

1 High-Dimensional Single-Cell Multimodal Landscape of Human Carotid Atherosclerosis

2
3 Alexander C. Bashore^{1†}, Hanying Yan^{2†}, Chenyi Xue¹, Lucie Y. Zhu¹, Eunyoung Kim¹, Thomas
4 Mawson¹, Johana Coronel¹, Allen Chung¹, Sebastian Ho¹, Leila S. Ross¹, Michael Kissner³,
5 Emmanuelle Passegue³, Robert C. Bauer¹, Lars Maegdefessel^{4,5,6}, Mingyao Li^{2†}, Muredach P.
6 Reilly^{1,7†}

7
8 ¹Division of Cardiology, Department of Medicine, Vagelos College of Physicians and Surgeons,
9 Columbia University, New York.

10
11 ²Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania
12 Perelman School of Medicine, Philadelphia, PA

13
14 ³Columbia Stem Cell Initiative, Department of Genetics and Development, Columbia University
15 Irving Medical Center, New York

16
17 ⁴Department of Vascular and Endovascular Surgery, Technical University Munich, Germany

18
19 ⁵German Center for Cardiovascular Research (DZHK), partner site Munich Heart Alliance

20
21 ⁶Karolinska Institute, Department of Medicine

22
23 ⁷Irving Institute for Clinical and Translational Research, Columbia University Irving Medical
24 Center, New York, NY

25
26 †These authors contributed equally

27
28 *Correspondence: mpr2144@cumc.columbia.edu or mingyao@penncmedicine.upenn.edu

30 **Abstract**

31

32 Background:

33

34 Atherosclerotic plaques are complex tissues composed of a heterogeneous mixture of cells.
35 However, we have limited understanding of the comprehensive transcriptional and
36 phenotypical landscape of the cells within these lesions.

37

38 Methods:

39

40 To characterize the landscape of human carotid atherosclerosis in greater detail, we combined
41 cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) and single-cell RNA
42 sequencing (scRNA-seq) to classify all cell types within lesions (n=21; 13 symptomatic) to
43 achieve a comprehensive multimodal understanding of the cellular identities of atherosclerosis
44 and their association with clinical pathophysiology.

45

46 Results:

47

48 We identified 25 distinct cell populations each having a unique multi-omic signature, including
49 macrophages, T cells, NK cells, mast cells, B cells, plasma cells, neutrophils, dendritic cells,
50 endothelial cells, fibroblasts, and smooth muscle cells (SMCs). Within the macrophage
51 populations, we identified 2 proinflammatory subsets that were enriched in IL1B or C1Q
52 expression, 2 distinct TREM2 positive foam cell subsets, one of which also expressed
53 inflammatory genes, as well as subpopulations displaying a proliferative gene expression
54 signature and one expressing SMC-specific genes and upregulation of fibrotic pathways. An in-
55 depth characterization uncovered several subsets of SMCs and fibroblasts, including a SMC-
56 derived foam cell. We localized this foamy SMC to the deep intima of coronary atherosclerotic
57 lesions. Using CITE-seq data, we also developed the first flow cytometry panel, using cell
58 surface proteins CD29, CD142, and CD90, to isolate SMC-derived cells from lesions. Last, we
59 found that the proportion of efferocytotic macrophages, classically activated endothelial cells,
60 contractile and modulated SMC-derived cell types were reduced, and inflammatory SMCs were
61 enriched in plaques of clinically symptomatic vs. asymptomatic patients.

62

63 Conclusions:

64

65 Our multimodal atlas of cell populations within atherosclerosis provides novel insights into the
66 diversity, phenotype, location, isolation, and clinical relevance of the unique cellular
67 composition of human carotid atherosclerosis. This facilitates both the mapping of
68 cardiovascular disease susceptibility loci to specific cell types as well as the identification of
69 novel molecular and cellular therapeutic targets for treatment of the disease.

70 Introduction

71
72 Atherosclerosis, the major underlying cause of cardiovascular diseases (CVD), is a chronic
73 inflammatory disease that is initiated by the infiltration and modification of circulating low-
74 density lipoproteins into the subendothelial space. As the disease initiates and progresses,
75 continuous recruitment of circulating immune cells contributes to a complex inflammatory
76 microenvironment. Additionally, smooth muscle cells (SMCs) migrate from the arterial media
77 into the developing lesion, where they undergo phenotypic modulation and transition to a
78 variety of cell fates including foam cells, macrophage-like cells, synthetic extracellular matrix-
79 producing fibrotic cells, and osteogenic-like cells. Emerging genetic and experimental data
80 suggest that specific SMC-derived cell types are detrimental and contribute to plaque instability
81 (e.g., macrophage and osteogenic types), whereas others are protective and contribute to
82 plaque stability (e.g., fibrotic cell types)¹. Increased proportions of pro-inflammatory cells
83 relative to matrix-producing cells within the lesion appear to increase plaque vulnerability and
84 the likelihood of rupture, leading to clinical manifestations such as myocardial infarction and
85 ischemic stroke².

86
87 Histological studies have identified culprit lesions as having a thin fibrous cap, large necrotic
88 core, and an overabundance of immune cells like macrophages and T cells^{3,4}. Moreover, recent
89 genome-wide association studies (GWAS) have identified many genetic loci associated with
90 CVD⁵, and many of these have been mapped to genes that modulate disease specifically
91 through SMC-derived cell, endothelial cell (EC), or immune cell functions. However, these
92 discoveries have translated into new therapies since the cellular mechanisms driving this
93 complex process have yet to be elucidated fully, cell and molecular mechanistic and causal
94 studies are lacking for most loci, and because mechanistic mouse models may not fully
95 recapitulate the complexity of human atherosclerosis⁶.

96
97 Recently, there has been an effort to characterize all cell types in human atherosclerosis by
98 utilizing rapidly evolving single-cell technologies to identify cell type-specific candidate genes
99 and molecular mechanisms and to drive novel therapies for atherosclerosis. These studies have
100 relied primarily on single-cell RNA sequencing (scRNA-seq), which uncovers cellular
101 heterogeneity by identifying subpopulations of cells with distinct transcriptional profiles in
102 atherosclerotic plaques. In one study, a multi-omic approach identified innate and adaptive
103 immune cell alterations in carotid plaques associated with clinical events⁸. In another important
104 study, a vast heterogeneity of all cell types was revealed using combined scRNA-seq and single-
105 cell ATAC sequencing to profile immune and nonimmune cells in human carotid
106 atherosclerosis⁹. To date, these studies have been relatively small in scale and have not
107 provided a precise reference of cell types and canonical protein markers to facilitate clinical and
108 therapeutic translation. To achieve greater precision, we performed CITE-seq utilizing a large
109 panel of antibodies on carotid atherosclerotic plaques in combination with scRNA-seq. Here we
110 report the most comprehensive phenotypic and transcriptional landscape of all cell types within
111 advanced carotid atherosclerosis and identify cell-type specific protein markers as well as
112 perturbed activation states associated with cardiovascular events in the largest set of human
113 samples to date.

115 **Methods**

116

117 **Human studies**

118 All human subjects research in this study, including the use of human tissues, conformed to the
119 principles outlined in the Declaration of Helsinki. All patient information was de-identified. For
120 single-cell studies, tissues from human carotid atherosclerotic plaques were collected from
121 twenty-one patients undergoing carotid endarterectomy surgery. These human subject studies
122 were performed with approval (protocol number AAAJ2765) of the local Institutional Review
123 Board (IRB) of Columbia University Irving Medical Center, and written informed consent was
124 obtained from all participants. Exclusion criteria include current infection, known immune
125 system disorder, and active or recent (within last three months) radiation, chemotherapy,
126 hormone-based, and/or immunotherapy treatment for cancer. **Table S1** summarizes the
127 demographic and clinical characteristics of the entire cohort, and subgroups to perform
128 comparative analysis. Symptomatic patients were defined as having a stroke or transient
129 ischemic attack (TIA) within 6 months of surgery. Asymptomatic patients had no history of
130 ischemic events. For arterial immunohistochemistry, human coronary artery tissues were
131 obtained from either donor hearts rejected for transplantation or explanted hearts during
132 cardiac transplant surgery. These human subject studies were performed with approval
133 (protocol number AAAR6796) of the local Institutional Review Board (IRB) of Columbia
134 University. Written informed consent was obtained upon admission before surgery.

135

136 **Atherosclerotic plaque cell dissociation for single-cell RNA-seq and CITE-seq preparation**

137 Plaque specimens at endarterectomy were placed in phosphate buffered saline (PBS),
138 immediately transported to the laboratory, washed three times in PBS and minced into small
139 pieces. The tissue was then digested in RPMI 1640 containing Liberase (10 U/mL, Sigma cat no
140 5401127001), Elastase (8 U/mL, Sigma cat no E7885-20MG), and DNase (120 U/mL,
141 Worthington cat no LS006331) at 37°C for 1 hour. The cell suspension was filtered through a 70-
142 µm cell strainer, washed with double the volume of RPMI 1640 with 10% FBS, centrifuged at
143 400g for 5 min at 4°C, washed once with sterile-filtered FACS buffer (PBS + 2% FBS + 5mM EDTA
144 pH 8.0 + 20mM HEPES + 1mM sodium pyruvate), and spun down at 400g for 5 min at 4°C. For
145 CITE-seq, plaque specimens were dissociated the same way as for scRNA-seq and then
146 resuspended in 49µL of FACS buffer and 1µL of TruStain FcX (BioLegend cat no 422302) and
147 blocked for 10 minutes at room temperature. Following incubation, lyophilized oligo-
148 conjugated antibodies (**Table S2**) were reconstituted in 50µL of FACS buffer and were added to
149 cell suspension and incubated for 30 minutes at 4°C. Then samples were washed 3 times in
150 FACS buffer with centrifugation at 400g for 5 min at 4°C. Single cell suspensions were stained
151 with Calcein AM (Thermo Fisher Scientific C34852) and Propidium Iodide (BioLegend 421301),
152 filtered through a 40-µm cell strainer and immediately brought to the flow cytometry core for
153 sorting. Viable cells were sorted on a BD FACSAria II and selected as being positive for Calcein
154 AM (488 nm excitation and 530/30 bandpass filter) and negative for Propidium Iodide (561 nm
155 excitation and a 610/20 bandpass filter). Doublets were excluded using forward light scatter
156 (FSC) signal area and height as customary. The FACSAria was set up using the 100 µm nozzle at
157 20 PSI pressure, and cells were sorted using a one-drop purity (4-way Purity) mode. Cells were

158 collected in 1.5mL Eppendorf tubes containing DMEM/f12 + 10% FBS before pelleting and
159 counting for 10x Genomics Chromium input.

