Supplementary Methods

Isolation of nuclei from frozen tissue samples

For all experiments approximately 50 mg of frozen coronary artery tissue was ground into a fine powder using pre-chilled Mortar and Pestle with dry ice and liquid nitrogen. Frozen coronary artery tissue powder was then transferred back to a pre-chilled 1.5 mL microcentrifuge tube and kept on dry ice until all samples were broken down. We added 1 mL of cold 1X Homogenization Buffer (5 mM $CaCl_2$, 3 mM $Mg(Ac)_2$, 10 mM Tris pH 7.8, 320 mM sucrose, 0.1 mM EDTA, 0.1% NP-40, 0.1 mg/mL BSA, Roche cOmplete protease inhibitor) and the tubes were gently inverted 5 times and the powder gently pipetted up and down using a wide-bore 1 mL pipette tip set to a volume of 1 mL. The samples were then immediately transferred to cold, pre-chilled 1 mL glass Dounce homogenizers on ice.

We performed Dounce homogenization (10 strokes with Pestle A (Loose) and 20 strokes with Pestle B (Tight)) on ice and passed the lysate through a 70 µm Falcon strainer (Corning). The flow-through was collected and transferred to a chilled 2 mL Lo-Bind microcentrifuge tube (Eppendorf) and centrifuged at 4°C for 1 minute at 100 g. This supernatant was transferred to a new 2 mL Lo-Bind microcentrifuge tube (Eppendorf) and an OptiPrep (Iodixanol)/sucrose (Sigma) gradient was generated the same way as in the Omni-ATAC protocol[1]. 400 µl of sample was first mixed thoroughly with 400 µl of 50% Iodixanol by pipetting (to give 25% Iodixanol). Next, we layered 600 µl of 29% Iodixanol underneath and then 600 µl of 35% Iodixanol underneath the 29% layer. Samples were centrifuged for 20 minutes at 10,000 g at 4°C and the top layers were aspirated down to within 300 µl of the nuclei band.

At this stage we carefully took the band containing the nuclei (setting the pipette volume to 100 µl) and added the nuclei to 1.3 mL of cold Nuclei Wash Buffer (10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 1% BSA, 0.1% Tween-20) in a 1.5 mL Lo-Bind microcentrifuge tube. The microcentrifuge tube was inverted gently 5 times, nuclei gently mixed by pipetting (setting the pipette volume to 1 mL), and contents passed through a 40 µm Falcon cell strainer (Corning) into a new 1.5 mL Lo-Bind microcentrifuge tube (Eppendorf). Nuclei were pelleted by centrifugation for 5 minutes at 500 g at 4°C and supernatant carefully removed. Finally, this nuclei pellet was gently resuspended in 100 µl of the Nuclei Buffer provided with the kit (diluted from 20X Stock to 1X working concentration with nuclease-free water) by gently pipetting up and down. Samples and nuclei were kept on ice for all steps of the nuclear isolation. For each sample we measured the nuclei concentration by taking the mean of two separate counts using Trypan blue (Thermo Fisher) and the Countess II instrument (Thermo Fisher). Post cell lysis we generally observed less than 5% Live cells when visualizing with the Countess, consistent with proper lysis.

Single-nucleus ATAC GEM preparation

The full protocols for the single cell ATAC-seq data generation are available at the following link: https://support.10xgenomics.com/single-cell-atac.

We first performed a pilot study (4 samples) using the original Chromium Single Cell ATAC Reagent Kit protocol (10x Genomics, PN-1000111). The remaining samples were processed using the Chromium Next GEM Single Cell ATAC Reagent Kit v1.1 (10x Genomics, PN-1000175 and PN-1000176) protocol. Transposed nuclei (15 µl) were mixed with Barcoding Enzyme/Reagent Master Mix (60 µl) for a total of 75 µl per sample. This Transposed Nuclei + Master Mix volume was subsequently added to the 10x Genomics Chromium GEM chip. 50 µl of Gel Beads and 40 µl of Partitioning Oil were subsequently added to the chip and the chip covered with the 10x Gasket. The assembled chip was then run on the 10x Genomics Chromium Controller instrument to generate Gel Beads in Emulsions (GEMs). 100 µl of GEMs were then incubated in a thermal cycler (105°C lid temperature) using the following cycling conditions: 72°C for 5 minutes; 98°C for 30 seconds; 12 cycles of 98°C for 10 seconds, 59°C for 30 seconds, 72°C for 1 minute; then holding at 15°C. After incubation, GEMs were cleaned up using Dynabeads MyOne and SPRIselect reagent to produce a volume of 40 µl.

Single-nucleus ATAC library generation

Each sample was uniquely indexed with the Chromium i7 Multiplex Kit N, Set A, 96 rxns (10x Genomics) PN-1000084. The index sequences for this kit are listed in **Supplementary Table 2**. The sample index PCR reaction was prepared by mixing 40 µl of single-nucleus ATAC library with 57.5 ul of Sample Index PCR Mix and 2.5 µl of an individual primer from the Chromium i7 Sample Index N, Set A kit. The subsequent PCR amplification conditions were as follows: 98°C for 45 seconds; then 11 cycles of 98°C for 20 seconds, 67°C for 30 seconds, and 72°C for 29 seconds; 72°C for 1 minute; then holding at 4°C. After the sample index PCR we performed a final double sided size selection using SPRIselect beads. Before sequencing all snATAC libraries were run on the Agilent TapeStation High Sensitivity D1000 or D5000 ScreenTape. Library sizes are shown in **Supplementary Figure 1**.

Differential analysis between SMCs and fibromyocytes

Data processing

In order to explore differential regulatory profiles in SMCs and fibromyocytes, we performed another round of snATAC-scRNA-seq integration (using the data from Wirka et al. Nature Medicine 2019, GEO accession GSE131780)[2]. To delineate modulated SMCs (fibromyocytes) with high resolution and to further validate results from Wirka et al, we slightly increased the stringency of quality control for scRNA-seq data including coverage as an additional metric. Genes expressed in less than 5 cells were filtered out. Cells expressing < 500 and > 2500 genes, and with < 2000 UMIs were also trimmed from the dataset to prune defective cells, multiplets or cells with low coverage. Moreover, cells containing <1% and > 5% of reads mapping to the mitochondrial genome were discarded. Upon discarding lower quality cells, 7209 high quality cells remained for subsequent analysis. Read counts normalization and dimensionality reduction was performed as described above. To optimize clustering, several resolutions (granularity parameters) were applied to avoid under- or over-clustering of the data.

A resolution of 1.2-1.7 yielded a consistent number of clusters and upon identifying the differentially expressed genes for each, clusters could be successfully annotated using the human scRNA cluster-specific gene lists provided by Wirka et al 2019.

