 386 genes were upregulated in the AA and 56 genes in the DT (Fig. 2b and Supplementary Data 1). RT-qPCR in independent, paired AA and DT samples verified differential expression of 6 out of 6 genes tested (Den, Lum, Pde1c, Gpc3, 3632451006Rik and Hoxa7; Fig. 2c). Overall, 88/442 genes identified as region-specific by RNA-seq were previously found to be differentially expressed in AA and DT VSMCs by microarray analysis. Notably, 44% of the microarray-identified genes were lowly expressed (normalised log$_2$-transformed read count $<4$) and their regional specificity was less reproducible by RNA-seq compared with genes expressed at higher levels (Supplementary Fig. 3). Similar to previous reports, the genes upregulated in the DT population compared with the AA included many developmental regulators, including 15 Hox genes corresponding to a posterior identity, and were significantly enriched for the gene ontology (GO) terms “anterior/posterior pattern specification” and “regionalisation and developmental process” (Supplementary Data 2a). In contrast, genes upregulated in AA VSMCs were enriched for terms associated with immune response, cell proliferation and migration (for example Lbp, Tgfb1 and Mmp12; Fig. 2b and Supplementary Data 2b), consistent with the increased response of AA cells to mitogens and inflammatory cytokines. Additionally, genes typically associated with the synthetic VSMC phenotype (Spp1, Pde1c) were expressed at higher levels in the AA compared with the DT VSMCs (Fig. 2b). In summary, population-level analysis of VSMC region-specific transcription identified
Fig. 1 Single-cell RNA-seq analysis of vascular smooth muscle cells. a Schematic of the approach. Cells from the medial layer are enzymatically digested to obtain a single-cell suspension. Single-cell cDNA libraries are then generated, followed by sequencing and data analysis. b Violin plots showing the log2-transformed normalised expression of VSMC marker genes across the profiled 143 cells (top), as well as of housekeeping genes with similar mean expression levels (lower panel). c Mapping of single-cell VSMC transcriptomes (light blue), as well as transcriptomes from control VSMC (Sca1−, dark blue) and adventitial (Adv) cell (Sca1−, orange; Sca1+, red) samples (tube controls) on a two-dimensional PCA space. d Dot plot showing the log2-transformed read counts detected for each gene (black dots) when pooling across all single-cell samples versus the read counts detected with bulk RNA-seq. The dashed line shows a linear regression fit. e PCA plot summarising the single-cell expression profiles for ex vivo VSMCs (blue) and in vitro cultured VSMCs (green, data from Gene Expression Omnibus accession GSE79436, Adhikari et al.13)

Fig. 2 Bulk RNA-seq of VSMCs from the aortic arch and descending thoracic aorta. a Schematic representation of the aorta, indicating the aortic arch (AA) and descending thoracic aorta (DT). b Volcano plot showing significance (−log_{10} p-value) versus relative gene expression in VSMCs from the DT versus the AA. Genes showing significant differences in expression (adjusted p-value < 0.01, log2 fold change > 1) are labelled in yellow (upregulated in DT) and red (upregulated in AA). Gene names for selected genes are indicated. Three independent samples were analysed for each region (AA and DT samples paired). c Relative expression (log2(DT/AA)) of selected genes determined by RT-qPCR (black dots/bars, n = 4) and bulk RNA-seq (grey dots/bars). Error bars indicate s.e.m. from four independent RT-qPCR experiments. Rik*: 3632451O06Rik
molecular determinants that may underlie the reported regional differences in VSMC growth factor response and disease susceptibility.

Vascular region identity manifests in individual VSMCs. The scRNA-seq dataset described above (Fig. 1) comprised 79 AA and 64 DT cells, allowing comparison of the transcriptional signatures associated with VSMCs from these regions at a single-cell level. Consistent with the region-specific gene signatures identified by bulk expression analysis, genes such as *Pde1c* and *Hand2* were almost exclusively detected in cells from the AA region and *Hoxa7* in cells from the DT region (Fig. 3a, top panel). However, other genes showing region-dependent expression levels in the bulk analysis were expressed at similar levels in a subset of cells from both regions, while the number of expressing cells was greater in either the DT (*Tnc, Calcrl*) or AA region (*Aspn, Gpc3*) (Fig. 3a, lower panel). Therefore, the observed region-specific gene signatures could reflect the presence of different proportions of cells expressing these genes in each region. Alternatively, and not mutually exclusively with the above, each individual VSMC may bear an inherent regional signature.

To distinguish between these possibilities, we investigated whether VSMC regional identity can be reliably predicted from single-cell profiles. To this end, we employed random forest (RF) analysis (Fig. 3b), a nonparametric tree-based machine learning method that is widely used for classification based on expression data, including scRNA-seq. First, 25% of cells were randomly selected and set aside for the final testing of the prediction algorithm. Data for the remaining 108 cells was used to rank differentially expressed genes detected in the bulk RNA-seq analysis, based on their power to predict regional identity in individual cells. The influence of each of the top 30 predictive genes on RF classification accuracy is shown in Fig. 3c and 3d. The area under the ROC curve (AUC) was 0.9967, indicating excellent performance (Fig. 3d). A PCA plot based on the 30 genes used in the final classifier, showing AA (red) and DT (yellow) cells. Cells in the 25% test subset are circled in black.
Supplementary Data 3a. These 30 genes were used in the final RF classifier; using additional genes for classification did not significantly increase classifier performance. The resulting RF model showed very high predictive power on the unseen 25% of the data, as demonstrated by the ROC curve (Fig. 3d), correctly classifying 17/18 AA cells and 14/15 DT cells set aside for model testing (Supplementary Data 3b). Consistent with this, a PCA plot based on the 30 RF classifier genes showed a clear segregation of AA from DT cells (Fig. 3e). The highly robust performance of the scRNA-seq-based classifier strongly indicates that features of regional identity are borne by individual cells, rather than reflect the differential composition of cell populations between the regions.

VSMC heterogeneity within vascular regions. We next examined whether cells within a specific vascular region showed heterogeneity with respect to expression of other genes. To identify variably expressed genes robustly, we assessed the variance of log-transformed scRNA-seq counts for each gene across cells, relative to that expected from the estimated technical noise (modified from ref. 22; Fig. 4a; see Methods for details). We identified 113 highly variably expressed genes (HVGs) in AA and 79 HVGs in DT VSMCs (adjusted p-value < 0.05; Supplementary Data 4ab). Notably, the HVG scores between the two regions were correlated (r = 0.67; Supplementary Fig. 4a) and only one of the identified HVGs (Wif1) was detected as differentially expressed between the two regions.

Many of the identified HVGs encode factors that have been directly linked to VSMC phenotype or have documented roles in cardiovascular disease, inflammation and VSMC biology (Fig. 4b and Supplementary Data 4c). For example, one of the top HVGs, Rgs5, was detected as differentially expressed between the two regions.