160

161 **CITE-seq library preparation**

162 CITE-seq libraries were prepared as previously described^{7,10} following 10x Genomics 3' v3
163 protocol according to manufacturer's instructions for cDNA amplification using 0.2 μM of ADT
164 additive primer (5'CCTTGGCACCCGAGAATTCC). The supernatant from the 0.6x SPRI cleanup
165 was saved and purified with two rounds of 2x SPRI, and final product was used as a template to
166 produce ADT libraries. Antibody tag libraries were generated by PCR using Kapa Hifi Master Mix
167 (Kapa Biosciences KK2601), 10 mM 10x Genomics SI-PCR primer
168 (5'AATGATACGGCGACCACCGAGATCTACTCTTCCCTACACGACGCTC), and Small RNA RPIx
169 primer (5'CAAGCAGAAGACGGCATAACGAGATxxxxxxGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA
170 with xxxxxx denoting one of the four following sequences: CGTGAT, ACATCG, GCCTAA,
171 TGGTCA). Following amplification, Antibody tag libraries were cleaned up with 1.6x SPRI.
172 Subsequently, ADT quality was verified using a DNA high sensitivity assay on an Agilent 2100
173 bioanalyzer.

174

175 **scRNA-seq and CITE-seq data pre-processing**

176 scRNA-seq and CITE-seq FASTQ files were pre-processed by 10x Genomics Cell Ranger 5.0.1 to
177 generate count matrices with quality controls (QCs). Reference version 2020-A with GENCODE
178 v32 annotation was used. For CITE-seq libraries, an additional feature barcode reference was
179 included that contained IDs and sequences of 283 TotalSeq-A antibody-derived tags. Suffix was
180 added to gene symbols when multiple genes with different gene IDs shared the same gene
181 symbol.

182

183 **Quality control (QC), normalization, and highly variable gene (HVG) selection**

184 Using the Scanpy package in Python¹¹, we first removed any cells expressing less than 200
185 genes, and then any genes expressed in less than 3 of the remaining cells with functions
186 "pp.filter_cells" and "pp.filter_genes". We also annotated the group of mitochondrial genes
187 and computed metrics with function "pp.calculate_qc_metrics". Cells with total UMI counts
188 over 40,000, or total mitochondrial genes proportion over 30%, or total unique genes over
189 6,000 were filtered out. Cells that pass QC filtering have gene expression normalized using a
190 two-step procedure: (1) Cell level normalization, in which the UMI count for a given gene in
191 each cell is divided by the total UMI count across all genes in the cell, multiplied by the median
192 total UMI count across all cells, and transformed to a natural log scale. (2) Gene level
193 normalization, in which the cell level normalized values for each gene are standardized by
194 subtracting the mean and dividing by the standard deviation across all cells within the same
195 batch for the given gene. HVGs were selected based on the log-normalized UMI counts using
196 the approach introduced by Stuart and Butler¹² and implemented in the
197 "pp.highly_variable_genes" function with "batch_key" parameter in the Scanpy package. The
198 selected HVGs were used for downstream clustering analysis.

199

200 **Clustering and differential gene and protein expression analysis**

201 CarDEC¹³, a published deep learning tool for scRNA-seq data batch effect correction, gene
202 expression denoising, and clustering of cells, was applied on 15 scRNA-seq samples to get
203 denoised gene expression on the z-score scale. CarDEC_CITE, a modified version of CarDEC for
204 CITE-seq data which preprocesses the gene and protein expression data separately, was utilized
205 on 6 CITE-seq samples, generating denoised feature expression on the z-score scale and
206 identifying 25 clusters.

207
208 Marker genes and proteins were detected using “tl.rank_genes_groups” in the Scanpy package
209 for each cluster identified, and the cell types were assigned based on the top features. Major
210 cell lineages were annotated using knowledge of known protein and gene markers for each cell
211 type from previously published flow cytometry and scRNA-seq studies^{14,15}.

212 213 **Integrative analysis of scRNA-seq and CITE-seq data**

214 sciPENN is our published deep learning tool for integrative analysis of scRNA-seq and CITE-seq.
215 Leveraging information learned from CITE-seq data, sciPENN can predict protein expression in
216 scRNA-seq and transfer clustering labels from CITE-seq into scRNA-seq data¹⁶. Using sciPENN,
217 we integrated scRNA-seq and CITE-seq data into a shared latent low-dimensional space, which
218 was further dimension reduced and visualized using functions “sc.tl.pca”, “sc.pp.neighbors” and
219 “sc.tl.umap” within Scanpy package.

220 221 **Sub-clustering of macrophages, SMCs, and Fibroblasts**

222 Among the 25 clusters obtained from the CITE-seq data, cells labeled as Myeloid (clusters 0, 3,
223 9, 11, 13, 19, 22), SMCs and Fibroblasts (clusters 2, 5, 12, 17, 18) were extracted, respectively.
224 CarDEC was then performed to further cluster and identify the subpopulations.
225 “tl.rank_genes_groups” was used to get marker genes and cell types were assigned based on
226 them.

227 228 **Differential gene expression and proportional analysis between symptomatic and 229 asymptomatic plaques**

230 Among 21 samples, 13 are from symptomatic patients, and 8 are from asymptomatic patients.
231 For each cluster, pseudo-bulk RNA-seq analysis on the single-cell data was performed among
232 the symptomatic and asymptomatic groups using DESeq2¹⁷ package in R. Raw counts were
233 summed per gene per sample for cells in each cluster, pre-filtering was performed to keep
234 genes with at least 10 reads total, standard differential expression analysis was performed
235 using function “DESeq” and results were generated with function “results”. For each cluster, we
236 also calculated the total cell counts proportion per sample, and unpaired two sample t test was
237 performed using “ttest_ind” within scipy.stats package in Python to compare symptomatic and
238 asymptomatic groups.

239 240 **Pathway analysis**

241 Cell-type specific differentially expressed (DE) genes, with a Benjamini-Hochberg adjusted p-
242 value of less than 0.01 and log₂ fold change greater than 1, were uploaded into Ingenuity
243 Pathway Analysis (IPA) software (QIAGEN). IPA analysis reported the -log₁₀ p values and Z-

244 scores of canonical pathways and upstream regulators. The highest predicted upregulated
245 canonical pathways and upstream regulators with a $-\log_{10}$ p value > 1.3 for each cell population
246 were selected for the respective analysis and visualized with either a heat map or radar plot.
247 Full list of pathways can be found in **Table S4**.

248

249 **Gene scoring analysis**

250 To better characterize genes and pathways that are activated or repressed in cell populations,
251 we performed gene scoring analysis in human carotid plaque scRNA-seq data. We first compiled
252 lists of genes involved in multiple cellular functions and pathways, including apoptosis, DNA
253 damage, macrophage efferocytosis, senescence, UPR response, and inflammasome. Cell cycle
254 genes, including 43 S phase genes and 54 G2/M phase genes from Tirosh et al.¹⁸, were
255 extracted from Seurat package¹⁰, where gene symbols were updated. UMI counts were
256 normalized to a target total count of 10,000 per cell and then log-transformed. Gene scoring
257 was performed using “scanpy.tl.score_genes” function from Scanpy 1.9.1¹¹ with parameters
258 $n_bins = 24$ and $ctrl_size = 100$. Cell cycle scoring was performed using
259 “scanpy.tl.score_genes_cell_cycle” function supplied with S phase genes and G2/M phase
260 genes above.

261

262 Gene scores were compared between cells from asymptomatic subjects and symptomatic
263 subjects in each of the SMC, modulated SMC, fibroblast, macrophages, and endothelial cell
264 populations. Wilcoxon rank-sum test was performed on each score between asymptomatic and
265 symptomatic subjects in each cluster. P-values were adjusted by Benjamini-Hochberg method.
266 Scores with adjusted P-value < 0.05 were considered significantly different between
267 asymptomatic and symptomatic subjects.

268

269 To compare gene score between multiple populations, we performed Kruskal-Wallis test across
270 macrophage populations, SMCs and fibroblasts subpopulations. P-values were adjusted by
271 Benjamini-Hochberg method in each set of analyses. For gene scores with adjusted P-value
272 < 0.05 , Dunn’s test was performed to identify pairs of populations that had significant
273 differences in gene score distribution. P-values from Dunn’s test were also adjusted by the
274 Benjamini-Hochberg method.

275

276 **M1 and M2 macrophage signature analysis**

277 To characterize M1- and M2-macrophage functions in macrophage subpopulations, we utilized
278 lists of M1 and M2 signature genes from Martinez et al.¹⁹. Genes from membrane receptors,
279 cytokines and chemokines, apoptosis-related genes, solute carriers, enzymes, extracellular
280 mediators, and DNA-binding factors were ordered by M1:M2 expression ratio. 54 genes with
281 higher M1-macrophage expression and 43 genes with higher M2-macrophage expression were
282 used to score macrophage subpopulations. Gene score comparison between populations was
283 performed as described above.

284

285 **Flow cytometry**

286 Flow cytometry was performed on a frozen cell suspension from a plaque that was previously
287 submitted for scRNA-seq. Following thawing, cells were washed with FACS buffer and
288 centrifuged at 400g for 5 mins. The plaque cells were then blocked with TruStain FcX
289 (BioLegend cat no 101319) for 10 mins at room temperature. Subsequently, BD Brilliant Stain
290 Buffer (BD Biosciences 563794) and the following fluorescently conjugated antibodies were
291 added to the cells for 20 minutes at 4°C: CD45-APC/Cy7 (BioLegend 368515), CD142-APC
292 (BioLegend 365205), CD31-BV421 (BioLegend 303123), CD90-BUV496 (BD Biosciences 741160),
293 and CD29-PerCP-eFluor 710 (Thermo Fisher Scientific 46-0299-41). Cells were then washed 3
294 times with FACS buffer and viability dyes Calcein AM and Propidium Iodide were added. The
295 fluorescently labeled plaque cells were measured on a Novocyte Penton and analyzed with
296 FlowJo software.

297

298 **Immunohistochemistry of human coronary arteries**

299 The left anterior descending coronary artery was excised from the myocardium and
300 immediately placed into ice-cold RPMI, washed with RPMI to remove any blood, and cut into
301 roughly 5mm sections and embedded in OCT. Blocks were immediately snap-frozen and stored
302 at -80°C. For sectioning, tissue blocks were allowed to equilibrate at -20°C, sectioned into 8µm
303 thick sections, and mounted onto Superfrost Plus Microscope Slides.