Differential accessibility, annotation and promoter analysis

Using cell type groupings defined by scRNA-seq label transfer, peaks with differential accessibility between Fibromyocytes and traditional Smooth Muscle Cells (SMC) were identified using a Wilcoxon-test as implemented in the ArchR package[3]. Peaks called during this analysis had a width of 500 bp. The threshold for differential peak significance was set at FDR <= 0.05 and Log2 fold change > 1, resulting in a total of 5681 significantly upregulated peaks and 2121 downregulated peaks. For differentially accessible promoter analysis, the promoter coordinates of protein coding genes were extracted using the R packages ensembldb v2.14.0[4] and EnsDb.Hsapiens.v86 v2.99.0 (Rainer J (2017). *EnsDb.Hsapiens.v86: Ensembl based annotation package*. R package version 2.99.0). These promoter coordinates were overlapped with upregulated and downregulated peak coordinates using the R package GenomicRanges v1.42.0[5].

As an additional approach for differential peak annotation, protein coding gene coordinates were extracted with ensembldb. Upregulated and downregulated peaks were annotated with the nearest protein coding gene using GenomicRanges v1.42.0. This annotation was validated using the R package ChIPseeker v1.26.0 [6] along with TxDb.Hsapiens.UCSC.hg38.knownGene v3.10.0 (Team BC, Maintainer BP (2019). *TxDb.Hsapiens.UCSC.hg38.knownGene: Annotation package for TxDb object(s)*. R package version 3.4.6).

Differential motif and Gene Ontology (GO) enrichment

Differential upregulated and downregulated peaks for fibromyocytes were converted into BED files and tested for TF motif enrichment using the command line tool HOMER v4.10 (http://homer.ucsd.edu/homer/ngs/) [7]. The findMotifsGenome.pl script was used to search for known and *de-novo* motifs. The analysis was run using default values, with the exception of the parameter “-size”, that was set to “-size given” to define peaks width from the BED files data instead of arbitrarily defining a constant value. The same BED files were used for region set enrichment analysis using the Genomic Regions Enrichment of Annotations Tool (GREAT v4.0.4) (http://great.stanford.edu/public/html/index.php)[8] using the whole genome as background (reported results are from the GO database).

Correlation of promoter profiles and integrated expression

For the coronary artery scRNA-seq data[5], individual cells from the different cell types were separated. For each cell type, we subsequently took the average of scRNA-seq read counts from all the related cells to represent the average expression index. For the snATAC data, individual cells from different cell types were similarly separated. For each cell type, reads from all the related cells were collected and piled up to generate pseudo-bulk data. The average

signal across +/- 3 kb centered on gene transcription start sites (TSS) were calculated as the promoter accessibility. Finally, for each cell type observed in both the scRNA and snATAC-seq datasets, we compared the promoter accessibility and average expression in log2 scale using scatter plots. The correlation coefficients were also calculated and labeled. This material is shown in **Extended Data Figure 3c**.

Super enhancer analysis

We used published H3K27ac ChIP-seq data in human coronary artery smooth muscle cells from our previous study [9] to identify the super enhancers in SMCs. In detail, the H3K27ac ChIP-seq reads were mapped to the hg38 genome with Bowtie2 (v2.3.5.1)[10] with default parameters (and -x 2000 as an additional parameter). The high quality (MAPQ > 30) unique mappable reads were kept for subsequent analyses. The H3K27ac genome-wide domains were identified with SICER (v2)[11] with default parameters. Next, the SICER peaks >= 10 kb were selected and those with non-overlaps with transcription start sites (TSS) +/- 3kb were defined as super enhancers. When comparing the effect of super enhancers to snATAC marker genes (from different cell types), we calculated regulatory potential (RP)[12] on each gene to estimate the regulatory effect of super enhancers on a given gene set. Only the reads/signal on the super enhancer regions were included in calculating the RP score. The RP scores were then log transformed (i.e., log2(RP+1)) and z-normalized for fair comparison. The normalized RP scores were subsequently plotted for different marker gene sets. Results are presented in **Supplementary Figure 4** and **Supplementary Table 5**.

Bulk coronary artery ATAC-seq

We generated bulk ATAC-seq libraries using coronary arteries from the same set of patients analyzed by snATAC-seq. We again used the Omni-ATAC protocol[1] to isolate nuclei from frozen coronary artery samples, but transposed 50,000 nuclei/sample for all bulk ATAC-seq reactions. After the density gradient centrifugation step using OptiPrep (Iodixanol)/sucrose (Sigma), 50,000 nuclei were transferred to cold ATAC RSB (Resuspension Buffer; 10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$) + 0.1% Tween-20 and centrifuged for 5 minutes at 4°C at 500 g. The supernatant was carefully removed and nuclei were resuspended in 50 µl of transposition mix (25 µl 2X TD buffer, 2.5 µl Tn5 transposase (Illumina), 16.5 µl PBS, 0.5 µl 1% digitonin, 0.5 µl 10% Tween-20), pipetted up and down 6 times, followed by shaking at 1000 rpm for 30 minutes at 37°C. ATAC-seq reactions were purified using the Zymo DNA Clean and Concentrator-5 Kit. The ATAC-seq libraries were amplified as per Supplementary Protocol 1 in the Omni-ATAC study using indexed oligos from Supplementary Table 1 of the original ATAC-seq paper [13]. Library size distributions were evaluated using Agilent TapeStation analysis and quantified using Qubit 3.0 fluorometer (Thermo Fisher Scientific) and KAPA qPCR (Kapa Biosystems). Libraries were multiplexed and sequenced to approximately 200 million paired reads per library on an Illumina NovaSeq.

Bulk coronary ATAC-seq libraries were analyzed using the PEPATAC pipeline[14] (version 0.8.6). Within this pipeline, adapter trimming was performed using Skewer[15] and reads were first pre-

aligned to the mitochondrial genome and a list of human repeats using Bowtie2 (-k 1 -D 20 -R 3 -N 1 -L 20 -i S,1,0.50-k 1 -D 20 -R 3 -N 1 -L 20 -i S,1,0.50). After discarding mitochondrial and repeat sequences, trimmed fastq files were mapped to the human hg38 genome using Bowtie2 (--very-sensitive --X 2000)[10]. Samtools[16] was used to remove poor quality reads (-q 10) and duplicate sequences were removed using the samblaster tool[17]. Peaks were called on sorted, deduplicated bam files using MACS2 (-f BED -q 0.01 --shift -100 --extsize 200 --nomodel --keep-dup all)[18] and ignoring the hg38 blacklist of peaks. Due to tissue availability we could not repeat bulk ATAC-seq for some bulk ATAC-seq libraries with lower quality. We were left with bulk ATAC-seq libraries for 35 out of the 41 individuals for downstream analyses.