**Fig. 4** VSMCs show heterogeneous expression of genes implicated in cardiovascular disease. a Scatter plots showing the mean-variance relationship of log2-transformed normalised expression levels for each gene, with colour-highlighting of genes showing highly variable expression in the aortic arch (AA, red, top panel) and descending thoracic aorta (DT, yellow, lower panel). b Bar graph showing the implication of the identified highly variable genes (HVGs) in AA (red), DT (yellow) or both regions (green) in functions related to VSMC biology based on published literature (see Methods for details). c Dot plot showing log2-transformed normalised counts detected in individual VSMCs from AA (red) and DT (yellow) for selected genes that show variable expression across single cells. d, e t-SNE plot visualising a 10X Chromium dataset generated from 2846 unselected cells (gated as live using Zombie NIR staining and singlets using doublet discrimination) from the whole aortas of three tamoxifen-labelled Myh11-CreERT2/Confetti animals (pooled). d Clusters generated using graph-based clustering are colour-coded as indicated and adventitial (Adv), endothelial (EC) and VSMC (VSMC) populations are labelled. e Log-transformed expression levels of selected HVGs identified in AA and DT populations based on Fluidigm C1 data (Rgs5, Irf1, Atf3, Nfkbia), shown using a scale from light to dark grey.
variable genes could underlie the functional heterogeneity of variation in expression levels in the vessel media. These highly regulation and vascular disease development show cell-to-cell directly contribute to disease predisposition of the aortic arch. expression manifests at the level of individual VSMCs and may cytometry analysis (Supplementary Fig. 5b). To eliminate possible cells (Fig. 4c). Other genes showing atherosclerosis8 to VSMC biology, given, for example, that only a small number of (also identi top HVGs in both AA and DT regions. As shown in Fig. 4d, t-SNE analysis of this expression (including Myh11/Cnn1, Cdhr5 and Pdgfra, respectively, Supplementary Fig. 4b). Myh11-expressing VSMCs further segregated into seven clusters, highlighting the heterogeneity of this population. Remarkably, these clusters showed differential expression of 13% of the HVGs identified above, including Rgs5, Irf1, Nkx2.5 and Adamts1 (22/147, Fisher’s exact test \( p = 2 \times 10^{-15} \)), and two genes prioritised using the DM method (Rgs5 and Plaur) based on Fluidigm C1 data (Fig. 4e, see Methods for details).

Jointly, these analyses demonstrate that region-specific gene expression manifests at the level of individual VSMCs and may directly contribute to disease predisposition of the aortic arch. Additionally, many genes previously implicated in VSMC regulation and vascular disease development show cell-to-cell variation in expression levels in the vessel media. These highly variable genes could underlie the functional heterogeneity of VSMCs, such as that observed in atherosclerosis8–10.

A subset of VSMC-lineage cells express Stem Cell Antigen 1. The Ly6a/Sca1 gene, which marks stem and progenitor cells in the vasculature29–34 and other tissues35 was identified as heterogeneous expressed in VSMCs with the DM method above (Supplementary Fig. 5a). Consistent with this, Ly6a/Sca1 expression was detected in individual cells in the VSMC cluster in the 10X Chromium whole aorta analysis, in addition to the expected abundant Ly6a/Sca1 expression in the adventitial and endothelial clusters (Fig. 5a). Furthermore, 0.5–1% of medial cells (\( n > 10 \)) stained positively with anti-Sca1 antibody in flow cytometry analysis (Supplementary Fig. 5b). To eliminate possible artefacts from antibody cross-reactivity, we also analysed transgenic Sca1-GFP animals, which express GFP from the Ly6a/Sca1 promoter37. Similar to the observations from antibody staining, 0.2–1.6% of medial cells from transgenic animals expressed GFP (Supplementary Fig. 5c, \( n = 4 \)). To test whether GFP-expressing cells from the medial layer of Sca1-GFP animals have a VSMC identity, we sorted GFP-positive cells from either medial or adventitial samples and immunostained them for aSMA (a contractile VSMC marker). As expected, none of the 60 analysed GFP+ cells from the adventitial sample expressed aSMA. In contrast, 25–86% of medial GFP+ cells stained positive for aSMA (Fig. 5b, c). Taken together, these results suggest the existence of a rare subpopulation of Sca1-expressing VSMCs within the medial layer of healthy animals.

To further confirm the lineage affiliation of Sca1+ medial cells, we analysed Sca1 expression in the aortic media of mouse models that allow for genetic lineage labelling of VSMCs (Myh11-CreER2/EYFP38 and Myh11-CreER2/Confetti38). These animals express the tamoxifen-inducible CreER2 recombine selectively in smooth muscle cells under control of the Myh11 promoter. Activation of the Cre recombinase by tamoxifen administration therefore generates in VSMC-specific deletion of a fluorescent reporter at either single-colour (EYFP) or multicolour reporters (Confetti), resulting in stable fluorescent protein expression (Fig. 6a). The specificity of the lineage label to the media was confirmed by analysis of 72 arterial sections from six Confetti-labelled animals (Fig. 6b, c), which detected only seven lineage-positive cells in the adventitial layer and one cell in the vascular endothelium. In contrast, recombination was efficient in medial cells (70–95% for the Confetti reporter and 40–90% for the EYFP reporter, Fig. 6d), consistent with previous data8. We confirmed by flow cytometry that a significant proportion of Sca1-expressing cells (S+) isolated from the media of Myh11-CreER2/EYFP mice expressed the EYFP lineage label (Fig. 6e). We also detected a small number of Ly6a/Sca1+ cells in a 10X Chromium analysis of sorted Myh11-lineage labelled cells (Supplementary Fig. 6a), confirming the existence of a rare population of Sca1-positive VSMCs in the media.

Single-cell expression profiles of Sca1+ VSMC-lineage cells. Due to the low frequency of Ly6a/Sca1+ VSMCs, the number of cells profiled using 10X Chromium technology was insufficient for robust analysis. We therefore index-sorted cells from tamoxifen-treated Myh11-CreER2/Confetti animals based on Sca1 expression and profiled their transcriptomes using the Smart-seq2 protocol (Fig. 7a). To ensure that only single cells were analysed, we gated for cells expressing only a single colour of the lineage label (L) and confirmed singlet isolation by confocal microscopy (Supplementary Fig. 6b). In total, data for 92 Sca1-positive lineage-positive cells (S+L+), as well as 27 Sca1-positive/lineage-negative (S+L–) and 36 “conventional” Sca1-negative/lineage-positive (S–L+) control cells, from six independent cell pools profiled in three experiments passed stringent quality control filters (Supplementary Fig. 6c–e, see Methods for details) and were used in further analysis.

Ly6a/Sca1 transcripts were only detected in a subset of S+ cells (41/119, 34%, Fig. 7b), which could be due to either a lower sensitivity of scRNA-seq compared with FACS analysis or biological factors such as differences between transcript and protein stability. While all S–L+ cells expressed Myh11 mRNA, as expected, S+L+ cells showed a broad range of Myh11 transcript levels (Fig. 7b). Myh11 was also expressed (more than 10 reads) in 6/27 of S+L- cells, suggesting that these cells represent the fraction of VSMCs that did not successfully recombine the lineage label after tamoxifen treatment. Conventional VSMCs (S–L+; Fig. 7c, magenta triangles) and the majority of non-lineage-labelled Sca1-expressing cells (S+L–;
The cVSMC network was expressed highly in conventional S
contractile markers (highest level of co-expression consisted of 24 genes, including key analysis (WGCNA, Fig. 7d and S7c, d). The module showing the work modules according to weighted gene co-expression network Fig. 7a,b and Supplementary Data 6), which fell into three net-
VSMC populations. We identi-
blue circles) formed distinct clusters in a PCA based on the levels
of the 500 most variably expressed genes. In contrast, Sca1+
VSMC-lineage cells (S+L+; yellow squares) showed remarkable heterogeneity, with some S+L+ cells clustering together with conventional VSMCs, and others spreading away from them. These findings provided the initial evidence that VSMC-lineage Sca1+ cells are distinct from their non-VSMC-lineage counter-
parts and may show heterogeneity with respect to the contractile VSMC phenotype.

**Increased expression of VSMC response genes in Sca1+ VSMCs.** To identify gene signatures underlying the observed heterogeneity of S+L+ cells, we initially employed the robust HVG method used above for the analysis of the AA and DT VSMC populations. We identified 52 genes showing highly variable expression within the S+L+ population (Supplementary Fig. 7a,b and Supplementary Data 6), which fell into three network modules according to weighted gene co-expression network analysis (WGCNA, Fig. 7d and S7c, d). The module showing the highest level of co-expression consisted of 24 genes, including key contractile markers (Myh11, Acta2, Tagln and Cml1; Fig. 7d), and we will refer to it as the “cVSMC network”. As expected, cVSMC network was expressed highly in conventional S+L+ VSMCs, variably across S+L+ cells and at low levels in S+L- cells (Supplementary Fig. 7e).