304

305 For immunofluorescence, the sections were fixed with 4% paraformaldehyde for 20 minutes at
306 room temperature. After fixing, the sections were blocked in PBS containing 1% BSA and 10%
307 goat serum for 1 hour. The sections were then subsequently incubated with the following
308 primary antibodies, CD90 (1:200, Abcam ab181469) and Smoothelin (1:1000, Abcam
309 ab219652), overnight at 4°C in a humidified chamber. The next day, the sections were washed
310 with PBS and incubated with the appropriate secondary antibodies and LipidTOX Red (1:200,
311 Thermo Scientific H34476) for 1 hour at room temperature. Finally, the sections were
312 counterstained with DAPI and mounted with Invitrogen Prolong Glass Antifade Mountant
313 (Invitrogen P36982). Images were acquired with a Nikon A1 confocal microscope.

314 **Results**

315

316 **High-dimensional multi-omic profiling identifies 25 distinct cell populations in human**
317 **atherosclerotic plaques**

318

319 To examine the single-cell transcriptome and cellular immunophenotypes of human
320 atherosclerotic plaques, we performed multi-omic profiling on carotid endarterectomy tissue
321 from 21 patients. Tissue samples were digested enzymatically, and cells from six of these
322 patients were labeled with a panel of 274 CITE-seq antibodies before sequencing. Data were
323 integrated from 88,093 cells across all patients (CITE-seq and scRNA-seq) using our deep
324 learning model sciPENN, which also predicted protein expression in the query scRNA-seq
325 dataset from the CITE-seq reference dataset¹⁶ (**Fig 1A**). All cells were visualized by uniform
326 manifold approximation and projection (UMAP), revealing 25 distinct cell populations (**Fig 1B**).
327 Utilizing well-established canonical proteins, we identified 2 EC populations and 18 leukocyte
328 populations (**Fig 1C** and **Suppl Fig 1**), including all major immune cell types such as
329 macrophages, T cells, NK cells, B cells, plasma cells, mast cells, dendritic cells, and neutrophils.
330 Within macrophages, we observed 5 subtypes (Macrophage 1-5) that had high, near-ubiquitous
331 expression of major lineage markers such as CD64, CD11c, CD14, and MHCII. Three populations
332 of CD4⁺ T cells (CD4⁺ T Cell 1-3) and 2 populations of CD8⁺ T cells (CD8⁺ T Cell 1 and 2),
333 determined by CD3, CD7, CD2, and CD194 expression, were identified. Cluster 15, known to be
334 phenotypically like T cells, was distinguished by the specific expression of the NK cell markers
335 CD56 and CD16. B Cell 1, B Cell 2, and plasma cells were identified by the expression of CD19
336 and CD20, and KIT⁺ mast cells were identified by expression of CD117. Two small dendritic cell
337 populations (clusters 19 and 22) that comprised plasmacytoid (pDC) and conventional DCs (cDC)
338 characterized by the expression of CD303 and CD141, respectively, were also identified. We
339 classified cluster 24 as neutrophils, which can be difficult to observe in scRNA-seq data, by the
340 expression of CD15. Similarly, we identified two EC clusters as having specific expression of the
341 proteins CD31, CD34, CD144, and CD49b.

342

343 Because SMCs and fibroblasts have not been immunophenotyped extensively, we used
344 canonical gene markers to identify these cell types (**Fig 1D** and **Suppl Fig 2**). To differentiate
345 SMCs and fibroblasts that clustered together by phenotypic similarities, we examined the
346 expression of known SMC genes ACTA2, MYH11, and CNN1, as well as fibroblast-specific genes
347 COL1A2, LUM, and LOXL1. This process revealed that clusters 5 and 12 (SMC 1 and 2) were
348 contractile SMCs with the highest expression of all SMC-related genes, whereas clusters 17
349 (SMC 3) and 18 (Modulated SMC) had lower expression of SMC genes. Modulated SMC had
350 higher expression of fibroblast genes relative to SMC 1 and 2, indicating that this SMC-derived
351 population has undergone phenotypic modulation. Cluster 2, with high expression of COL1A2,
352 LUM, and LOXL1, encompassed all fibroblasts in our dataset.

353

354 Next, we calculated the proportion of each cluster for each of the 21 individual subjects (**Suppl**
355 **Figure 3A and B**). On aggregate, macrophages comprised 32% of the cells in our dataset, T cells
356 25%, SMCs 14%, fibroblasts 13%, and ECs 8% (**Fig 1E**). In contrast, other scRNA-seq studies have
357 identified T cells as the dominant cell type in human carotid plaques^{8,9}, perhaps due to

358 differences in digestion and isolation protocols or patient heterogeneity. Leukocytes comprised
359 the largest proportions of cells within our carotid atherosclerosis dataset, consisting mainly of
360 macrophages and T cells.

361

362 **Multimodal analysis provides deep understanding of plaque cellular phenotypes**

363

364 To extend our findings beyond known cell surface protein markers, we used our CITE-seq data
365 to identify the distinct gene and protein signatures for all 25 populations depicted in **Fig. 1B**. To
366 assess the added value of incorporating both RNA and protein information into our analysis, we
367 calculated the gene and protein correlation for every gene-protein pair in our dataset for each
368 individual cluster. As shown in **Fig 2A**, overall, there was a high degree of correspondence
369 between protein and gene expression (**Table S3**), although it is well-established that there can
370 be variability between protein levels and their coding transcripts²⁰. In several circumstances,
371 the protein expression contributed critically to the identification of cell types when the RNA
372 was not highly expressed (**Suppl Fig 4**). For example, established neutrophil protein markers
373 CD15 and CD16 were highly expressed and specific although their corresponding genes, *FUT4*
374 and *FCGR3A*, were not detectable.

375

376 Our five macrophage clusters segregated into three major phenotypic classes. Macrophages 1
377 and 5 had high expression of FOLR2, cadherin 11 and CD93, suggesting roles in inflammation,
378 phagocytosis, and cell adhesion. In contrast, Macrophages 3 and 4 had high expression of CD36
379 and TREM1, suggesting a role in lipid uptake and metabolism. However, Macrophage 3 uniquely
380 expressed CD35, a receptor for complement activation. Macrophages 2 and 3 shared
381 expression of CLEC12A and CD192 (CCR2), indicating that they are more inflammatory and likely
382 recently recruited from circulation (**Suppl Fig 5A**). Macrophage 4 had the highest expression of
383 genes involved in lipid metabolism such as *ABCA1*, *LPL*, *FABP5*, and *APOE*, whereas
384 macrophages 2 and 3 had the highest expression of inflammatory genes such as *IL1B*, *NLRP3*,
385 and *CCR2* (**Suppl Fig 5B**). Macrophages 1 and 5, notably, had intermediate expression of a
386 substantial proportion of genes involved in both functions. Because there was considerable
387 overlap in the gene expression profile of all macrophage subtypes, we identified which
388 upregulated and downregulated genes were unique to each population (**Suppl Fig 5C**). Using
389 this gene set, we performed pathway analysis to identify in more detail the predicted functional
390 differences between these populations (**Suppl Fig 5D**). This suggested both substantial
391 functional heterogeneity as well as degrees of overlap of plaque macrophages.

392

393 CITE-seq profiling was particularly helpful for characterizing SMC, fibroblasts, and ECs using
394 novel protein phenotypic markers. The SMC populations (SMC 1-3) had relatively uniform
395 expression of the proteins CD29, CD142, and EGFR. The modulated SMC population also had
396 relatively uniform expression of CD29 and EGFR but lacked expression of CD142. In contrast,
397 fibroblasts expressed CD90 to a similar degree as the modulated SMC populations, but
398 specifically expressed PDPN and CD13 (**Fig 2B**). Using the combination of these markers, we
399 developed a novel flow cytometry gating strategy to discriminate between SMC 1-3, modulated
400 SMCs, and fibroblasts (**Fig 2C**).

401

402 Our clustering also revealed two distinct EC populations (**Fig 1B and C**). To differentiate these,
403 we identified proteins that were expressed uniquely on one but not the other, e.g., CD49f,
404 CD112, and CD200 were expressed on endothelial cell 1 while CD62P, CD61, and CD201 were
405 highly expressed on endothelial cell 2 (**Fig 2D**). To delineate these cell types further, we
406 performed a differential expression analysis followed by pathway analysis using inferred
407 functional differences (**Fig 2E**). This analysis revealed that the endothelial cell 1 population
408 represented activated and inflamed endothelium, suggesting that this cluster promotes
409 leukocyte extravasation. In contrast, the endothelial cell 2 population had upregulation of
410 fibrotic pathways, suggesting that this cluster may be predisposed to undergo endothelial to
411 mesenchymal transition.

412

413 **Cerebrovascular events are associated with dysregulation of macrophages as well as specific** 414 **smooth muscle cell, and endothelial cell subpopulations**

415

416 UMAP visualization of samples revealed that the majority of clusters overlapped between
417 asymptomatic and symptomatic patient groups (**Fig 3A, 3B**). However, some striking differences
418 were identified, including reductions in the proportions of Macrophage 1, Endothelial Cell 1,
419 SMC 2, and modulated SMC populations, and a modest increase in SMC 3 population in
420 symptomatic patients (**Fig 3C**). The reduction in Macrophage 1, a macrophage population with
421 efferocytotic functions (**Suppl Fig 6A**), suggests that the capacity to clear dead and dying cells is
422 compromised. The decrease in the activated Endothelial Cell 1 population may suggest
423 increased plaque erosion. The decrease in modulated SMCs and SMC 3 subpopulation increase
424 suggests a shift to pro-inflammatory SMCs (**Suppl Fig 6B**). Unlike other studies⁸, we did not see
425 alterations in the proportion of T cell subpopulations (**Suppl Fig 6C**). Contrary to findings in
426 previous publications⁸, differential expression analysis revealed only three significant DE genes
427 associated with clinical status: *S100A8*, *DDT*, and *PTGS1* (**Suppl Fig 6D**).

428

429 To assess biological processes that are associated with clinical events in more depth, we
430 performed a gene scoring analysis for macrophages, SMC, and fibroblasts. Generally,
431 macrophages in symptomatic plaques were more senescent (**Fig 3D**), exhibited a higher level of
432 ER stress through PERK activation, had a higher degree of inflammasome activation, but also
433 were pro-resolving (**Suppl Fig 6E**). Like macrophages, SMCs and fibroblasts from symptomatic
434 plaques were more senescent and PERK was more activated than were those from
435 asymptomatic plaques (**Suppl Fig 6F**). Also, SMCs and fibroblasts were more glycolytic
436 suggesting a metabolic shift in plaques associated with clinical events, as has been noted in
437 plaques with features of plaque instability in mice²¹ (**Fig 3E**).