To generate a bulk ATAC-seq counts matrix, we used bam files from these 35 individuals and the snATAC peak coordinates (in saf format) as input for featureCounts[19] with the -p flag for paired-end mode. For the bulk caQTL analysis for replication (n=35), we used the same conditions as the single cell caQTL analysis in RASQUAL[20]. We used the default RASQUAL createASVCF.sh script, the VCF file for these individuals (variants > 5% minor allele frequency), and the bulk ATAC bam files to generate the bulk allele-specific VCF file. rasqualTools was used to make compatible read count and sample specific offset files. For the bulk samples we only retained ATAC peaks with an average of 10 or more read counts across individuals. We again tested association for all variants within a +/- 10 kb window and ran RASQUAL with the -t flag to output the top associated SNP for each peak. We adjusted for age, sex, and the first three principal components of ancestry in the covariate file (-x flag). To obtain the null distribution of bulk q values we performed 5 different permutation runs for each cell type using the --random-permutation flag to break the relationship between genotype and peak accessibility. Finally, we adjusted for multiple testing using FDR corrections using the same method as for cell type caQTLs (https://github.com/natsuhiko/rasqual/issues/21).

Histological analysis and quantitation of atherosclerosis

For all histological staining, frozen sections (8 μm thickness) were prepared from OCT embedded human coronary artery segments adjacent to those used for snATAC profiling. A minimum of two sections per sample were blindly stained with each of the following stain, Oil Red O (ORO), Picro-Sirius red (PSR) and Hematoxylin and Eosin (H&E) at the UVA Research Histology Core. Briefly, for ORO staining, frozen sections were fixed in 10% Neutral Buffered Formalin solution, washed and stained in Oil Red O solution (Poly Science #s2120) for 5 min. After washing, slides were then stained in Hematoxylin solution (Richard Allen #7221) for 1 min before being rinsed and mounted with an aqueous mounting medium. For H&E, slides were stained using the Hematoxylin 360 reagents manufactured by Leica in an automated Gemini Stainer. For PSR, slides were placed in Picro-Sirius red solution (Direct Red 80 in saturated aqueous solution of picric acid) for 1 hour, rinsed in deionized water and were washed twice in acidified water. Slides were then dehydrated in ethanol, cleared in xylene and mounted.

Whole slide images were then captured at approximately 100,000 x 30,000 pixel resolution using a Hamamatsu NanoZoomer S360 Digital Slide Scanner C13220 at the Biorepository and Tissue Research Facility at UVA. Visualizations, annotations and quantifications were blindly

conducted within the PathcoreFlow workspace or using custom automated and semi-automated approaches as further detailed below.
A total of 18 donors across lesion stages were independently examined by two experimenters, and categorized into three different disease stages according to the Stary classification guidelines as follows (Cat 1: N=3; Cat 2: N=5; Cat 3: N=10).

To quantitate the lipid droplet area in ORO digitized images, we created a binary mask with the white region representing the target color (red) and black representing all other colors. An automated analysis was performed using a custom Python script using OpenCV package. The original images in the RGB (Red, Green, Blue) channel were first converted to the HSV (Hue, Saturation, Value) channel. In OpenCV, Hue has values from 0 to 180, Saturation and Value from 0 to 255. Thus, OpenCV uses HSV ranges between (0-180, 0-255, 0-255). Lipid droplets in ORO-stained images are predominantly red, so we first set the cv2.inRange function with the range of HSV values from true red color (0,255,255). To accommodate for color variations, we considered a range of HSV values for the red color. By checking the HSV color map in OpenCV, the red color has hue values in the range of 0 to 10. We also set the intensity value v in range(20,255) and saturation s in range(100,255). We used the cv2.inRange to generate a mask that has a value of 255 for pixels where the HSV values fall within the specified color range and a value of 0 for non-color pixels with values outside this interval. Pixel area measurements were converted to mm squared and normalized to the total tissue mask area.

To quantitate intima-media thickness (IMT) and distinguish the arterial wall layers, collagen was visualized using Picrosirius Red staining. We developed a semi-automated approach to quantitate the IMT from blinded Picrosirius Red digitized whole slide images. First, we used a manual free hand tool to trace the lumen, intima and media layers. The lumen boundaries were used to calculate the lumen area, lumen centroid and minimum and maximum lumen diameter. To accurately measure the IMT, we utilized the centroid as the center for remapping the three layer boundary image to a polar coordinate space. The mean, minimum and maximum IMT were then determined for each section.

Sample size estimation

While the overall sample size of this study was determined based on tissue availability, we performed several post-hoc power calculations. In the conventional sample size estimation for detecting cell type marker genes, we regarded snATAC accessibility-based gene scores of the marker gene in the target cell type as distribution 1 and the gene scores of the same marker gene in the rest of cell types as distribution 2. Subsequent power calculations for t-test were conducted based on these two distributions. We used the Python “statsmodels” package (v0.14) to carry out sample size power calculations applying a stringent cutoff (power=0.99, alpha=0.01) to ensure high confidence.

To demonstrate that we had sufficient power to detect cell types from our combined snATAC data, we performed single cell sample size calculation using SCOPIT [21] (**Supplementary Figure 7**). In detail, we estimated the sample size with the “Retrospective” mode on the

SCOPIT web server (http://navinlab.com/SCOPIT/) and set the “number of cells must be sequenced” to 50.

To confirm that we had the appropriate sample size needed for SMC and fibromyocytes differential accessibility comparisons, we performed a similar canonical 2-sample t-test power analysis as described above. Upon extracting ArchR differential accessibility-based gene scores using the same parameters as in the differential peak analysis (see above methods section), we calculated effect sizes for each gene using the gene score means and pooled standard deviations across cells from the SMC and fibromyocyte groups. To calculate the sample sizes required to reject the null hypothesis (no significant difference in accessibility-based gene scores between the two cell groups) at a given power value, we developed a custom python script using the statsmodels package. As before, we used stringent cutoffs for both the statistical power (0.99) and significance level (type 1 error rate, alpha = 0.01) to show the robustness of the differential accessibility-based analysis. As a result, we found that the required sample size for 93% of differential genes (based on accessibility-based gene scores) including fibromyocyte markers such as *FBLN2, LUM*, *F2R* and *TNFAIP6* was well below the number of cells available for the two groups (SMC n=6518, fibromyocytes n=2512) (**Supplementary Data 9**). We also calculated the mean required sample sizes across all differential genes for a wide range of power values (0.05-0.99) to further confirm that we were well-powered to detect significant accessibility differences between SMCs and fibromyocytes (**Supplementary Figure 7)**.