The reduced levels of cVSMC markers in a subpopulation of S
+L- cells that did not cluster with conventional S−L+ VSMCs (Fig. 7e) prompted us to search for genes that are upregulated as the contractile identity is lost. We assigned each cell a “cVSMC score” based on expression of the cVSMC network genes (Supplementary Data 7; see Methods for details). Using negative binomial regression, we detected 312 genes whose expression positively correlated with the cVSMC score in S+L+ cells (cVSMCpos), and 303 genes showing a negative correlation (cVSMCneg) (likelihood ratio test, fdr-adjusted p-value < 0.05; see Methods; Fig. 7f and Supplementary Data 8).

Genes that correlated positively with the cVSMC score were generally expressed at higher levels in S−L- cells compared with S+L- cells and showed variable expression in S+L- cells (Fig. 7g). As expected, cVSMCpos genes showed enrichment for ontology terms describing functional features of contractile VSMCs such as “muscle contraction” (Tpm1, Mylk), “actin filament organisation” (Actn4, Dstn) and “cell-substrate adhesion” (Sdc4, Itgb1) (Fig. 7h and Supplementary Data 9). In contrast, genes showing a negative correlation with the cVSMC score (cVSMCneg) were enriched for ontology terms associated with vascular development, potentially reflecting a less differentiated state compared with contractile...
VSMCs (Fig. 7h). Consistent with this, cVSMCneg genes included classical "synthetic" VSMC markers Mgp, Col8a1 and Spp1 and showed enrichment for terms associated with functional characteristics of phenotypically switched VSMCs, including migration (such as Pak3, Igf1 and Igfbp5), proliferation ("cell growth", "wound healing") and secretion of extracellular matrix components ("vesicle organisation" and "extracellular matrix structure") (Fig. 7i and Supplementary Data 9). We also found enrichment for genes associated with activation of many signalling pathways (PI3K, small GTPases, Tgf-beta, chemokines), suggesting that the cVSMCneg gene signature reflects a responsive cell state.

Notably, a number of GO terms (including cell adhesion, migration and ERK1/2 signalling, Fig. 7h) showed enrichment among both cVSMCpos and cVSMCneg genes. However, we noted that in these cases the positively and negatively correlated genes promote different functions within the same broad pathways. For example, inspection of the focal adhesion pathway annotated in the KEGG database suggested that cVSMCpos genes enhance stress fibre formation, while cVSMCneg genes promote cell proliferation (e.g., Pak3, Supplementary Fig. 8a). Similarly, in the MAPK pathway, genes that correlate positively with the cVSMC score are generally inhibitory, whereas negatively correlated genes promote growth factor signalling (Supplementary Fig. 8b).

Collectively, our findings demonstrate that S+L+ cells are a heterogeneous cell population showing progressively lower levels of contractile VSMC genes (cVSMCpos) and higher expression of cVSMCneg genes that are associated with responses to diverse signals.

**Sca1 expression is a hallmark of VSMC stimulation.** To test whether Sca1 expression is linked to VSMC phenotypic modulation, we initially compared levels of Ly6a/Sca1 transcripts in VSMCs isolated from healthy tissue with cultured VSMCs, representing a partial model of phenotypic switching. Increased Ly6a/Sca1 expression in the cultured cells was evident by both RT-qPCR analysis (Fig. 8a) and in scRNA-seq datasets (Fig. 8b).

The increased Sca1 expression in cultured VSMCs could result from selective expansion of pre-existing Sca1+ medial cells or upregulation of Ly6a/Sca1 transcription in contractile VSMCs upon phenotypic switching. To address this, we sorted Sca1 (GFP)-negative medial cells from Sca1-GFP animals and analysed GFP expression at different timepoints during in vitro culture. At day 3, no GFP expression was observed in the medial sample, while it was abundantly detected in sorted Sca1(GFP)-positive
Myh11-CreERT2/Confetti lineage labelling

FACS sorting based on Sca1 and lineage label

Smart-seq2 scRNA-seq

Single-cell suspension

Sca1 staining

Myh11 Ly6a/Sca1

Log expression

S–L+
S+L+
S+L–
S–L–

Sca1

S+L–
S+L+
S–L+

S+L–
S–L+
S+L+

cVSMC score

S+L+ cells ordered by cVSMC score

Regulation of pi3k signalling

Pos. reg. of response to external stimulus

Pos. reg. of TGF-β production

Reg. of vasculature development

Vesicle organisation

Reg. of cell shape

Small GTPase mediated signal transduction

Chemokine-mediated signalling pathway

Cell chemotaxis

ERK1 and ERK2 cascade

Extracellular structure organisation

Response to lipoprotein particle

Cell growth

Smooth muscle cell migration

Pos. reg. of cell migration

Cellular response to TGF-β stimulus

Angiogenesis

Neg. reg. of smooth muscle cell prolif.

Wound healing

Reg. of actin filament-based movement

Cell−substrate adhesion

Muscle hypertrophy

Integrin-mediated signalling pathway

Blood circulation

Muscle cell differentiation

Actin filament organisation

Potassium ion transport

Myh11

Flna

Cav1

Myl9

Cnbp

Cd81

Cav1

Myl9

Cnbp

Cd81

Fhl1

Myh11

Confetti

0
5
10
15

Myh11 Ly6a/Sca1

0
10
20

PC1

PC2

Myh11

Mgp

Colla1

Cacna1

Igf1

Pak3

Vcam1

Col8a1

0
2000
4000

0
2000
4000

Myh11

0
2000
4000

2.5
0.0
−2.5

Expression score

Expression

Gene ratio

adj. p-value

0.01
0.02
0.03
0.04
0.05
0.06
0.08

Gene ratio

pos

neg

235
289
adventitial cells (Fig. 8c). However, longer exposure to culture conditions (7 days and more) induced GFP expression in medial cells, and on day 11, 15–28% were GFP-positive (Fig. 8c, d).

To test whether upregulation of Ly6a/Sca1 is also observed in vivo, we analysed Sca1 expression by flow cytometry at different timepoints after induction of VSMC lineage labelling in Myh11-CreERT2/EYFP mice. We observed a small but significant increase in the fraction of VSMC-lineage cells that were Sca1-positive over time, as confirmed by logistic regression (Fig. 8c).

The identified trend appeared stronger, rather than weaker, when the confounding effects of animal age were regressed out (Supplementary Fig. 9), suggesting that it was not due to age-related differences.

Finally, we tested whether Sca1 upregulation is also a feature of phenotypic modulation of VSMCs in vivo using a vascular injury model. Lineage-labelled Myh11-CreERT2/EYFP animals underwent carotid ligation injury, which results in the acute down-regulation of contractile genes and activation of proliferation. Flow cytometric analysis of dispersed carotid arteries from these lineage-labelled animals demonstrated that Sca1 expression is induced in 10–45% of VSMCs (EYFP+) 8 days after ligation (Fig. 8f–h).

Taken together, these results provide in vitro and in vivo evidence that VSMCs upregulate Sca1 in response to stimuli such as culture conditions and vascular injury.

**Shared signature of healthy and plaque Sca1+ VSMCs.** The evidence for expanded numbers of Sca1+ medial cells following vascular injury (Fig. 8f–h) prompted us to investigate this population in atherosclerotic plaques. To this end, we used tamoxifen-labelled Myh11-CreERT2/Confetti mice on an ApoE−/− background, where atherosclerosis is induced by feeding a cholesterol-rich diet. We isolated and digested atherosclerotic plaques from these samples with 10X Chromium technology. As shown in the t-SNE plot (Fig. 9a), most of the VSMC-derived plaque cells formed one large population, which could be further subdivided into nine clusters (clusters 0–8; Fig. 9a). Additionally, a clearly distinct population (cluster 9; Fig. 9a, pink) was formed by cells expressing macrophage markers (e.g., Cldn8, Lys2 and Fcgr1g).