438

439 **Deep sub-clustering reveals diverse macrophages phenotypes in carotid plaques**

440

441 Initial clustering identified 5 plaque macrophage clusters (Macrophage 1-5) (**Figs. 1 and 2**). To
442 explore cellular diversity more deeply, we performed a sub-clustering analysis on macrophage
443 and DC populations. This revealed 10 cell clusters, 7 macrophage subpopulations, 2 DC clusters,
444 and residual T cells expressing CD3 (**Fig 4A and Suppl Fig 7A and B**). From the initial clustering
445 (**Suppl Fig 7C**), Macrophage 1 mapped to clusters 0 and 4, Macrophage 2 mapped to cluster 1,

446 Macrophage 3 mapped to cluster 3, and Macrophage 4 mapped to clusters 2 and 5. In contrast,
447 Macrophage 5 was distributed throughout the subclusters. Cluster 7 was comprised of cells for
448 every macrophage population in the initial analysis. Clusters 0 and 1 were highly inflammatory,
449 expressing genes *C1Q* and *IL1B*, respectively, and enriched in inflammatory pathways (**Fig 4B**).
450 Clusters 2 and 4 expressed foam cell marker genes, in which cluster 2 expressed *ABCA1*, *LPL*,
451 *CD36* and *TREM2* most prominently, whereas cluster 4 expressed *APOE* in addition to
452 inflammatory genes like *CCL18* and *C1Q* genes. Additionally, apoptotic (cluster 3), proliferative
453 (cluster 5), and *ACTA2*⁺ (cluster 7) subpopulations were identified. Cluster 3 showed
454 upregulation of the granzyme A apoptotic pathway (**Fig 4B**), along with high expression of
455 mitochondrial genes (**Suppl Fig 7A and B**). The proliferative subpopulation had specific
456 expression of many proliferation markers including *MKI67*, *TUBB*, and *STMN1* (**Suppl Fig 7B**).
457 Finally, the *ACTA2*⁺ subpopulation (~6% of all macrophages) was, notably, characterized by high
458 expression of several SMC-related genes such as *MYOCD*, *ACTA2*, *MYH11*, and *CNN1* (**Suppl Fig**
459 **7B**), along with protein expression of CD64, CD11c, and MHCII proteins (**Suppl Fig 7D**),
460 suggesting that it could represent phenotypic switching of SMCs to macrophage-like cells. This
461 subpopulation also exhibited upregulation of fibrotic signaling and epithelial-mesenchymal
462 transition pathways (**Fig 4B**).
463

464 To establish regulatory features of each macrophage subpopulation, we conducted an
465 upstream analysis using top DE genes and identified potential transcriptional regulators (**Fig**
466 **4C**). The *C1Q*^{hi} (cluster 0) and *IL1B*^{hi} (cluster 1) subpopulations shared regulatory features, like
467 *STAT1* and *RELA*, involved in immune modulation^{22,23}. The proliferative subpopulation (cluster
468 5) exhibited the most specific regulatory features with high inferred regulation by the oncogene
469 *MYC*. The *ACTA2*⁺ subpopulation showed inferred transcriptional regulation by *MRTFB*, *MRTFA*,
470 *NOTCH3*, and *SMAD3*, known to control SMC differentiation²⁴⁻²⁷. To infer M1/M2 canonical
471 functional properties, we tested enrichment of M1- and M2-associated genes¹⁹ in our dataset
472 (**Fig 4D**). Only *IL1B*^{hi} macrophages showed enrichment in M1 genes. Conversely, the *C1Q*^{hi},
473 foamy 1, foamy 2, and apoptotic subpopulations exhibited strong expression of M2 genes. The
474 proliferative and *ACTA2*⁺ subpopulations did not show a strong M1 or M2 signature. These data
475 align with literature suggesting that the M1/M2 classification inadequately captures the *in vivo*
476 phenotypic diversity of macrophages in atherosclerotic lesions²⁸.
477

478 Although we did not observe any differences in macrophage subpopulations between
479 asymptomatic and symptomatic plaques (**Suppl Fig 7E**), we evaluated which cell populations
480 had upregulated gene expression for processes known to impact atherosclerosis. To do so, we
481 utilized a list of genes involved in inflammasome activation, DNA damage and efferocytosis to
482 perform gene score analyses for these processes (**Fig 4E**). As expected, the *IL1B*^{hi} subpopulation
483 was the only one with marked inflammasome activation. The proliferative subpopulation had
484 highest expression of genes involved in DNA damage, while each of the *C1Q*^{hi}, foamy 1, and
485 foamy 2 subpopulations had evidence for enhanced efferocytotic function.
486

487 **Deep sub-clustering reveals multiple SMC and fibroblast subclasses in plaques**

488

489 In initial clustering we identified 3 SMCs, 1 modulated SMC, and 1 fibroblast population (**Figs. 1**
490 **and 2**). This high-level clustering may not reveal the true heterogeneity within these cell types.
491 A recent study that performed snRNA-seq of human aortic tissue identified 7 SMC and 4
492 fibroblast populations²⁹. Therefore, we performed a deep sub-clustering analysis on SMCs and
493 fibroblasts that revealed 10 unique clusters with specific gene expression profiles (**Fig 5A**).
494 Using classical SMC and fibroblast phenotypic markers (**Fig 5B**), we found that clusters 0 and 2
495 (major SMC 1 and 2) represented the two major classes of contractile SMCs with the highest
496 expression of MYH11 and ACTA2. Two smaller contractile SMC populations, clusters 4 and 6
497 (minor SMC 1 and 2), were also identified. Minor SMC 1 had moderate expression of MYH11
498 and TAGLN, but very low expression of other contractile SMC genes. Minor SMC 2 had
499 moderate expression of *MYOCD*, *MYH11*, and *SMTN*, but low levels of other SMC genes. The
500 most striking population was cluster 5, which had low expression of SMC genes and the specific
501 expression of many genes involved in lipid metabolism, suggesting that this large population
502 may be SMC-derived foam cells (**Suppl Fig 9A**). This foamy SMC cluster mapped almost
503 exclusively back to the modulated SMC population in the original analysis (**Suppl Fig 9B**). We
504 identified 3 fibroblast populations, clusters 1, 3, and 8, determined by the expression of *BGN*,
505 *DCN*, and *COL1A1*. Cluster 9, the last cluster identified, represented a fibromyocyte, having the
506 highest expression of *VCAM1* and *LY6E* (**Suppl Fig 9C**).
507

508 To identify functional differences in the SMC subpopulations, we performed pathway analyses
509 of SMC and fibroblast sub-clusters. Consistent with annotation as prototypical contractile SMC
510 populations by DE gene analysis, our pathway analyses (**Fig 5C**) revealed that major SMC 1 and
511 2 were enriched for muscle functions, TGF- β signaling, and actin cytoskeleton signaling.
512 Coincident with downregulation of contractile SMC genes, minor SMC 1 had upregulated
513 pathways involved in mRNA degradation and cell cycle regulation, suggesting that it might be
514 an actively dividing and proliferating subpopulation. Fibroblast clusters had upregulation of
515 several fibrotic pathways, and we found upregulation of necroptotic and apoptotic pathways in
516 the fibromyocytes (**Fig 5D**).
517

518 Since our data and those from recent mouse and human studies suggest that many foam cells
519 may be SMC-derived^{30,31}, we used a combination of markers revealed by CITE-seq and
520 functional inference to design an immunostaining protocol to spatially locate foamy SMCs in
521 human coronary artery sections. As noted, the foamy SMC cluster mapped primarily onto the
522 modulated SMC cluster in our initial analysis (**Suppl Fig 9B**), expressed the SMC-specific gene
523 smoothelin (**Fig 5B**), and highly expressed the protein CD90 (**Suppl Fig 9D**). At
524 immunohistochemistry of human coronary artery sections with a combination of smoothelin,
525 CD90, and LipidTOX (to identify foam cells) (**Fig 5E**), we revealed a sub-cluster of neointimal
526 cells that stained for all three markers, validating, and localizing modulated SMC-derived foamy
527 cells in human atherosclerotic lesions.
528

529 Given differences in SMC 2, SMC 3, and modulated SMCs by clinical status in our initial
530 clustering (**Fig 3C**), we examined the cell distributions by clinical events in our deep cell type
531 clustering and found subtle decreases in the proportions of foamy SMCs and fibromyocytes,

532 with a significant increase in minor SMC 2 and increased PERK activation in symptomatic
533 plaques (**Fig 5F and Suppl Fig 9E**).

534 **Discussion**

535

536 Through the largest and most in-depth scRNA-seq and high-dimensional CITE-seq phenotypic
537 characterization of human carotid atherosclerosis to date, we provide novel insights into the
538 cellular and molecular pathogenesis of atherosclerotic CVD. Our study reveals a comprehensive
539 landscape of 25 cell types, including macrophages and their regulatory features,
540 immunophenotypic and functional characterization of SMCs and fibroblasts, and cell type-
541 specific perturbations associated with cerebrovascular events. We identified 7 macrophage, 3
542 fibroblast, and 7 SMC subpopulations, including 3 phenotypically modulated or ‘switched’ SMC
543 subpopulations. Each cell population displayed a distinct gene and protein expression signature,
544 reflecting their unique functions. For the first time, we also offer marker panels for immune-
545 isolation and localization of key modulated SMC populations to facilitate further functional and
546 translational studies.

547

548 Our work expands on recent advancements in flow cytometry and CyTOF that have improved
549 our understanding of atherosclerosis by revealing the diversity of cell types in this disease^{8,32,33}
550 and also contributes beyond those made to date by single-cell genomic technologies, as
551 exemplified in the work by Fernandez et al.⁸ and Depuydt et al.⁹. These studies have enhanced
552 our knowledge of cellular heterogeneity and regulatory networks in atherosclerosis^{8,9,26,34-43}.
553 Fernandez et al.⁸ used CyTOF, scRNA-seq, and selective CITE-seq to investigate immune cell
554 diversity in human carotid atherosclerosis, but their study lacked nonimmune cells and only
555 single-cell profiled 6 patients. Depuydt et al.⁹ expanded the scRNA-seq analysis to 18 patients,
556 including immune and nonimmune cells, and performed scATAC-seq to identify transcriptional
557 regulators. However, both studies lacked detailed phenotypic characterization of cell types
558 within lesions, which is critical for understanding their roles in the disease.

559

560 Despite extensive evidence in mouse and human atherosclerosis that SMC, SMC-derived cells,
561 and fibroblasts comprise the majority of cells within atherosclerotic lesions⁵³, single-cell
562 profiling studies in humans have focused primarily on leukocytes. To address this gap, we
563 performed a comprehensive characterization of SMCs and fibroblasts in human lesions. Our
564 analysis identified unique surface proteins, including CD29 and EGFR, that are expressed
565 universally across these cell types, while CD90 was specific to modulated SMCs and fibroblasts,
566 and CD142 was not expressed on the modulated SMC population. These findings enable
567 isolation of these cell types by flow cytometry and determination of the spatial location of each
568 cell type within atherosclerotic tissue, thereby providing critical insight into their functions and
569 potential cell-cell interactions. Sub-clustering analysis revealed previously unknown
570 heterogeneity, including two major contractile SMC and two modulated SMC subpopulations
571 (foamy and fibromyocyte). Mouse models suggest that fibromyocytes contribute to plaque
572 stabilization³⁵, while the role of SMC-derived foam cells remains unclear. Our preliminary
573 evidence suggests that SMC-derived foam cells are localized to the neointima of advanced
574 coronary lesions, not the fibrous cap, shedding light on their possible functions and lesion
575 domain interactions in atherosclerotic lesion subdomains.