We used the single-cell powerEQTL tool [22] to demonstrate that we had sufficient power in our caQTL analysis (**Supplementary Figure 7).** While this tool assumes a standard linear mixed effects model, we estimated the minimum power (0.90) to detect caQTL variants at 5% minor allele frequency given the following parameters: alpha=0.05, n tested SNPs = 250,000, minimum effect size = 0.3, n cells = 650 (mean n SMCs/sample), log standard deviation (sigma) = 0.13, and intra-class correlation (rho) = 0.5. These estimates assume normally distributed chromatin accessibility levels after pre-processing. Given that we used allele-specific and total read counts to map caQTLs via RASQUAL, we expect at least 1.6x improved power for true positive caQTL discovery over linear models, based on previous estimates[20].

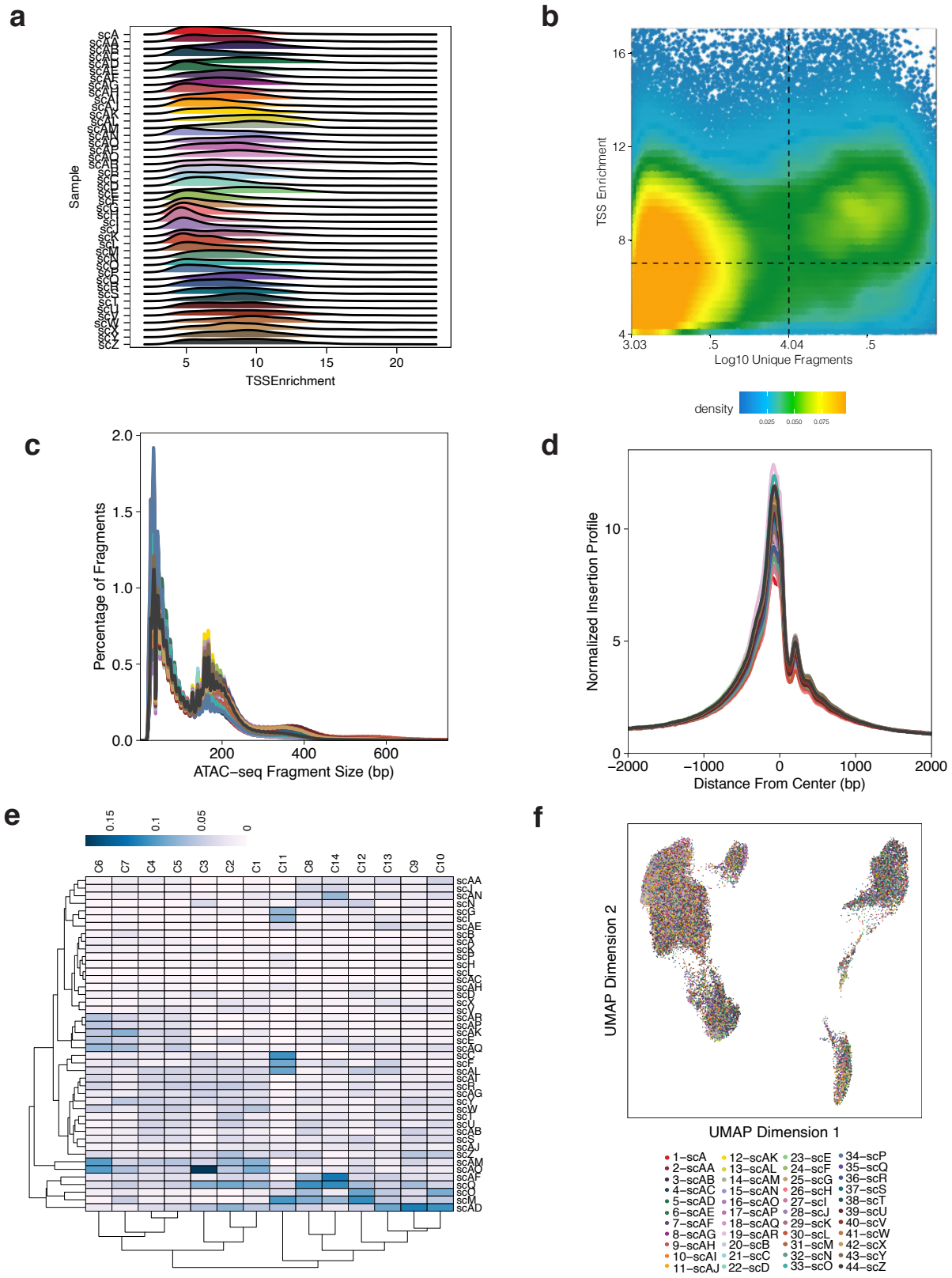
References

1. Corces, M. R. *et al.* An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
2. Wirka, R. C. *et al.* Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. *Nat. Med.* **25**, 1280–1289 (2019).
3. Granja, J. M. *et al.* ArchR is a scalable software package for integrative single-cell

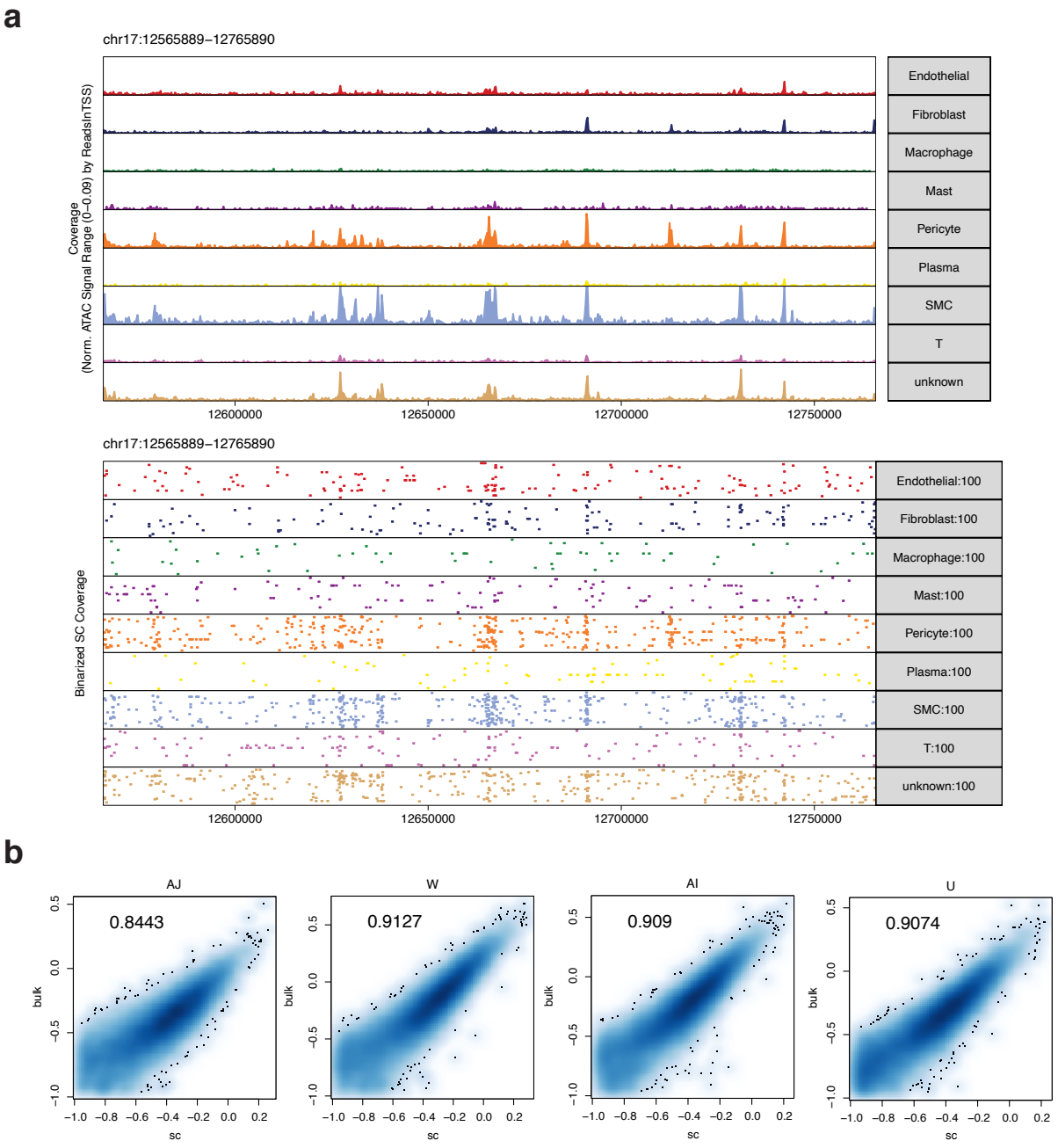
chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).