In contrast to the low frequency of Sca1+ expressing cells observed in healthy tissue, Ly6a/Sca1 was detected in a significant proportion of VSMC-derived plaque cells (Fig. 9b). Most Ly6a/Sca1+ cells were contained within cluster 7 (Fig. 9a, lilac), indicating that Sca1 expression marks a distinct type of VSMC-derived plaque cells. We next examined whether Sca1-expressing VSMC-lineage cells in the plaques express the transcriptional signatures of their Sca1-positive counterparts (S+L+) found in healthy vessels. We found that VSMCpos genes associated with the contractile phenotype were depleted in Sca1+ plaque cells compared with the majority of the profiled cells (Fig. 9c).

In contrast, cVSMCneg genes (associated with response to signals) showed enrichment in the Sca1-expressing cluster (Fig. 9d). These results demonstrate that Sca1-positive VSMC-lineage cells from the plaques share transcriptional features of the rare S+L+ cells detected in healthy animals.

Notably, cluster 9 (Fig. 9a, pink) and a subpopulation of cells in the main cluster, which do not express Ly6a/Sca1 (cluster 8; Fig. 9a, magenta), expressed cVSMCneg genes at higher levels than most plaque cells, suggesting that these cells could be related to the Sca1-positive population. Cluster 8 cells express chondrocytic genes (Sox9, Hspa), consistent with a calcifying phenotype, whereas cells within cluster 9 have features of macrophages, as mentioned above. Notably, cVSMCpos genes were expressed at reduced levels in both of these cell populations compared with the S+L+ cell population (Fig. 9c, e), which may reflect their further progression away from the contractile state compared with the Sca1-positive cells.

Previous studies have suggested that the diversity of VSMC-lineage cells in the atherosclerotic plaque arises from clonal expansion of either one or very few cells. However, it is not known whether Sca1-positive and Sca1-negative VSMC-derived plaque cells coexist in the same clones or, alternatively, arise from independent clonal origins. To test this directly, we stained plaques from Myh11-CreERT2/Confetti/ApoE−/− mice fed a cholesterol-rich diet for Sca1. We found that Sca1 is expressed in VSMC-derived cells within atherosclerotic lesions (Fig. 9h), consistent with previous observations. Sca1-positive Confetti+ cells were located in the core region of the plaque, whereas Sca1 was not detected in the plaque cap (Supplementary Fig. 10). Yet, the clones that contained Sca1+ Confetti+ cells spanned both the cap and core regions, suggesting that Sca1-positive and Sca1-negative VSMC-derived plaque cells originate from a single cell. Taken together, these results indicate that dynamic modulation of Sca1 expression in VSMC-lineage cells is a hallmark of plaque development.

**Discussion**

In this study, we have delineated VSMC heterogeneity and identified the molecular differences that may underlie their selective clonal expansion in disease. We observe considerable heterogeneity between VSMCs within and between vascular beds.
in healthy animals, and yet find that VSMCs uniformly show hallmarks of their vascular regional identity. We identify a rare population of Sca1-positive VSMC-lineage cells in healthy vessels and provide evidence from disease models that Sca1 expression marks VSMCs undergoing phenotypic switching.

The observed differential expression of transcription factors and signalling molecules between VSMCs from the aortic arch versus the descending thoracic aorta suggests distinct gene regulatory wiring between cells in these regions, which may underlie the reported regional differences in disease response43. This is consistent with classical transplantation experiments44 and ex vivo studies17 showing that gene expression and disease susceptibility of VSMCs is independent of vascular environment and growth conditions. In turn, the significant within-region VSMC heterogeneity in expression of genes involved in VSMC biology and disease, which we report, is consistent with that previously observed for G-protein coupled receptor expression by single-cell RT-qPCR45. Such heterogeneity is noteworthy as it indicates the existence of specific subsets of cells with particular disease-relevant properties. Sources of heterogeneity may include further
Fig. 8  Sca1 is upregulated in response to VSMC stimulation in vitro and in vivo. a Relative expression (log2-transformed) of Myh11 and Ly6a/Sca1 in ex vivo (black) and cultured mouse aortic VSMCs at passage 4–5 (red) determined by RT-qPCR, normalised to housekeeping gene expression (Hmbs). Lines show mean from analysis of three independent primary cultures, error bars show s.e.m. Differences in Myh11 (p = 0.001) and Ly6a/Sca1 (p = 5.1e−10) expression are statistically significant (student’s t-test). b The PCA plot of single-cell expression profiles for ex vivo VSMCs (squares) and cultured VSMCs (triangles) shown in Fig. 1e, with expression level of Ly6a/Sca1 colour-coded from light to dark grey. c, d Images (c) and GFP-signal quantification (d) of FACS-isolated medial cells from Sca1-GFP animals (sorted as GFP-negative, n = 4, top row in c) with Sca1-GFP adventitial (Adv, sorted as GFP- or GFP+ (middle row in c), tissue from four animals was pooled) and wildtype (WT) medial cell controls (sorted as GFP-, lower row in c). Cells were cultured for 11 days before fixation and confocal imaging. e Epifluorescence images showing GFP signal after 3 or 10 days of culture. Scale bars are 100 µm. d Quantification of GFP signal in each population showing the number of GFP-positive cells as a percentage of the total number of DAPI-positive cells. Images for quantification were taken in a single z plane. Individual replicates and their mean are indicated and error bars show s.e.m. f Logistic regression analysis of the relationship between Sca1+ and LCM+ cells and time after lineage labelling (logit-link logistic regression coefficient = 0.016+/−0.005 [mean+/−95% confidence interval], p-value based on Student’s distribution = 2.56e−10). Trendline and data points, colour-coded by animal age (black gradient), are shown. Age was not included in the regression model presented here; the model accounting for both time after labelling and age is shown in Supplementary Fig. 9. g, h FACS plots showing EYFP and Sca1 (APC) expression in cells from the left common carotid artery (LCCA) isolated from tamoxifen-labelled Myh11-CreERt2/EYFP no injury controls (f) or eight days after ligation (g). h The percentage of lineage labelled cells (EYFP+) that expressed Sca1 in the LCCA isolated from ligated and no injury controls. Dots represents data from independent animals (n = 5 for each group), lines show means, and error bars s.e.m.
developmental differences within the VSMC lineage, differences in the cells’ microenvironment and haemodynamic stress or stochastic factors. Analyses focused on individual genes are required to further investigate the mechanisms and functional consequences of VSMC heterogeneity in healthy vessels.

We explored the functional relevance of a rare subset of VSMC-lineage cells that express Sca1, which is encoded by one of the identified variably expressed genes (Ly6a/Sca1). We show that Sca1+ VSMC lineage cells constitute a heterogeneous cell population. Some of these cells exhibit a contractile VSMC (cVSMC) signature, as evidenced by expression of cVSMCpos genes. In contrast, other Sca1+ VSMC lineage cells express reduced levels of the contractile VSMC signature and instead manifest a “VSMC response signature” (represented by cVSMCneg genes), including markers of synthetic VSMCs. We further demonstrate that the transcriptional profile associated with these non-conventional Sca1+ VSMC-lineage cells in healthy vessels (VSMC response signatureHigh, cVSMC signatureLow) is shared by Sca1+ cells within atherosclerotic plaques. We therefore propose that Sca1+ VSMC-lineage cells within healthy tissue could represent a more plastic state, which would readily respond to injury and inflammation (Fig. 10). Consistent with this idea, we show that Sca1 is upregulated in VSMCs upon exposure to stimuli known to induce phenotypic switching. In contrast, Sca1 is not expressed in VSMC-derived plaque cells that display the contractile signature (including cells in the cap). Sca1 is also absent in VSMC-derived plaque cells that have adopted alternative identities, including those expressing chondrocyte markers, which nonetheless also display the VSMC response signature characteristic of Sca1-positive cells. We therefore propose that Sca1 expression during plaque development could mark an intermediate VSMC state, which might give rise to phenotypically distinct cells within lesions (Fig. 10).