576

577 Our extensive plaque characterization suggests that macrophages are the predominant
578 leukocyte population, contrary to findings from Fernandez et al. and Depuydt et al., who
579 identified T cells as the most abundant immune cell type in lesions. Previous scRNA-seq studies
580 of mouse and human lesions indicated a higher prevalence of myeloid cells than T cells^{35,38,39},
581 highlighting the need for standardization of sample processing, analytical methods, and
582 detailed cross-species comparisons. Using protein and gene expression, we identified distinct
583 phenotypic, functional, and activation states of plaque macrophages. Sub-clustering analysis
584 confirmed two proinflammatory macrophage subsets (*IL1B*^{hi} and *C1Q*^{hi}) with enriched
585 inflammatory pathways, consistent with other single-cell studies in different disease contexts⁴⁴.
586 The *IL1B*^{hi} subset exhibited the most distinct inflammatory gene and protein expression profile,
587 including the genes *IL1B*, *NLRP3*, *CCR2*, and the protein CD192. This suggests that *IL1B* targeting
588 with canakinumab, which reduced CVD events in the CANTOS trial⁴⁵, may act preferentially on
589 these macrophages. Of further translational relevance, our recent work in mice with
590 collaborators suggests that inhibiting *IL1B* in individuals with clonal hematopoiesis (CH) could
591 effectively reduce CH-driven CVD risk by targeting inflammatory macrophages⁴⁶.

592
593 Our macrophage findings support, in part, a recent transformative study in mice that identified
594 nonfoamy macrophages as having a distinct proinflammatory transcriptome and suggested that
595 lipid accumulation in macrophages may not be a driver of plaque inflammation and
596 vulnerability³⁶. Yet our data suggest complex functions for macrophage foam cell sub-types. We
597 discovered two distinct foamy macrophage types in lesions. Foamy 1 had a classic foam cell
598 phenotype with enriched lipid metabolism genes (*ABCA1*, *LPL*, *FABP5*) and high CD36 protein
599 expression. This population predominantly expressed *TREM2*, which is proposed to protect
600 against atherosclerosis by facilitating cholesterol metabolism through inducing expression of
601 *APOE*, *LPL*, and *LIPA*, while simultaneously reducing inflammation⁴⁷. However, in potential
602 conflict with this model, a recent preprint revealed that macrophage-specific deletion of *Trem2*
603 in mice reduced atherosclerosis progression and plaque burden⁴⁸ suggesting that *TREM2* may
604 be atherogenic at least under certain circumstances. These studies highlight the critical need for
605 additional research in this area e.g., to determine if *Trem2* agonism or antagonism and under
606 what circumstances is a viable approach for preventing atherosclerosis. In this context, our
607 Foamy 2 macrophage population expressed *TREM2* and certain foam cell genes (*APOE*) as well
608 as *C1Q* genes, indicating an intermediate state between Foamy 1 and inflammatory *C1Q*^{hi}
609 populations. In fact, we identified more than one cluster in our initial analysis (Macrophage 1
610 and 5) that had intermediate expression of genes involved in both inflammation and lipid
611 metabolism. Further research is needed to understand the functions of these foamy-
612 inflammatory macrophage types and their impact on disease progression and clinical events.

613
614 We found a macrophage subpopulation expressing proliferative genes, like the stem-like
615 macrophages identified in murine atherosclerosis through fate-mapping⁴⁹. Macrophage
616 proliferation is prominent in advanced stage atherosclerosis in mice⁵⁰ and can also occur
617 distinctly during lipid-induced regression and plaque stabilization⁵¹. Thus, this proliferative
618 population may be responsible for local macrophage proliferation in advanced human
619 atherosclerosis or may act as a transitional cell state giving rise to various phenotypes, including
620 macrophages involved in plaque regression and stabilization. These findings highlight the

621 diversity and complexity of macrophage subtypes within plaques⁵², emphasizing opportunities
622 but significant uncertainties in targeting specific macrophage populations with precision
623 therapeutics.

624
625 Our study supports the concept of marked cell type transitions in human atherosclerosis. We
626 identified a small (~6% of all macrophages) but distinct population of macrophages that
627 expressed SMC-specific genes and displayed a fibrotic pathway phenotype regulated by known
628 SMC transcription factors^{26,54}. Interestingly, these macrophages also expressed classic
629 macrophage proteins, which has been observed previously in advanced human coronary
630 atherosclerosis³⁰, suggesting they are of SMC origin. This finding aligns with previous mouse
631 studies showing SMC plasticity during atherosclerosis progression, including SMC-derived
632 macrophages, fibroblasts, osteogenic cells as well as foam cells^{31,34,35,39}. Additionally, we
633 identified a substantial population of SMC and likely SMC-derived cells that had a foam cell
634 gene signature, with specific expression of genes involved in lipid metabolism, agreeing with
635 recent studies suggesting that a large portion of foam cells within atherosclerotic plaques may
636 be of SMC origin^{30,31}. We also observed two distinct EC populations, with one showing
637 upregulated fibrotic pathway genes suggestive of endothelial-to-mesenchymal transition. Our
638 findings provide human evidence that supports recent studies in mice, highlighting the
639 heterogeneity and transitions of ECs during atherosclerosis³⁴. Importantly, this EC
640 pathophysiological process has been linked to increased disease severity in mouse models of
641 atherosclerosis and in human coronary disease^{55,56}.

642
643 We found that specific cell-type alterations may associate with symptomatic clinical events. Our
644 clinical analyses, however, may have limitations due to small sample sizes, patient selection
645 biases, and time delays between cerebrovascular events and surgeries. Despite limitations, we
646 observed decreases in efferocytotic macrophages, activated ECs, contractile SMCs, and foamy
647 modulated SMCs, along with an increase in inflammatory SMCs in symptomatic plaques. These
648 alterations indicate a loss of efferocytotic macrophages, erosion of the EC monolayer, and a
649 shift towards an inflammatory SMC/SMC-derived state, in plaque vulnerability². Additionally,
650 macrophages in symptomatic plaques showed increased senescence, also reported to associate
651 with plaque vulnerability^{57,58}. SMCs in symptomatic plaques exhibited an increased glycolytic
652 gene signature, extending prior findings of altered metabolism in plaques overall of high-risk
653 patients⁵⁹. As a corollary, alterations in bioenergetic mechanisms in mouse models of
654 atherosclerosis have been found to control SMC differentiation and to be associated with
655 features of plaque instability^{21,60}. These clinical associations provide the foundation for pre-
656 clinical and therapeutic studies targeting specific cell types and molecular programs to improve
657 outcomes in atherosclerotic CVD, particularly in high-risk patients unresponsive to lipid-
658 lowering therapy.

659
660 Although our dataset provides extensive insights into human carotid atherosclerosis, its sample
661 size is still limited and prone to technical biases and artifacts. Findings need to be further
662 corroborated in human coronary arteries and for cardiac events. To gain more confidence in
663 identifying causal alterations associated with clinical CVD events, it is also crucial to increase the
664 number of patient samples and diversity, and to utilize advanced computational and integrative

665 genomics methods and large-scale human genetics. These integrative studies can help identify
666 specific cell types and molecular programs that play a causal role, enabling mechanism-based
667 translation for clinical and therapeutic applications.

668
669 In summary, we present the most advanced in-depth multimodal atlas of advanced human
670 carotid atherosclerosis, reveal a vast cell-type heterogeneity of macrophages, SMCs, and
671 fibroblasts, and identify cell type-specific alterations associated with clinical events. We
672 characterized SMC and fibroblasts in unprecedented depth, revealing 3 populations of
673 phenotypically modulated SMCs, including a foamy population that resides in the deep intima
674 of the developing lesion. Further, we have uncovered complex macrophage subpopulations
675 with distinct gene, protein, and transcriptional regulatory features. These data are foundational
676 for future studies mapping CVD susceptible loci to specific cell types, with the goal of identifying
677 new CVD targets specific for certain cell populations fundamental to the progression and
678 regression of disease. Such integrative approaches provide the analytical and conceptual
679 framework for the discovery of novel therapeutics that target specific cell populations in CVD.

680 681 **Acknowledgements**

682
683 The scRNA-seq and CITE-seq (10x Genomics) was performed in the JP Sulzberger Columbia
684 Genome Center, supported in part through the National Institutes of Health/National Cancer
685 Institute Cancer Center Support Grant P30CA013696, and used the Genomics and High
686 Throughput Screening Shared Resource. Flow cytometry experiments described in this article
687 were performed in the Columbia Stem Cell Initiative Flow Cytometry core facility at Columbia
688 University Irving Medical Center. Confocal images were collected in the Confocal and
689 Specialized Microscopy Shared Resource of the Herbert Irving Comprehensive Cancer Center at
690 Columbia University, supported by National Institutes of Health (NIH) grant no. P30 CA013696
691 (National Cancer Institute).