4. Rainer, J., Gatto, L. & Weichenberger, C. X. ensembldb: an R package to create and use Ensembl-based annotation resources. *Bioinformatics* **35**, 3151–3153 (2019).
5. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
6. Yu, G., Wang, L.-G. & He, Q.-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
7. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
8. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
9. Miller, C. L. *et al.* Integrative functional genomics identifies regulatory mechanisms at coronary artery disease loci. *Nat. Commun.* **7**, 12092 (2016).
10. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
11. Zang, C. *et al.* A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952–1958 (2009).
12. Wang, S. *et al.* 51 BETA. *Nat. Protoc.* **8**, 2502–2515 (2013).
13. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21.29.1-21.29.9 (2015).
14. Smith, J. P. *et al.* PEPATAC: an optimized pipeline for ATAC-seq data analysis with serial alignments. *NAR Genom. Bioinform.* **3**, lqab101 (2021).
15. Jiang, H., Lei, R., Ding, S.-W. & Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).

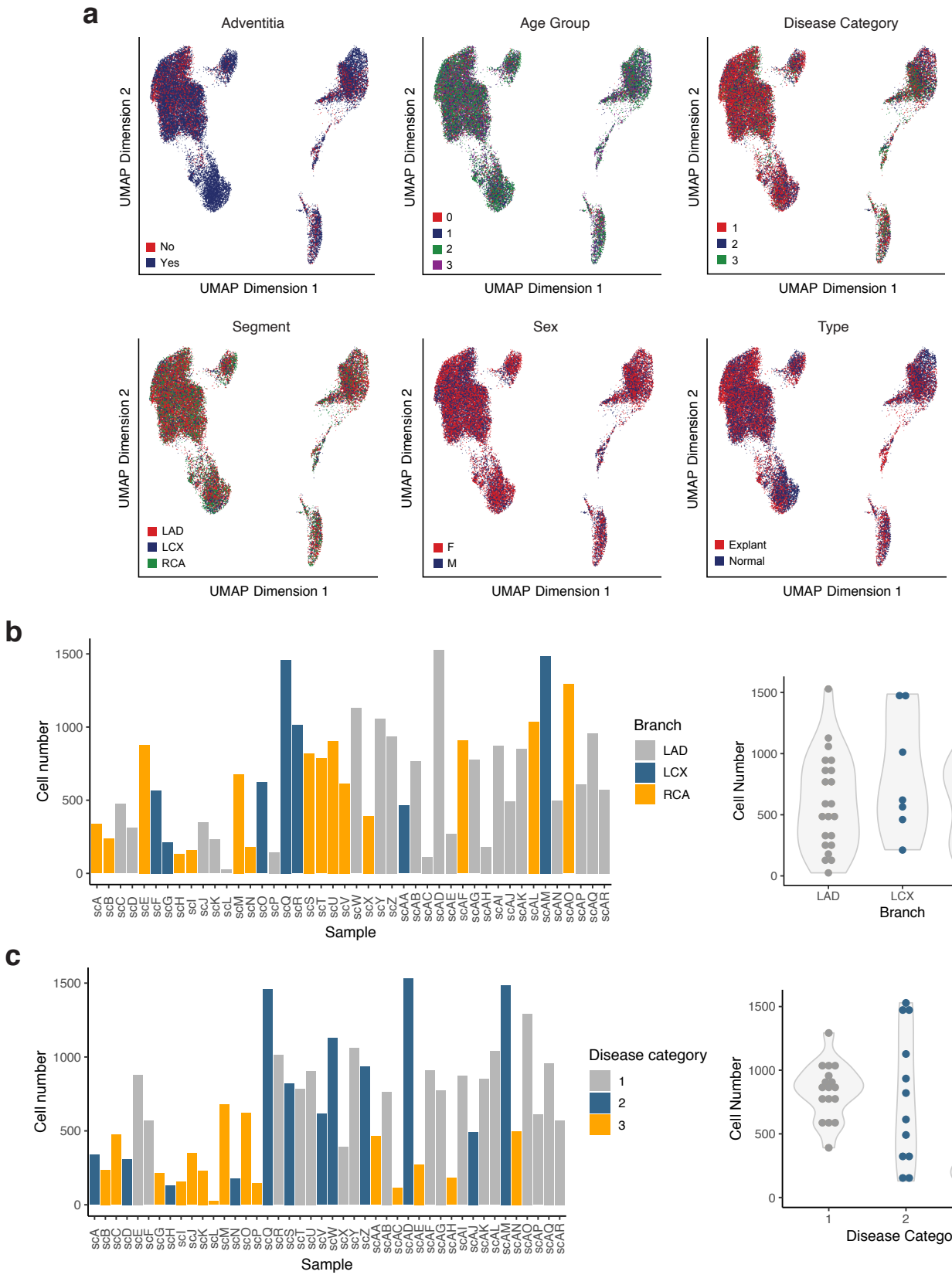
16. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
17. Faust, G. G. & Hall, I. M. SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics* **30**, 2503–2505 (2014).
18. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).
19. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
20. Kumasaka, N., Knights, A. J. & Gaffney, D. J. Fine-mapping cellular QTLs with RASQUAL and ATAC-seq. *Nat. Genet.* **48**, 206–213 (2016).
21. Davis, A., Gao, R. & Navin, N. E. SCOPIT: sample size calculations for single-cell sequencing experiments. *BMC Bioinformatics* **20**, 566 (2019).
22. Dong, X. *et al.* powerEQTL: An R package and shiny application for sample size and power calculation of bulk tissue and single-cell eQTL analysis. *Bioinformatics* (2021) doi:10.1093/bioinformatics/btab385.