Sca1-expressing cells in the vessel wall have been reported previously. In particular, a medial “side population” isolated on the basis of reduced Hoechst staining (reflecting the expression of the ABCG2 transporter) was found to be enriched for Sca1-positive cells. These “side population” cells expressed low levels of lineage-specific marker genes, but upregulated VSMC or endothelial markers in vitro in response to PDGF-BB/VEGF, respectively, indicating an increased plasticity. Expression of Abcg2 was detectable in 6/92 of the S+L+ cells we analysed, suggesting some overlap between these cells and the “side population”. S+L+ cells also showed increased expression of some endothelial-associated genes, including Vcam1 and Fli1, but whether endothelial cell function can be induced in these cells remains to be tested. Vascular wall-derived Sca1-expressing cells capable of differentiating into SMCs in vitro and in vivo have also been detected in the adventitia (AdvSca1). However, transcriptional profiling of a subset of the AdvSca1 cells (marked by Sonic Hedgehog (Shh)-response gene Gli1) showed expression of genes that are either not expressed in S+L+ cells (e.g., Eng/Cd105 and Sox2), or that form part of the cVSMC signature, which is expressed in both S+L+ cells and contractile VSMCs (Smtn, Tagln, Ccn7, Itgb1)31. Finally, Sca1 is expressed by a population of angiogenic progenitors residing in the vascular wall. The markers of these cells (Pecam1/Cd31, Kit/Cki, Cdh5/VE-Cadherin, Eng/Cd105) do not form part of the VSMC response signature identified in S+L+ cells, but notably, were detected in many S+L- cells. In conclusion, the transcriptional signature of S+L+ cells suggests that medial Sca1-expressing cells are distinct from adventitial and angiogenic progenitor cells. However, further analyses are required to comprehensively compare the previously reported cell populations in the vessel wall with the Sca1+ VSMC-lineage cells detected in our study.

A number of independent lineage tracing experiments have demonstrated that cells expressing contractile VSMC markers

---

**Fig. 9** The transcriptional signature of S+L+ cells from healthy vessels is expressed in VSMC-derived plaque cells. a t-SNE visualisation of 10X Chromium dataset (3314 cells) of lineage-positive cells isolated from aortic plaques and the underlying media in Myh11-CreERT2/Confetti mice that were first tamoxifen-treated and then fed a cholesterol-rich diet for 14 or 18 weeks (tissue from three and two animals respectively was pooled for the analysis). The cells are grouped into a large population composed of 9 clusters, including one enriched for chondrocytic genes (cluster 8, magenta), and a smaller, distinct population, which was enriched for macrophage genes (cluster 9, pink). b t-SNE plot of 10X Chromium dataset from a with Ly6a/Sca1 expression colour-coded from light to dark grey. c, d t-SNE plot of 10X Chromium from a with expression scores for cVSMCpos (c) or cVSMCneg (d) genes colour-coded on a blue to red gradient. e Violin plots showing the expression score of cVSMCpos genes in cells mapping to each cluster shown in a. f Confocal image of a cryosection of the descending thoracic aorta from a Myh11-CreERT2/Confetti/ApoE+/− mouse fed a cholesterol-rich diet for 30 weeks and immunostained for Sca1. Signals for fluorescent Confetti proteins (GFP, green; RFP, red; YFP, yellow and CFP, blue) are shown in i–ii, nuclear DAPI (white) is shown in all panels, and anti-Sca1 (magenta) is shown in i, ii, and iv. The region outlined in i is magnified in ii–iv, with arrows pointing to cells that are double-positive for the Confetti lineage label (RFP) and Sca1. Image in i is a maximum projection of 16 z-slices (2 μm each) and ii–iv show a single 2 μm Z-slice. Scale bars are 40 μm (i) or 15 μm (ii–iv).

---

**Fig. 10** Model of VSMC priming and inflammation-induced phenotypic switching. Within the healthy vessel, priming of contractile VSMCs induces expression of Sca1 (black symbols) and Response Signature (RS, blue) genes, as well as the downregulation of the cVSMC signature (cVSMC, red). We propose that Sca1-expressing VSMC-lineage cells are hyper-responsive to inflammation and constitute an intermediate plastic population that potentially gives rise to phenotypically distinct, Sca1-negative VSMC-derived cells in the plaque (cap cells displaying the cVSMC phenotype, and core cells co-expressing the Response Signature with chondrocyte- (magenta) or macrophage-like genes (pink)).
(Myh11, Tagln) generate the majority of cells within atherosclerotic lesions8,9,10,11,12,13,14. However, other studies suggest that specialised Scal1+ progenitor populations also give rise to vascular lesions15,16,17,18,19. Possibly alongside VSMCs in a context-dependent manner20,21. Our observation that a small proportion of Myh11-lineage labelled VSMCs in healthy vessels express Scal1, and that Scal1 is upregulated in VSMCs in response to stimulation, may provide an explanation for these apparently discrepant observations. The co-expression of Scal1 and the VSMC response signature with the Myh11-lineage label both in healthy vessels and within plaques suggest that observed Scal1+ cells in vascular disease models are likely to originate from phenotypically modulated VSMCs. However, development of specific, dual-lineage tracing models is required to examine the functional role of the previously reported progenitor populations and the Scal1+ cells we have identified here in order to resolve this definitively.

The hallmarks of VSMC heterogeneity revealed in our study enable the isolation and functional analysis of specific VSMC subpopulations, paving the way for selective targeting of causative vessels and atherosclerotic plaques will enable further exploration of genetic and regulation of disease priming. Our study also highlights candidate genes for functional testing of how VSMC heterogeneity affects the response to inflammatory signals in disease and can be used to identify corresponding plastic cell populations in humans for targeting and diagnosis. Beyond the immediate findings of our study, the single-cell transcriptomes generated here for the whole aorta and VSMC-lineage cells in healthy vessels and atherosclerotic plaques will enable further exploration of vascular cell heterogeneity and function.

Methods

Animals and tissue dissection. All experiments were carried out according to the UK Home Office regulations under project licence PPL 7007565 and have been approved by the University of Cambridge Ethical Reviews Committee. C57Bl/6 animals were purchased from Charles River. Myh11-CreER2, Rosa26-Confetti (Confetti), Rosa26-EYFP (EYFP), Apo-ε-1, Scal1-GFP animals were described previously22,23,24,25,26,27,28. Males (the Myh11-CreER2 transgene is Y-linked) received 10 intraperitoneal injections of 1 mg/ml tamoxifen in corn oil, typically at 6 weeks of age, if not indicated otherwise, for lineage labelling. Aortas from young healthy male mice (8–14 weeks), unless otherwise indicated, were dissected free of fat and connective tissue. The aortic arch (AA, the aortal segment from just left of the branch point of the left subclavian artery) and descending thoracic (DT, the straight aortal segment from after the arch has clearly ended until the diaphragm) aortic segments were isolated and processed separately in experiments comparing vascular regions. The animals used for analysis of Sca1+ cells over time in vivo, were 9.5-5 weeks of age and were identified by confocal imaging. For assessment of Confetti signal outside of the medial layer (defined as the area between the internal and external elastic lamina), four cryosections were analysed for each of two different parts of three vascular regions (AA, DT and carotid arteries) of six Myh11-CreER2/Confetti animals either 1 or 6 weeks after tamoxifen labelling (three animals per time point, 72 sections in total). No difference in recombination frequency or detection of Confetti-positive cells in the endothelial and adventitial layers were detected between the two chase timings.