692 693 **Sources of Funding**

694
695 A.C.B. is supported by the National Institute of Health Postdoctoral Training in Arteriosclerosis
696 fellowship (5T32HL007343). M.L. is supported by National Institutes of Health grant Nos.
697 R01GM125301, R01HL113147, R01HL150359, and R21HL156234. M.P.R. is supported for this
698 work by National Institutes of Health grant Nos. R01HL113147, R01HL150359, and
699 R01HL166916. A.C. is support by American Heart Association Predoctoral Fellowship 909206.
700 R.C.B. is supported for this work by National Institutes of Health grant Nos. R01HL141745 and
701 R01DK134026

702 **References**

703

- 704 1. Libby, P. The changing landscape of atherosclerosis. *Nature* **592**, 524-533 (2021).
- 705 2. Bentzon, J.F., Otsuka, F., Virmani, R. & Falk, E. Mechanisms of plaque formation and
706 rupture. *Circ Res* **114**, 1852-1866 (2014).
- 707 3. Burke, A.P., *et al.* Coronary risk factors and plaque morphology in men with coronary
708 disease who died suddenly. *N Engl J Med* **336**, 1276-1282 (1997).
- 709 4. Finn, A.V., Nakano, M., Narula, J., Kolodgie, F.D. & Virmani, R. Concept of
710 vulnerable/unstable plaque. *Arterioscler Thromb Vasc Biol* **30**, 1282-1292 (2010).
- 711 5. Aragam, K.G., *et al.* Discovery and systematic characterization of risk variants and genes
712 for coronary artery disease in over a million participants. *Nat Genet* **54**, 1803-1815
713 (2022).
- 714 6. Getz, G.S. & Reardon, C.A. Animal models of atherosclerosis. *Arterioscler Thromb Vasc*
715 *Biol* **32**, 1104-1115 (2012).
- 716 7. Stoeckius, M., *et al.* Simultaneous epitope and transcriptome measurement in single
717 cells. *Nat Methods* **14**, 865-868 (2017).
- 718 8. Fernandez, D.M., *et al.* Single-cell immune landscape of human atherosclerotic plaques.
719 *Nat Med* **25**, 1576-1588 (2019).
- 720 9. Depuydt, M.A.C., *et al.* Microanatomy of the Human Atherosclerotic Plaque by Single-
721 Cell Transcriptomics. *Circ Res* **127**, 1437-1455 (2020).
- 722 10. Hao, Y., *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587
723 e3529 (2021).
- 724 11. Wolf, F.A., Angerer, P. & Theis, F.J. SCANPY: large-scale single-cell gene expression data
725 analysis. *Genome Biol* **19**, 15 (2018).
- 726 12. Stuart, T., *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902
727 e1821 (2019).
- 728 13. Lakkis, J., *et al.* A joint deep learning model enables simultaneous batch effect
729 correction, denoising, and clustering in single-cell transcriptomics. *Genome Res* **31**,
730 1753-1766 (2021).
- 731 14. Cossarizza, A., *et al.* Guidelines for the use of flow cytometry and cell sorting in
732 immunological studies (third edition). *Eur J Immunol* **51**, 2708-3145 (2021).
- 733 15. Dobnikar, L., *et al.* Disease-relevant transcriptional signatures identified in individual
734 smooth muscle cells from healthy mouse vessels. *Nat Commun* **9**, 4567 (2018).
- 735 16. Lakkis, J., *et al.* A multi-use deep learning method for CITE-seq and single-cell RNA-seq
736 data integration with cell surface protein prediction and imputation. *Nat Mach Intell* **4**,
737 940+ (2022).
- 738 17. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion
739 for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
- 740 18. Tirosh, I., *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-
741 cell RNA-seq. *Science* **352**, 189-196 (2016).
- 742 19. Martinez, F.O., Gordon, S., Locati, M. & Mantovani, A. Transcriptional profiling of the
743 human monocyte-to-macrophage differentiation and polarization: new molecules and
744 patterns of gene expression. *J Immunol* **177**, 7303-7311 (2006).

- 745 20. Liu, Y., Beyer, A. & Aebersold, R. On the Dependency of Cellular Protein Levels on mRNA
746 Abundance. *Cell* **165**, 535-550 (2016).
- 747 21. Newman, A.A.C., *et al.* Multiple cell types contribute to the atherosclerotic lesion fibrous
748 cap by PDGFRbeta and bioenergetic mechanisms. *Nat Metab* **3**, 166-181 (2021).
- 749 22. Ivashkiv, L.B. IFNgamma: signalling, epigenetics and roles in immunity, metabolism,
750 disease and cancer immunotherapy. *Nat Rev Immunol* **18**, 545-558 (2018).
- 751 23. Li, Q. & Verma, I.M. NF-kappaB regulation in the immune system. *Nat Rev Immunol* **2**,
752 725-734 (2002).
- 753 24. Staus, D.P., Blaker, A.L., Taylor, J.M. & Mack, C.P. Diaphanous 1 and 2 regulate smooth
754 muscle cell differentiation by activating the myocardin-related transcription factors.
755 *Arterioscler Thromb Vasc Biol* **27**, 478-486 (2007).
- 756 25. Lockman, K., *et al.* Sphingosine 1-phosphate stimulates smooth muscle cell
757 differentiation and proliferation by activating separate serum response factor co-
758 factors. *J Biol Chem* **279**, 42422-42430 (2004).
- 759 26. Cheng, P., *et al.* Smad3 regulates smooth muscle cell fate and mediates adverse
760 remodeling and calcification of the atherosclerotic plaque. *Nat Cardiovasc Res* **1**, 322-
761 333 (2022).
- 762 27. Morrow, D., *et al.* Notch and vascular smooth muscle cell phenotype. *Circ Res* **103**,
763 1370-1382 (2008).
- 764 28. Williams, J.W., Giannarelli, C., Rahman, A., Randolph, G.J. & Kovacic, J.C. Macrophage
765 Biology, Classification, and Phenotype in Cardiovascular Disease: JACC Macrophage in
766 CVD Series (Part 1). *J Am Coll Cardiol* **72**, 2166-2180 (2018).
- 767 29. Chou, E.L., *et al.* Aortic Cellular Diversity and Quantitative Genome-Wide Association
768 Study Trait Prioritization Through Single-Nuclear RNA Sequencing of the Aneurysmal
769 Human Aorta. *Arterioscler Thromb Vasc Biol* **42**, 1355-1374 (2022).
- 770 30. Allahverdian, S., Chehroudi, A.C., McManus, B.M., Abraham, T. & Francis, G.A.
771 Contribution of intimal smooth muscle cells to cholesterol accumulation and
772 macrophage-like cells in human atherosclerosis. *Circulation* **129**, 1551-1559 (2014).
- 773 31. Wang, Y., *et al.* Smooth Muscle Cells Contribute the Majority of Foam Cells in ApoE
774 (Apolipoprotein E)-Deficient Mouse Atherosclerosis. *Arterioscler Thromb Vasc Biol* **39**,
775 876-887 (2019).
- 776 32. Shankman, L.S., *et al.* KLF4-dependent phenotypic modulation of smooth muscle cells
777 has a key role in atherosclerotic plaque pathogenesis. *Nat Med* **21**, 628-637 (2015).
- 778 33. Hamers, A.A.J., *et al.* Human Monocyte Heterogeneity as Revealed by High-Dimensional
779 Mass Cytometry. *Arterioscler Thromb Vasc Biol* **39**, 25-36 (2019).
- 780 34. Alencar, G.F., *et al.* Stem Cell Pluripotency Genes Klf4 and Oct4 Regulate Complex SMC
781 Phenotypic Changes Critical in Late-Stage Atherosclerotic Lesion Pathogenesis.
782 *Circulation* **142**, 2045-2059 (2020).
- 783 35. Wirka, R.C., *et al.* Atheroprotective roles of smooth muscle cell phenotypic modulation
784 and the TCF21 disease gene as revealed by single-cell analysis. *Nat Med* **25**, 1280-1289
785 (2019).
- 786 36. Kim, K., *et al.* Transcriptome Analysis Reveals Nonfoamy Rather Than Foamy Plaque
787 Macrophages Are Proinflammatory in Atherosclerotic Murine Models. *Circ Res* **123**,
788 1127-1142 (2018).

- 789 37. Cochain, C., *et al.* Single-Cell RNA-Seq Reveals the Transcriptional Landscape and
790 Heterogeneity of Aortic Macrophages in Murine Atherosclerosis. *Circ Res* **122**, 1661-
791 1674 (2018).
- 792 38. Zerneck, A., *et al.* Meta-Analysis of Leukocyte Diversity in Atherosclerotic Mouse
793 Aortas. *Circ Res* **127**, 402-426 (2020).
- 794 39. Pan, H., *et al.* Single-Cell Genomics Reveals a Novel Cell State During Smooth Muscle Cell
795 Phenotypic Switching and Potential Therapeutic Targets for Atherosclerosis in Mouse
796 and Human. *Circulation* **142**, 2060-2075 (2020).
- 797 40. Conklin, A.C., *et al.* Meta-Analysis of Smooth Muscle Lineage Transcriptomes in
798 Atherosclerosis and Their Relationships to In Vitro Models. *Immunometabolism* **3**(2021).
- 799 41. Ord, T., *et al.* Single-Cell Epigenomics and Functional Fine-Mapping of Atherosclerosis
800 GWAS Loci. *Circ Res* **129**, 240-258 (2021).
- 801 42. Slenders, L., *et al.* Intersecting single-cell transcriptomics and genome-wide association
802 studies identifies crucial cell populations and candidate genes for atherosclerosis. *Eur*
803 *Heart J Open* **2**, oeab043 (2022).
- 804 43. Ma, W.F., *et al.* Enhanced single-cell RNA-seq workflow reveals coronary artery disease
805 cellular cross-talk and candidate drug targets. *Atherosclerosis* **340**, 12-22 (2022).
- 806 44. Mulder, K., *et al.* Cross-tissue single-cell landscape of human monocytes and
807 macrophages in health and disease. *Immunity* **54**, 1883-1900 e1885 (2021).
- 808 45. Ridker, P.M., *et al.* Antiinflammatory Therapy with Canakinumab for Atherosclerotic
809 Disease. *N Engl J Med* **377**, 1119-1131 (2017).
- 810 46. Fidler, T.P., *et al.* The AIM2 inflammasome exacerbates atherosclerosis in clonal
811 haematopoiesis. *Nature* **592**, 296-301 (2021).
- 812 47. Endo-Umeda, K., *et al.* Myeloid LXR (Liver X Receptor) Deficiency Induces Inflammatory
813 Gene Expression in Foamy Macrophages and Accelerates Atherosclerosis. *Arterioscler*
814 *Thromb Vasc Biol* **42**, 719-731 (2022).
- 815 48. Patterson, M.T., *et al.* Trem2 Promotes Foamy Macrophage Lipid Uptake and Survival in
816 Atherosclerosis. *bioRxiv*, 2022.2011.2028.518255 (2022).
- 817 49. Lin, J.D., *et al.* Single-cell analysis of fate-mapped macrophages reveals heterogeneity,
818 including stem-like properties, during atherosclerosis progression and regression. *JCI*
819 *Insight* **4**(2019).
- 820 50. Robbins, C.S., *et al.* Local proliferation dominates lesional macrophage accumulation in
821 atherosclerosis. *Nat Med* **19**, 1166-1172 (2013).
- 822 51. Gerlach, B.D., *et al.* Efferocytosis induces macrophage proliferation to help resolve
823 tissue injury. *Cell Metab* **33**, 2445-2463 e2448 (2021).
- 824 52. Vallejo, J., Cochain, C., Zerneck, A. & Ley, K. Heterogeneity of immune cells in human
825 atherosclerosis revealed by scRNA-Seq. *Cardiovasc Res* **117**, 2537-2543 (2021).
- 826 53. Wissler, R.W. The arterial medial cell, smooth muscle, or multifunctional mesenchyme?
827 *Circulation* **36**, 1-4 (1967).
- 828 54. Cheng, P., *et al.* ZEB2 Shapes the Epigenetic Landscape of Atherosclerosis. *Circulation*
829 **145**, 469-485 (2022).
- 830 55. Chen, P.Y., *et al.* Endothelial-to-mesenchymal transition drives atherosclerosis
831 progression. *J Clin Invest* **125**, 4514-4528 (2015).