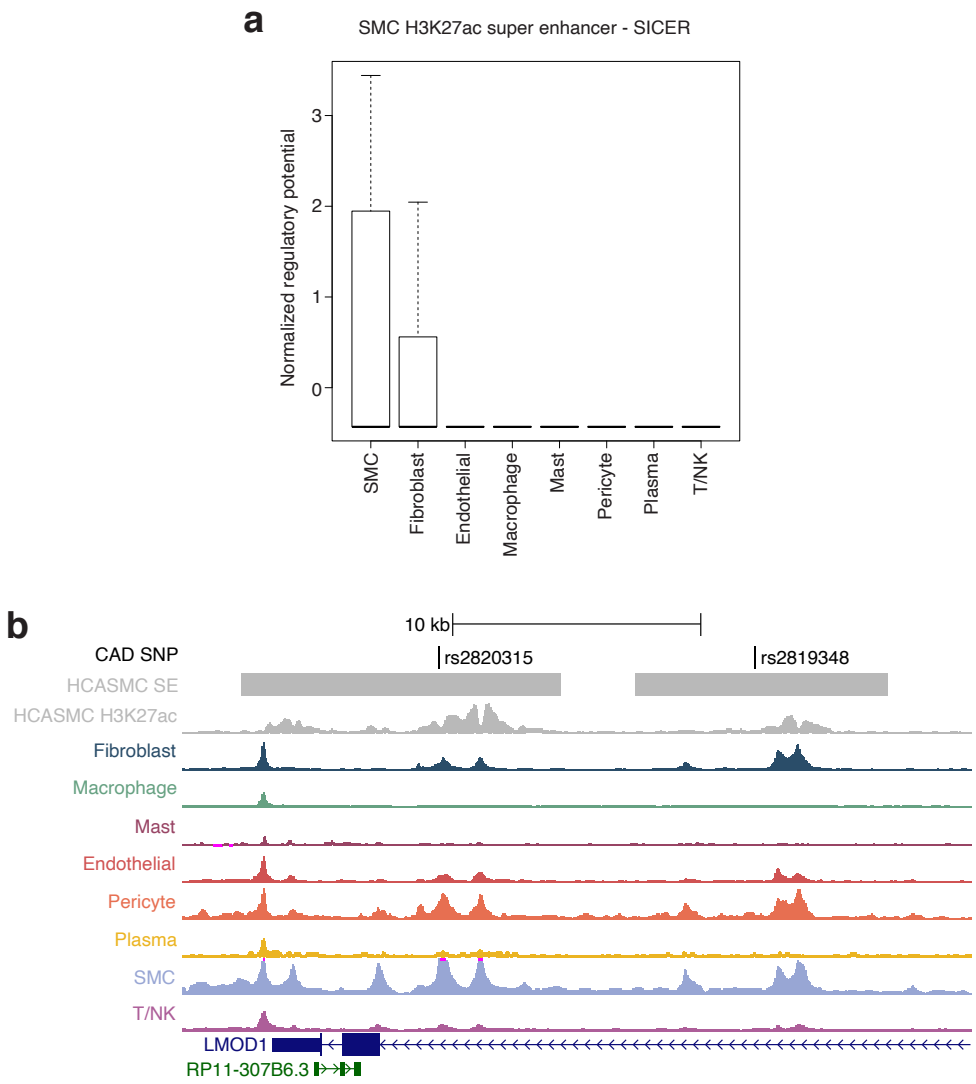
Supplementary Figure 1. Quality control (QC) for all coronary artery single nucleus ATAC libraries. (a) Distribution of transcription start site (TSS) enrichment scores for cells across all snATAC samples. (b) Density plot of TSS enrichment versus number of unique fragments (Log10) for total nuclei across all samples. We retained nuclei with TSS enrichments greater than or equal to 7 and more than 10,000 unique fragments. (c) Distribution of fragment sizes for all samples demonstrate typical ATAC-seq nucleosomal periodicity for cells passing QC. (d) TSS enrichment profiles for each coronary artery sample for nuclei passing QC. (e) Clustering all snATAC-seq samples according to cluster assignments. (f) UMAP of nuclei colored by corresponding sample.



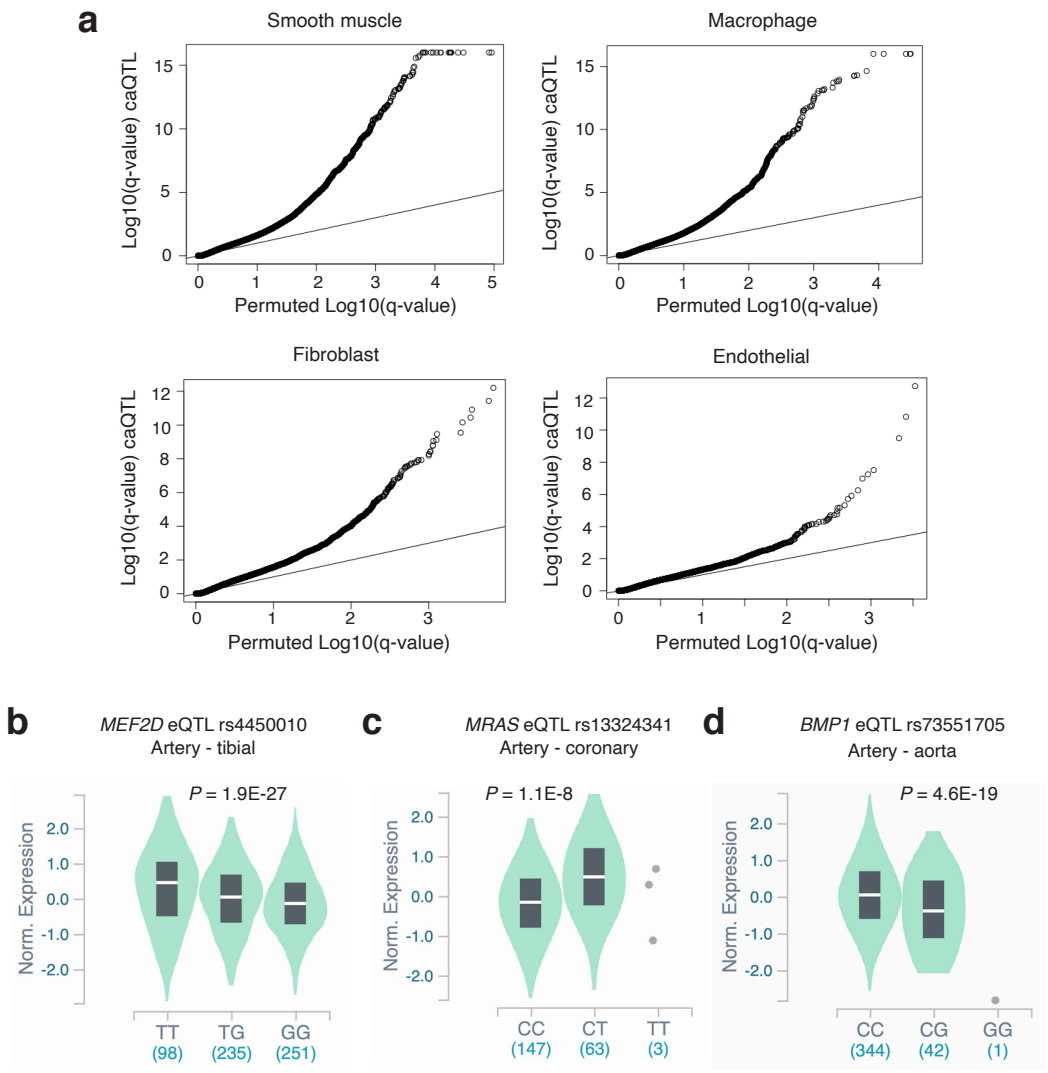
Supplementary Figure 2. (a) Genome browser tracks comparing accessibility profiles for aggregated nuclei for each cell type (top) with accessible fragments from 100 randomly selected nuclei in each cell type (bottom). Shown is the myocardin (MYOCD) locus that is an established smooth muscle cell marker gene. Genome browser tracks were plotted using ArchR. (b) Correlation of log2 normalized bulk and single-nucleus ATAC signal intensity for matched patient samples. Pearson correlation r values are shown for each sample.