Imaging and image analysis. Live cultured cells were imaged using an Olympus IX71 fluorescence microscope with HCImage Live software (Fig. 8c). Confocal imaging was done using a Leica SP8 with LASX software and a 40x objective (63x, 1.4 numerical aperture, 7.66 µm/pixel) or a 100x objective (25x, 0.75 numerical aperture, 3.48 µm/pixel). Representative images are shown in Fig. 5c. Image analysis was performed using ImageJ or Imaris software. Staining intensities in Fig. 5c were calculated from maximum projections of 14-24 2 µm z-stacks, representative images are shown in Fig. 5b.

Bulk gene expression analysis. Dissected aortas were immediately transferred to RNA-later followed by isolation of A and DT segments before manual removal of the adventitial and endothelial cell layers. The cleaned medial layer from 3 to 5 animals was then lysed in Trizol (Thermo-Fisher) and RNA isolated. For bulk RNA-seq, the extracted RNA was cleaned on a RNeasy column (Qiagen) and quality-assessed on a Bioanalyzer (Agilent, RNA integrity number [RIN] 7.8-9). Sequences obtained from paired libraries were mapped to the total RNA reference genome using TopHat53 and assembled using StringTie54. StringTie and StringTie2 were used to quantify the expression levels of known and novel transcripts. The assembled transcriptome was uploaded to the UCSC Genome Browser for visualisation.

Analysis of Sca1 expression after carotid ligation surgery. Tamoxifen-injected Myh11-CreER2/Rosa26-EYFP animals were subjected to carotid ligation surgery26,29. Animals were anaesthetised using 2.5–3% isoflurane (by inhalation) and given a pre-operative analgesic (Temgesic). The left common carotid artery was tied firmly with one knot using 6-0 silk suture just below the bifurcation point and animals sacrificed 8 days after surgery. The carotid arteries were removed, dissected free from adipose and connective tissue and whole arteries were digested as described above until a single-cell suspension was reached. The dissociated cells were washed once in FACS buffer (1% BSA in PBS), incubated with FCx (1:100, Biogenoid) to block FcRII binding, stained with APC-conjugated anti-Scal1 (1:100, Miltenyi 130-120-343) or isotype control antibody (1:100, Miltenyi 130-102-655) and fixed through a 40 µm cell strainer. Stained cells were analysed on an Articulation (BD Bioscience) and FlowJo V10 software was used for data analysis and quantification.

Analysis of aortic cryosections. After tissue dissection, aortas were fixed in 4% paraformaldehyde in PBS for 20 min, followed by overnight cryoprotection in 30% sucrose in PBS, 1 h incubation in a solution of 50% TissueTek OCT C in 30% OCT and the remaining OCT and cultured at −80 °C. C.T. Serial transverse cryosections (12 µm) were then cut onto Superfrost Plus micro-scope slides, washed in PBS, stained with DAPI at 1 µg/ml in PBS and mounted in RapiClear 1.52 followed by confocal imaging. Cryosections from the aorta of animals fed a cholesterol-rich diet were stained for Scal1 before mounting: cryosections were washed in PBS, blocked for 1 h (10% goat serum in 0.1% BSA in PBS), incubated overnight at 4 °C with anti-Scal1 (1:200, Biogenoid 108101) or Alexo Fluor 647-coupled anti-rat IgG2a (1:200, Biogenoid 400526) in blocking solution, washed in PBS and incubated with Alexo Fluor 647-coupled anti-rat antibody (1:1000, Abcam ab51067) for one hour, washed in PBS, incubated with DAPI (1 µg/ml in PBS) to counterstain nuclei and mounted in RapiClear 1.52. For quantification of Confetti signal outside of the medial layer (defined as the area between the internal and external elastic lamina), four cryosections were analysed from each of two different parts of three vascular regions (AA, DT and carotid arteries) of six Myh11-CreER2/Confetti animals either 1 or 6 weeks after tamoxifen labelling (three animals per time point, 72 sections in total). No difference in recombination frequency or detection of Confetti-positive cells in the endothelial and adventitial layers were detected between the two chase timings.
Detection of highly variable genes (HVGs). HVG analysis was performed separately on AA, DT and S+L cells. We used a strategy based on variance decomposition of log-transformed normalised expression counts into technical variance (learned from data as a function of log-transformed mean count for the AA and DT cells and estimated using ERCC spike-in controls for S+L cells) and biological variance of interest. Variance components were estimated for each HVG, and genes with log-fold change > 1 and fdr-adjusted \( p \)-value < 0.01 were identified as differentially expressed.

Single-cell RNA-sequencing. Fluidigm C1: Single-cell suspensions of median AA and DT samples from wild type C57BL/6 males were prepared as described above and processed using a Fluidigm C1 system. Cell suspensions (100 cells/µl) were loaded on medium-sized (17–25 µm) Auto Prep Arrays (Fluidigm) and processed according to the manufacturer’s instructions. The loaded arrays were visually assessed under an inverted microscope to select capture sites containing a single cell, yielding a 40–70% capture efficiency. Cells included in the analysis were from two independent experiments from each aortic region. Cells were processed using the SMARTer Ultra® Low RNA Kit (Clontech) and amplified cDNA was isolated from the arrays.

Smart-seq2: For the analysis of the Scai+/ control population isolated from Myh11-CreERT2/Confetti animals, single-cell suspensions of the median AA and 5 to 7 animals per experiment were prepared and stained as described above. Individual Scai+/ and control Scai−negative cells expressing a single Confetti lineage label (L−1) or Scai+ cells expressing no Confetti marker (L−) were FACS-sorted on an Aria-Fusion flow cytometer (BD Bioscience) into separate wells of a 96-well plate. Scai+ and Scai− cells were included on the same plate and processed together to generate amplified cDNA using the Smart-seq2 protocol carried out as described previously65, with the following minor modifications: Primerscript (Clontech) was used for reverse transcription, 24 PCR cycles were used for amplification of cDNA, and ERCC control RNA (Invitrogen) was added (1,400,000,000 or 1,800,000,000 dilution in RT-mix). The analysed cells were from three independent experiments using six different medial cell suspensions.

10 Chromium: For analysis of whole aortas, aortas from three tamoxifen-labelled Myh11-CreERT2/Confetti males were dissected free of fat and connective tissue and subjected to enzymatic digestion. Single-cell suspensions were pooled, stained with 10,000 singlets, 10 µl of 0.1% BSA, 400 singlet, 20,000 singlet, live cells isolated by FACS-sorting on an Aria Fusion flow cytometer (BD Bioscience). Note that the Confetti lineage label was not used for sorting. For analysis of VSMC-lineage labelled cells from healthy vessels, medial single-cell suspensions from three tamoxifen-labelled Myh11-CreERT2/Confetti males were generated as described above and FACS sorted on an Aria Fusion flow cytometer (BD Bioscience) for expression of live (Zombie-negative), singlets expressing a single Confetti colour only (gating strategy provided in Supplementary Fig. 11). For 10X Chromium analysis of VSMC-derived atherosclerotic plaque cells, plaque cells were manually isolated by dissection from Myh11-CreERT2/Confetti/Ad5/Confetti animals that had been tamoxifen labelled and then fed a cholesterol-rich diet for 14 or 18 weeks (as described in Chappell et al.8). Plaque cells were digested as described56 and plaque cells from two-three animals per time point were pooled and processed in parallel, filtered through a 40 µm filter and 20,000 singlet cells expressing a single Confetti protein isolated by FACS-sorting on an Aria-Fusion flow cytometer (BD Bioscience) were provided as published in Supplementary Fig. 11, no selection for Scai-expression included). Sorted cells were pelleted and resuspended in 35 µl PBS to the GRCm38 mouse genome using Nextera library prep kit (Illumina). Libraries were analysed by paired-end.
The dependence of normalised expression counts for a given gene and cell on the cell’s distance along the PC1 axis $di$ was then assessed by negative binomial (NB) regression implemented in the R package VGAM\(^4\) v1.0-5, whereby the size parameter of the corresponding NB distribution was assumed to be the inverse fitted SCV obtained as described above. The significance of the regression fit was assessed by the likelihood-ratio test versus the intercept-only model and the resulting $p$-values were adjusted for multiple testing using the Benjamini-Hochberg procedure. Genes showing adjusted $p$-values below 0.05 were selected for further analysis.