- 832 56. Evrard, S.M., *et al.* Endothelial to mesenchymal transition is common in atherosclerotic
833 lesions and is associated with plaque instability. *Nat Commun* **7**, 11853 (2016).
- 834 57. Childs, B.G., *et al.* Senescent intimal foam cells are deleterious at all stages of
835 atherosclerosis. *Science* **354**, 472-477 (2016).
- 836 58. Tabas, I. & Lichtman, A.H. Monocyte-Macrophages and T Cells in Atherosclerosis.
837 *Immunity* **47**, 621-634 (2017).
- 838 59. Tomas, L., *et al.* Altered metabolism distinguishes high-risk from stable carotid
839 atherosclerotic plaques. *Eur Heart J* **39**, 2301-2310 (2018).
- 840 60. Mayr, M., *et al.* Proteomic and metabolomic analyses of atherosclerotic vessels from
841 apolipoprotein E-deficient mice reveal alterations in inflammation, oxidative stress, and
842 energy metabolism. *Arterioscler Thromb Vasc Biol* **25**, 2135-2142 (2005).
- 843

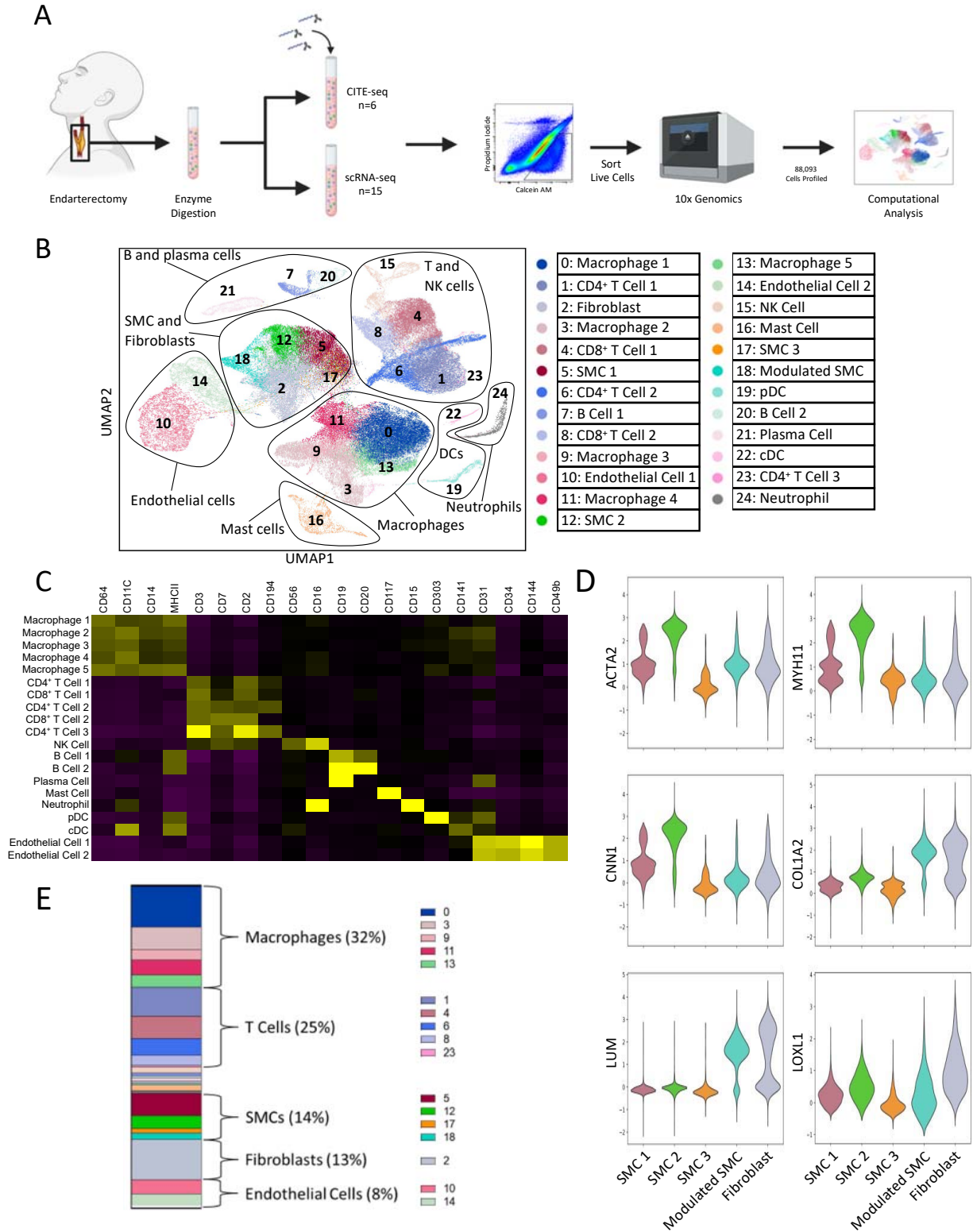
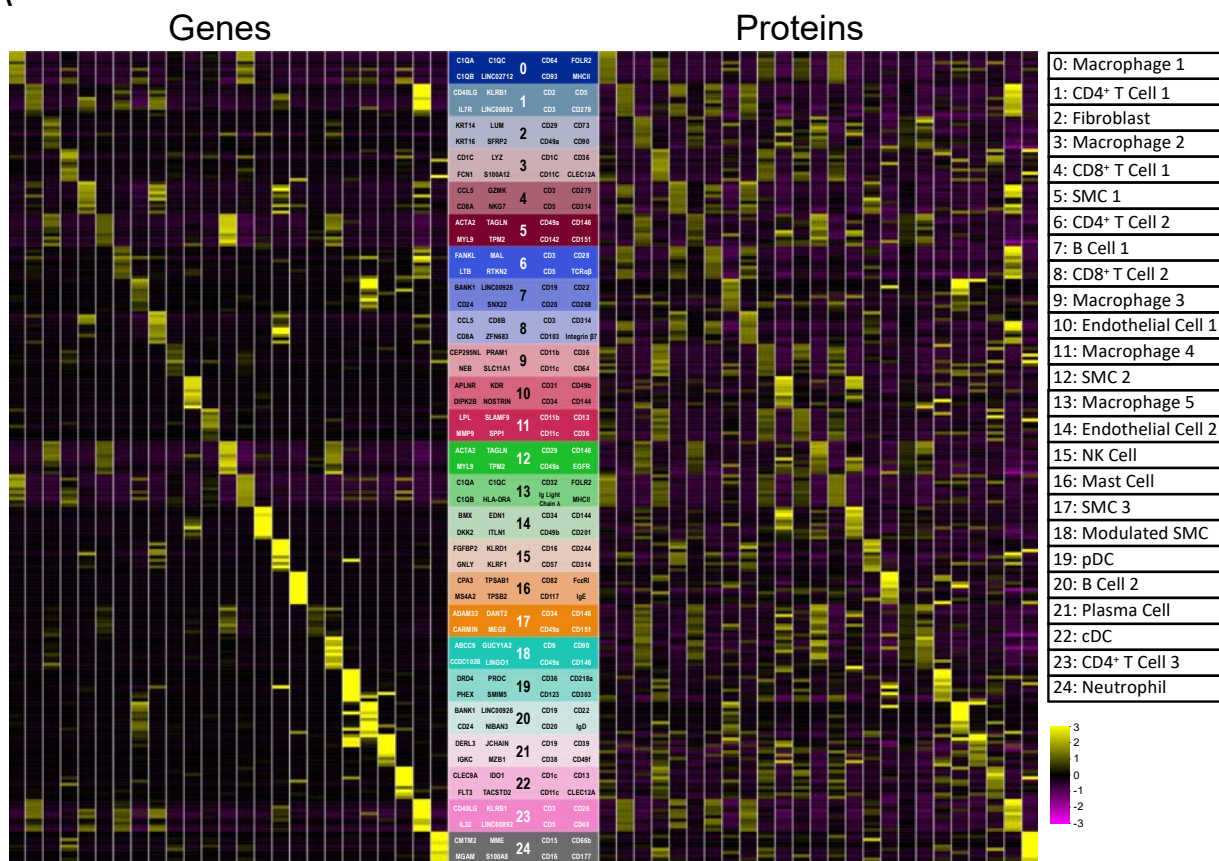
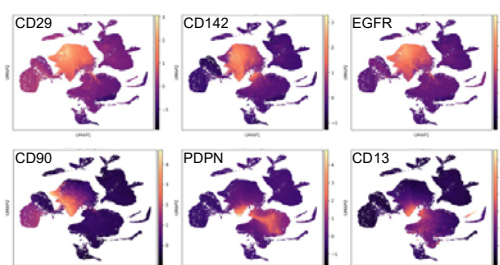


Figure 1. Multiomic analysis of human carotid atherosclerosis identifies 25 distinct cell populations. (A) Experimental design. (B) UMAP visualization of clustering revealed 25 cell populations. (C) Canonical Proteins to identify macrophages, T cells, NK cells, B Cells, Plasma Cells, Mast Cells, Neutrophils, dendritic cells, and endothelial cells. (D) Canonical Genes to identify SMCs and Fibroblasts. (E) Distribution of each major subclass across all samples.

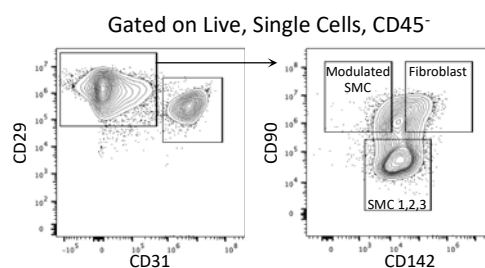
A



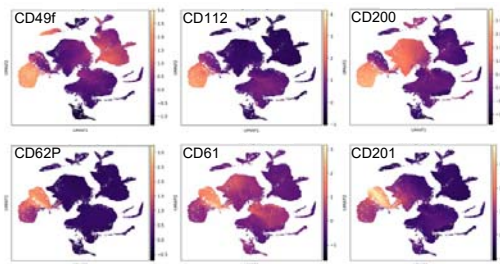
B



C



D



E

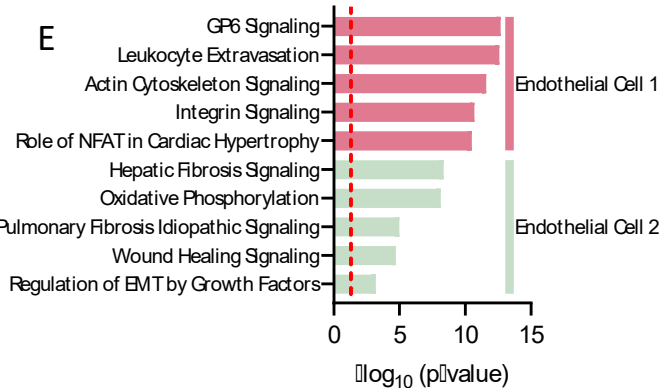


Figure 2. Multimodal biomarkers of cells within carotid atherosclerotic plaques. (A) Heatmap of all cell types identified in carotid atherosclerosis. Markers include top 10 mRNA (Left) and top 10 protein (Right) features identified by differential expression. (B) Feature plots displaying unique protein markers for SMCs (Top) and fibroblasts (bottom). (C) Flow cytometry validation of predicted protein markers for SMC 1-3, modulated SMC, and fibroblasts. (D) Feature plots displaying proteins to differentiate endothelial cell 1 and endothelial cell 2. (E) Pathway analysis comparing endothelial cell 1 and endothelial 2 using genes from differential expression analysis.