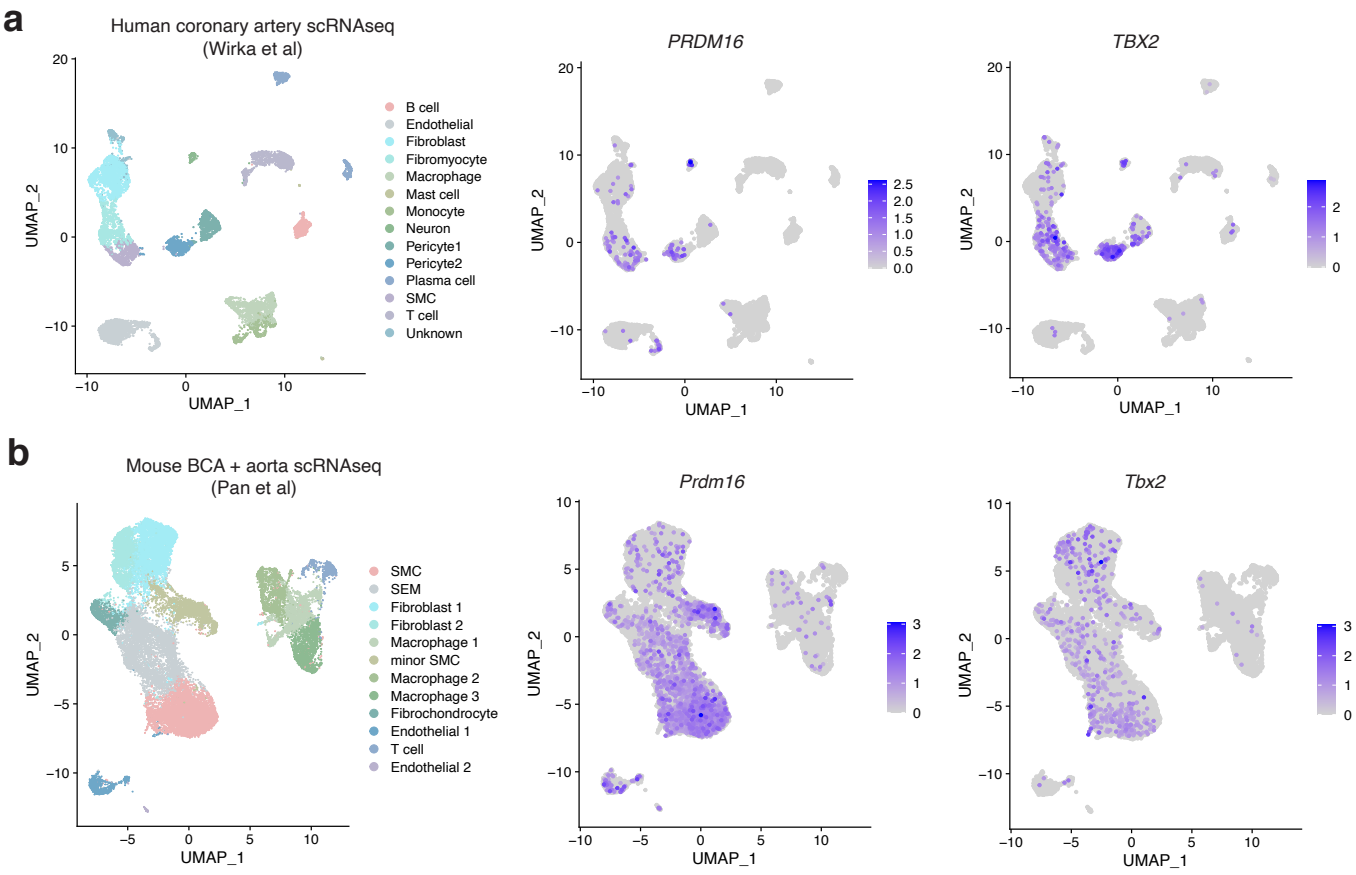
Supplementary Figure 3. (a) UMAP plots of coronary artery nuclei colored by sample and patient characteristics. Age group 0 (under 35), age group 1 (35-49), age group 2 (50-60), age group 3 (over 60). Disease categories described below. (b) Coronary samples colored according to derived coronary artery branch. Violin plots show overall distribution per category. (c) Coronary samples colored by disease category. Category 1: processed segment has no evidence of atherosclerosis; category 2: processed segment has no evidence of atherosclerosis but the patient has atherosclerosis; category 3: the processed segment is atherosclerotic.