**Analysis of the 10X Chromium single-cell RNA-seq data.** Raw sequencing reads were processed and aligned to the GRCh38 mouse genome through the 10X Genomics cellranger pipeline (v2.0 for VSMC-derived plaque cells and v2.1 for whole aorta cells and VSMC-lineage cells from healthy vessels). Quality control was performed based on total number of UMI counts per cell (atherosclerotic plaque dataset $>5000$ and $<20000$, whole aorta/lineage-positive medial cells $>1000$ and $<8000$ UMI counts), genes detected per cell (atherosclerotic plaque dataset $>1000$ and $<5000$, whole aorta/lineage-positive medial cells $>500$ and $<2500$ and percentage of mitochondrial reads per cell (atherosclerotic plaque dataset $<9$%, whole aorta/lineage-positive medial cells $<8$%). Reads from cells passing the quality control criteria were normalised and scaled using the NormalizeData (scale.factor = 10,000) and ScaleData functions from the CRAN R package Seurat\(^5\) v2.3.1. Principal component analysis was performed and the first five principal components used for clustering and t-distributed Stochastic Neighbour Embedding (t-SNE)\(^6\) visualisation using the Seurat package. Differential gene expression between clusters was determined using the “tobit” method of the FindMarkers function included in Seurat ($p_{adj} < 0.05$). For identification of genes upregulated in individual VSMC clusters in the analysis of cells from the whole aorta, minimum log fold change of 0.25 and minimum percentage of expressing cells of 25% were required for identification as an upregulated gene.

**Gene ontology analysis.** Gene ontology (GO) analysis of genes differentially expressed between AA and DT cells was carried out using the online Functional Annotation Clustering tool in The Database for Annotation, Visualisation and Integrated Discovery (DAVID\(^7\) v6.8) for all Biological Processes (BP) terms against all Mus musculus genes with medium classification stringency. Clusters were ranked by enrichment and fdr-adjusted $p$-values for the most significant term listed. GO enrichment analysis of genes positively and negatively correlated with cVSMC network expression was performed based on BP terms using clusterProfiler\(^8\) package in R v3.8.1. KEGG pathways\(^9\) in Supplementary Fig. 8 were visualised using the R package pathview\(^10\) v1.20.

**Code availability.** R packages were downloaded from CRAN and Bioconductor. Custom code will be made available upon request.

**Data availability**

Raw and processed sequencing data is available in the Gene Expression Omnibus (GEO) repository under accession number GSE117963. Previously published raw and processed sequencing data is available in the Gene Expression Omnibus (GEO) repository under accession number GSE79436. The dataset used and/or analysed during the current study are available in the European Nucleotide Archive (ENA) under the accession number PRJEB31966.