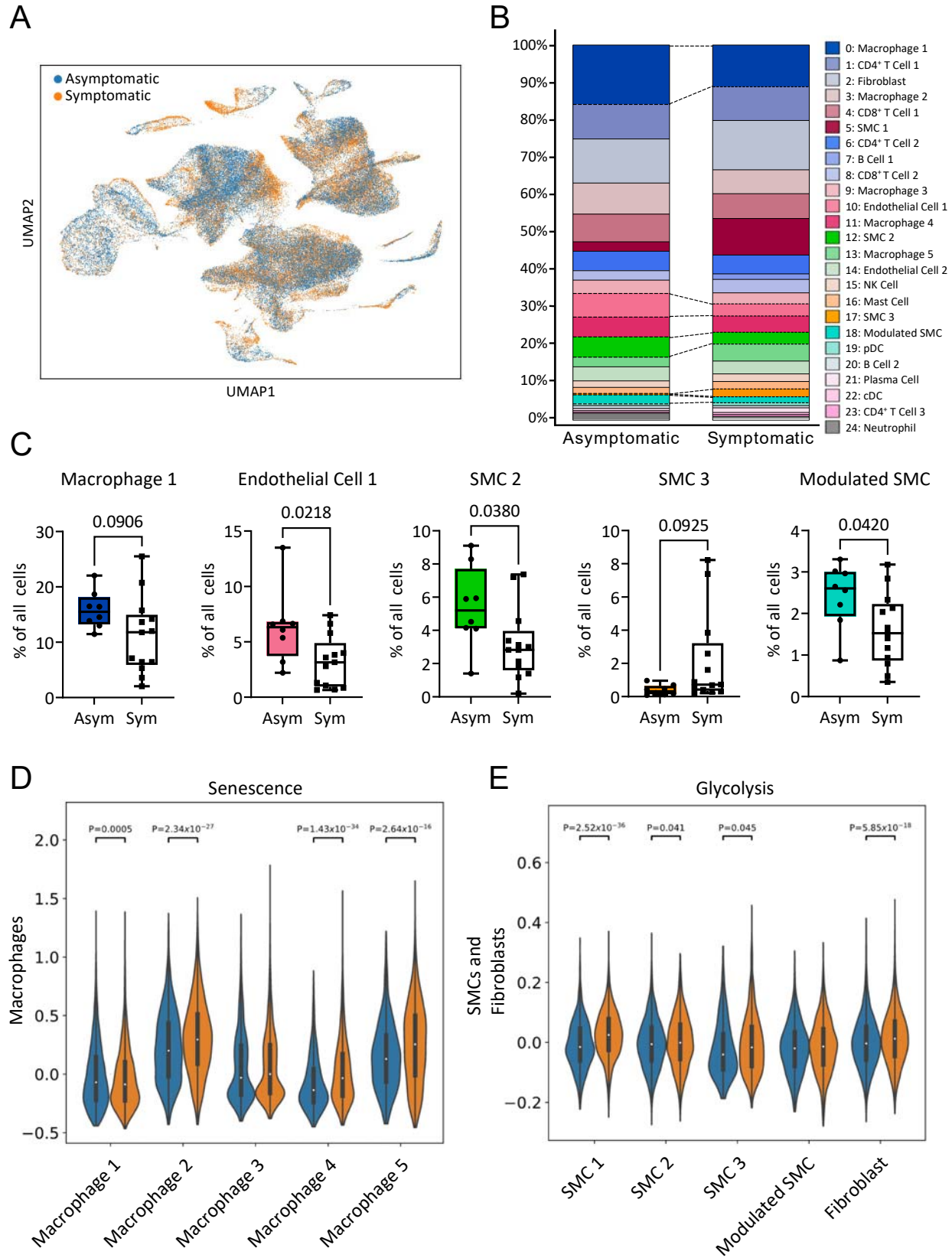


Figure 3. Cerebrovascular events are associated with alterations in the distribution of certain cell populations. (A) UMAP colored by clinical status. (B) Vertical bar graph showing proportion of all clusters by clinical status. (C) Box Plots showing distribution of select cell populations including macrophage 1, endothelial cell 1, SMC 2, SMC 3, and modulated SMC. (D) Gene scoring analysis comparing senescence in macrophage clusters in asymptomatic and symptomatic plaques. (E) Gene scoring analysis comparing glycolysis in SMC and fibroblast clusters in asymptomatic and symptomatic plaques.

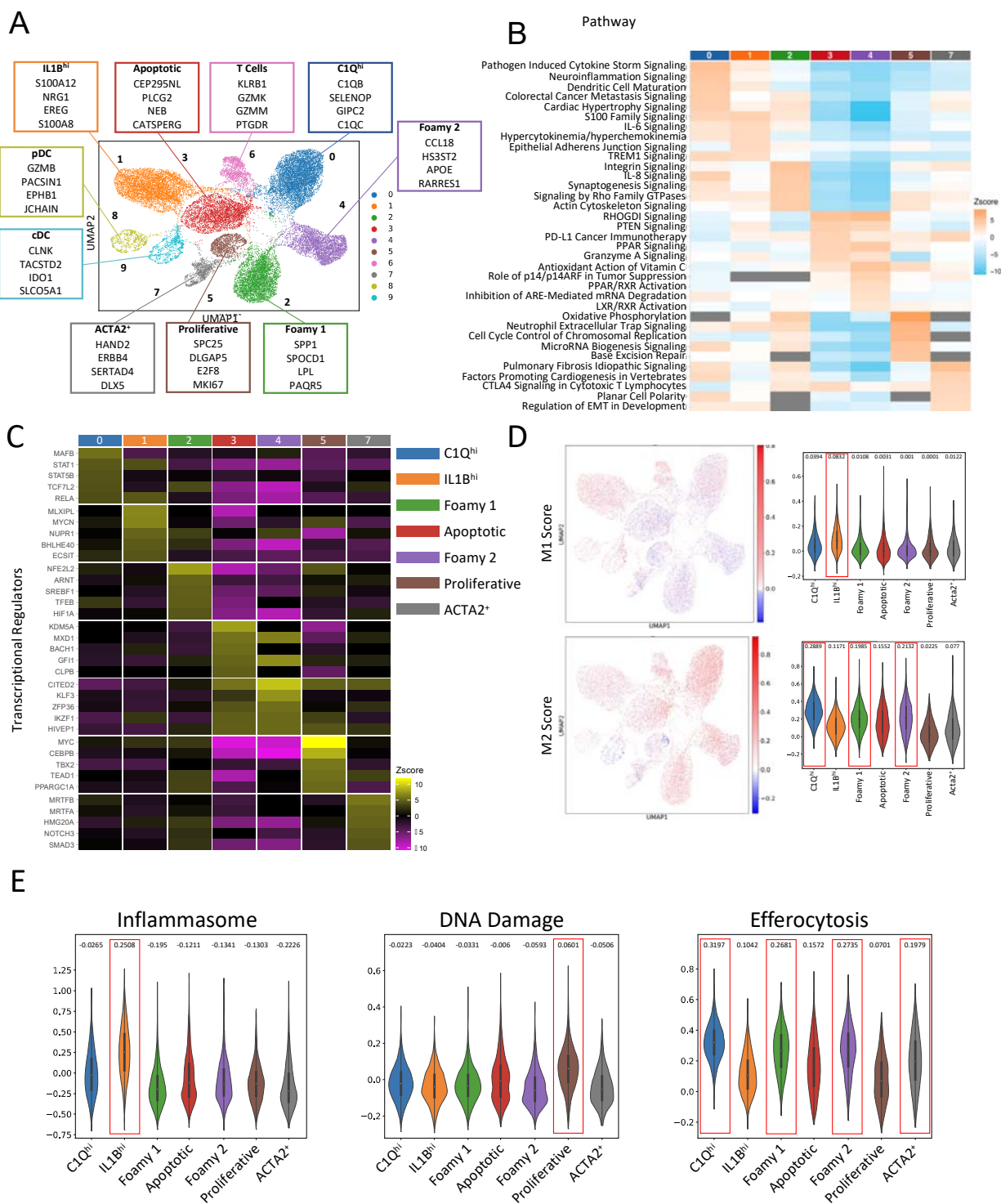


Figure 4. Sub-clustering analysis of myeloid cells reveals further macrophage phenotypic and functional heterogeneity. (A) Sub-clustering analysis of myeloid cells. Macrophage 1-5, pDC, and cDC clusters were selected from the initial clustering analysis, revealing 10 clusters. (B) Heatmap showing top upregulated pathways for each macrophage subpopulation based on top DE genes. (C) Heatmap showing top predicted transcriptional regulators for each macrophage subpopulation. (D) UMAP overlaying expression of common M1 and M2 signatures (left), and violin plots quantifying median expression for each macrophage subpopulation. (E) Gene scoring analysis for inflammasome activation, DNA damage, and efferocytosis across macrophage subpopulations.

Figure 5. Sub-clustering of SMC and Fibroblasts reveals distinct subpopulations. (A) Sub-clustering analysis of SMC and fibroblasts. SMC 1-3, modulated SMC, and fibroblast clusters were selected from the initial analysis, revealing 10 clusters. (B) Dotplot showing expression of classic SMC genes with corresponding phenotypes. (C) Radar plot showing pathways for SMCs obtained from DE genes. (D) Radar plot showing pathways for fibroblasts obtained from DE genes. (E) Immunohistochemistry staining targeting foamy SMC markers in coronary artery section. DAPI (blue), Smoothelin (green), LipidTOX (red), and CD90 (magenta). White arrows indicate Smoothelin⁺LipidTOX⁺CD90⁺ cells in neointima region of lesion. n=1. (F) Box plots showing distribution of select cell populations including foamy SMC, minor SMC 2, and fibromyocyte comparing asymptomatic and symptomatic plaques.