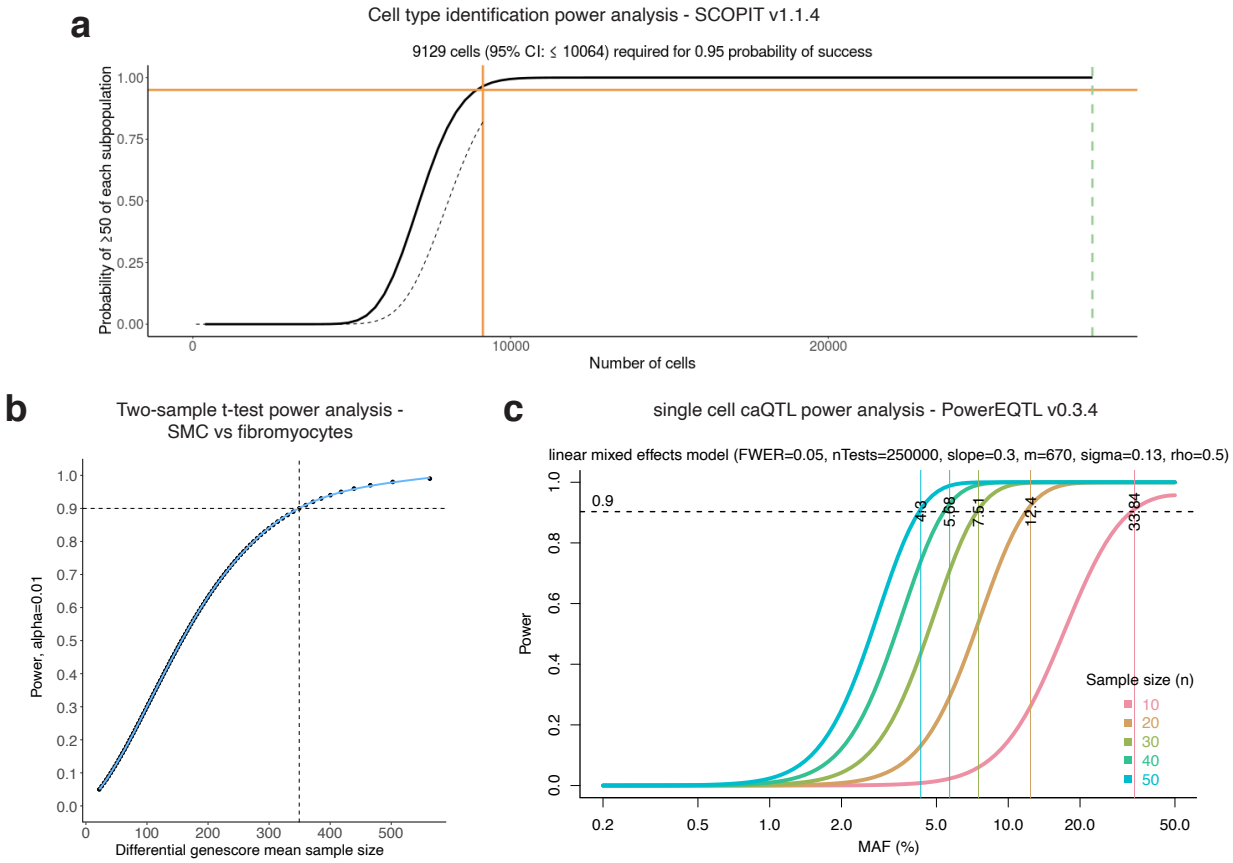
Supplementary Figure 4. (a) Normalized regulatory potential of long ATAC peaks (>10 kb) that are annotated as super enhancers (SE) from human coronary artery smooth muscle cell (HCASMC) enhancer histone modification H3K27ac, as detected using SICER. For the boxplot the centerline, bounds of box, top line and bottom line represent the median, 25th to 75th percentile range, 25th percentile - 1.5*interquartile range (IQR) and 75th percentile + 1.5*IQR, respectively. Number of cell type marker genes in the boxes (n): SMC:785; Fibroblast:477; Endothelial:499; Macrophage:1571; Mast:561; Pericyte:95; Plasma:1024; T/NK: 1138. (b) UCSC browser screenshot showing two fine-mapped CAD GWAS SNPs (rs2820315 and rs2819348) at the *LMOD1* locus overlapping super enhancers (grey boxes) identified from H3K27ac ChIP-seq in HCASMC. Individual coronary artery cell type specific chromatin accessibility tracks are also shown below.



Supplementary Figure 5. Chromatin acccessibility QTL quality control and top CAD replicated eQTLs. (a) Quantile-quantile (qq) plots of observed lead caQTL q values (log10) from RASQUAL against the permuted lead caQTL q values (log10) obtained using the '-t' option in RASQUAL. (b-d) eQTL violin plots for top CAD caQTL variants in GTEx artery tissues. In GTEx (v8) there are n=584 samples with donor genotypes for tibial artery, n=387 for aorta, and n=213 for coronary artery. Number of points for each genotype are shown in parentheses. Boxplot within the violin plot includes median (white line) and IQR, with the upper (75%) and lower (25%) quartiles shown. P-valules shown are nominal p-values determined from FasQTL based linear regression analysis (by GTEx) by testing genome-wide association of genotype with quantile normalized gene expression levels.



Supplementary Figure 6. (a) Single-cell RNA-seq Seurat v4 based analysis of human coronary artery atherosclerosis dataset (n=4) from Wirka et al. (Nat Med 2019). Cell type labels are based on those provided from the original manuscript. Feature UMAP plots for *PRDM16* and *TBX2* depict predominant expression in SMC and pericytes, with evidence of *PRDM16* expression in endothelial cells (EC). (b) Single-cell RNA-seq Seurat v4 based analysis of mouse atherosclerotic brachiocephalic artery (BCA) and ascending aorta samples from SMC lineage traced *Ldlr* knockout models from Pan et al. Circulation 2020. Feature UMAP plots for *Prdm16* and *Tbx2* depict predominant *Prdm16* and *Tbx2* expression in SMC populations, with evidence of *Prdm16* expression in EC and some macrophages.



Supplementary Figure 7. Power analysis results for cell type identification, differential accessibility, and caQTL. (a) SCOPIT v1.1.4 based power curve showing the probability of detecting at least 50 cells of the rarest cell type in coronary artery based on our snATAC data (Mast cell at a frequency of 0.07), assuming alpha =0.05. 99% probability of detection is achieved from a minimum of 9,129 total cells, as shown by the intersecting orange lines. (b) Two-sample t-test based power curve, showing the probability of detecting differential gene scores in smooth muscle cells (SMC) versus modulated SMC (fibromyocytes) as a function of mean sample sizes (calculated from the required sample sizes for all differential genes; see Methods) (c) Single-cell caQTL power analysis based on PowerEQTL v0.3.4, assuming a standard linear mixed effects model, alpha=0.05, nTests = 250,000 overlapping SNPs, minimum effect size=0.3, # cells = 670 (mean SMC/ sample). A minimum of 40 samples is required to achieve 90% power to detect caQTL variants at ~5% minor allele frequency (MAF).