**Received:** 1 February 2018
**Accepted:** 19 September 2018
**Published online:** 01 November 2018

**References**

1. Bennett, M. R., Sinha, S. & Owens, G. K. Vascular smooth muscle cells in atherosclerosis. *Circ. Res. 118*, 692–702 (2016).
2. Bennett, M. R. & Owens, G. K. Epigenetic control of smooth muscle cell differentiation and phenotypic switching in vascular development and disease. *Annu. Rev. Physiol. 74*, 13–40 (2012).
3. Rensen, S. S. M., Doevendans, P. A. F. M. & van Eys, G. J. J. M. Regulation and differentiation and phenotypic switching in vascular development and disease. *Annu. Rev. Physiol. 74*, 13–40 (2012).
4. Pfaltzgraff, E. R. & Bader, D. M. Heterogeneity in vascular smooth muscle cell embryonic origin in relation to adult structure, physiology, and disease. *Dev. Dyn. 244*, 410–416 (2015).
5. Qiu, J. et al. Biomechanical regulation of vascular smooth muscle cell functions: from in vitro to in vivo understanding. *J. R. Soc. Interface 11*, 20130852 (2014).
6. Majesky, M. W. Developmental basis of vascular smooth muscle diversity. *Arterioscler. Thromb. Vasc. Biol. 27*, 1248–1258 (2007).
7. Chappell, J. et al. Extensive proliferation of a subset of differentiated, yet plastic, medial vascular smooth muscle cells contributes to neointimal formation in the injury and atherosclerosis models. *Circ. Res. 119*, 1313–1323 (2016).
8. Jacobsen, K. et al. Diverse cellular architecture of atherosclerotic plaque derives from clonal expansion of a few medial SMCs. *JCI Insight 2*, e95890 (2017).
9. Misra, A. et al. Integrin beta3 regulates clonality and fate of smooth muscle-derived atherosclerotic plaque cells. *Nat. Commun. 9*, 2073 (2018).
10. Islam, S. et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res. 21*, 1160–1167 (2011).
11. Tung, F. et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods 6*, 377–382 (2009).
12. Adhikari, N. et al. Guidelines for the isolation and characterization of murine vascular smooth muscle cells. A report from the International Cardiovascular Translational Research. *J. Cardiovasc. Transl. Res. 8*, 158–163 (2015).
13. Chamley-Campbell, J., Campbell, G. R. & Ross, R. The smooth muscle cell in culture. *Physiol. Rev. 59*, 1–61 (1979).
14. Trigueros-Motos, L. et al. Embryological-origin-dependent differences in homeobox expression in adult aorta: role in regional phenotypic variability and regulation of NF-kB activity. *Arterioscler. Thromb. Vasc. Biol. 33*, 1248–1256 (2013).
15. Van Assche, T. et al. Transcription profiles of aortic smooth muscle cells from atherosclerosis-prone and -resistant regions in young apolipoprotein E-deficient mice before plaque development. *J. Vasc. Res. 48*, 31–42 (2011).
16. Cheung, C., Bernardo, A. S., Trotter, M. W., Pedersen, R. A. & Sinha, S. Generation of human vascular smooth muscle subtypes provides insight into embryological origin-dependent disease susceptibility. *Nat. Biotechnol. 30*, 165–173 (2012).
17. Cai, Y. et al. Role of cAMP-phosphodiesterase 1C signaling in regulating growth factor receptor stability, vascular smooth muscle cell growth, migration, and neonatal hyperplasia. *Circ. Res. 116*, 1120–1132 (2015).
18. Breiman, L. Random forests. *Mach. Learn. 45*, 5–32 (2001).
19. Macaulay, I. C. et al. Single-cell RNA-sequencing reveals a continuous spectrum of differentiation in hematopoietic cells. *Cell Rep. 14*, 966–977 (2015).
20. Lun, A. T. L., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res. 5*, 2122 (2016).
21. Gunaje, J. J., Bahrami, A. J., Schwartz, S. M., Daum, G. & Mahoney, W. M. PDGF-dependent regulation of regulator of G protein signaling-5 expression and vascular smooth muscle cell functionality. *Am. J. Physiol. Cell. Physiol. 301*, C478–C489 (2011).
22. Zhang, H. et al. Origin-specific epigenetic program correlates with vascular bed-specific differences in Rgs5 expression. *Faseb J. 26*, 181–191 (2012).
23. Kolodziejczyk, A. A. et al. Single cell RNA-sequencing of pluripotent states unlocks modular transcriptional variation. *Cell Stem Cell 17*, 471–485 (2015).
24. Braun, M., Pietsch, S., Schröir, K., Baumann, G. & Felix, S. B. Cellular adhesion molecules on vascular smooth muscle cells. *Cardiovasc. Res. 41*, 395–401 (1999).
25. Sprague, A. H. & Khalil, R. A. Inflammatory cytokines in vascular dysfunction and vascular disease. *Biochem. Pharmacol. 78*, 539–552 (2009).
26. Ziegenhain, C. et al. Comparative analysis of single-cell RNA sequencing methods. *Mol. Cell 65*, 631–643 (2017).
27. Passman, J. et al. A sonic hedgehog signaling domain in the arterial adventitia supports resident Sca1+ smooth muscle progenitor cells. *Proc. Natl Acad. Sci. USA 105*, 9349–9354 (2008).
28. Hu, Y. et al. Abundant progenitor cells in the adventitia contribute to atherosclerosis of vein grafts in ApoE-deficient mice. *J. Clin. Invest. 113*, 1258–1265 (2004).
29. Kramann, R. et al. Adventitial MSC-like cells are progenitors of vascular smooth muscle cells and drive vascular calcification in chronic kidney disease. *Cell Stem Cell 19*, 628–642 (2016).
30. Sainz, J. et al. Isolation of ’side population’ progenitor cells from healthy arteries of adult mice. *Arterioscler. Thromb. Vasc. Biol. 26*, 281–286 (2006).
31. Kokkinopoulos, I. et al. Adventitial SCA-1+ progenitor cell gene sequencing reveals the mechanisms of cell migration in response to hyperlipidemia. *Stem Cell Rep. 9*, 681–696 (2017).
32. Psalitis, P. N. & Simari, R. D. Vascular wall progenitor cells in health and disease. *Circ. Res. 116*, 1392–1412 (2015).
33. Holmes, C. & Stanford, W. L. Concise review: stem cell antigen-1: expression, function, and enigma. *Stem Cells 25*, 1339–1347 (2007).
36. Kotton, D. N., Summer, R. S., Sun, X., Ma, B. Y. & Fine, A. Stem cell antigen-1 expression in the pulmonary vascular endothelium. Am. J. Physiol. Lung Cell. Mol. Physiol. 284, L990–L996 (2003).
37. Ma, X., Robin, C., Ottersbach, K. & Dzierzak, E. The Ly-6A (Sca-1) GFP transgene is expressed in all adult mouse hematopoietic stem cells. Stem Cells 20, 514–521 (2002).
38. 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).
39. Cherezova, O. A. et al. Oxidized phospholipids induce type VIII collagen expression and vascular smooth muscle cell migration. Circ. Res. 104, 609–618 (2009).
40. Kanelisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30 (2000).
41. Kumar, A. & Lindner, V. Remodeling with neointima formation in the mouse carotid artery after cessation of blood flow. Arterioscler. Thromb. Vasc. Biol. 17, 2238–2244 (1997).
42. Shankman, L. S. et al. KLF4-dependent phenotypic modulation of smooth muscle cells has a key role in atherosclerotic plaque pathogenesis. Nat. Med. 21, 628–637 (2015).
43. VanderLaan, P. A., Reardon, C. A. & Getz, G. S. Site specificity of atherosclerosis: site-selective responses to atherosclerotic modulators. Arterioscler. Thromb. Vasc. Biol. 24, 12–22 (2004).
44. HAIMOVICI, H. & MAIER, N. Fate of aortic homografts in canine atherosclerosis. Study of fresh homograft and thoracic aortic implants into thoracic aorta: role of tissue susceptibility in anatherogenesis. Arch. Surg. 89, 961–969 (1964).
45. Kaur, H. et al. Single-cell profiling reveals heterogeneity and functional patterning of GPCR expression in the vascular system. Nat. Commun. 8, 15709 (2017).
46. Sawada, H., Rateri, D. L., Moorleghen, J. J., Majesky, M. W. & Daugherty, A. Smooth muscle cells derived from second heart field and cardiac neural crest reside in spatially distinct domains in the media of the ascending aorta-brief report. Arterioscler. Thromb. Vasc. Biol. 37, 1722–1726 (2017).
47. Nemenoff, R. A. et al. SDF-1 induction in mature smooth muscle cells by inactivation of PTEN is a criticalmediator of excacerbated injury-induced neointima formation. Arterioscler. Thromb. Vasc. Biol. 31, 1300–1308 (2011).
48. Feil, S. et al. Transdifferentiation of vascular smooth muscle cells to macrophage-like cells during atherogenesis. Circ. Res. 115, 662–667 (2014).
49. Majesky, M. W. et al. Differentiated smooth muscle cells generate a subpopulation of resident vascular progenitor cells in the adventitia regulated by Klf4. Circ. Res. 120, 296–311 (2017).
50. Tang, Z. et al. Differentiation of multipotent vascular stem cells contributes to vascular diseases. Nat. Commun. 3, 875 (2012).
51. Roostalu, U. et al. Distinct cellular mechanisms underlie smooth muscle turnover in vascular development and repair. Circ. Res. 122, 267–281 (2018).
52. Yuan, F. et al. Contribution of vascular cells to neointimal formation. PLoS One 12, e0168914 (2017).
53. Trappell, C., Pachter, L. & Saleberg, S. L. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1111 (2009).
54. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).
55. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. 9, 171–181 (2014).
56. Butcher, M. J., Herre, M., Ley, K. & Galkina, E. Flow cytometry analysis of immune cells within murine aorta. JoVE https://doi.org/10.3791/2481 (2014).
57. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015).
58. Hartung, J. A note on combining dependent tests of significance. Biom. J. 41, 849–855 (1999).
59. Luzzi, C. & Kleber, C. Punitroots: panels with unit roots (The R Foundation, New York, 2015).
60. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 41–47 (2018).
61. Maaten, L. V. D. & Hinton, G. Visualizing data using t-SNE. J. Mach. Learn. Res. 9, 2579–2605 (2008).
62. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57 (2009).
63. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16, 284–287 (2012).
64. Luo, W. & Brouwer, C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 29, 1830–1831 (2013).

Acknowledgements
The authors would like to thank Paula Freire-Pritchett for help with data analysis, Simon Andrews, Steven Wingett, Felix Krueger at the Babraham Institute’s Bioinformatics facility for help with data management and initial processing, Kristina Tabbada and Clare Murnane (Babraham Institute sequencing facility) for Illumina sequencing, Arthur Davis (Babraham Flow Facility), the Wellcome Trust-Medical Research Council, Institute of Metabolic Science, Metabolic Research Laboratories, Imaging core, Wellcome Trust Strategic Award [100574/Z/12/Z] for technical assistance and the Cambridge National Institute for Health Research Biomedical Research Centre, Phenotyping Hub for cell sorting, Alison Finigan (Cardiovascular Medicine Division, University of Cambridge) and Mark Lynch (Fluidigm) for technical assistance, Hashem Koohy for advice on random forest analysis and all members of the Spivakov and Jørgensen labs for helpful discussions. A.L.T., J.C. and J.L.H. are supported by British Heart Foundation (BHF) studentships (FS/15/62/32032, RE/13/6/30180, FS/15/38/31516) and L.D. and E.O. are supported by BBSRC DTP studentships. E.D. is supported by the ERC AdG 341096. H.F.J. and M.R.B. are supported by the BHF Centre of Regenerative Medicine (RM/13/3/30159), the BHF Cambridge Centre of Research Excellence (RE/13/6/30180) and a BHF Chair award (CH/20000003). M.S. is supported by core funding from the Medical Research Council of the UK.

Author contributions
M.S. and H.F.J. designed the project; A.L.T., J.C., J.L.H. and H.F.J. performed the experiments; L.D., E.O., P.O., A.L.T., M.S. and H.F.J. analysed the data; M.S. and H.F.J. wrote the paper with contributions from all authors; M.R.B. and E.D. contributed reagents and analysis tools; M.S., H.F.J. and M.R.B. secured project funding; M.S. and H.F.J. supervised the work.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-06891-x.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.