Single-cell RNA-seq analysis of human coronary arteries using an enhanced workflow reveals SMC transitions and candidate drug targets

Wei Feng Ma[1,2], Chani J. Hodonsky[2], Adam W. Turner[2], Doris Wong[2,3], Yipei Song[2,4], Nelson B. Barrientos[2,5], Jose Verdezoto Mosquera[2,3], Clint L. Miller[2,3,5]*

Affiliations

[1]Medical Scientist Training Program, University of Virginia, Charlottesville, VA 22908, USA.

[2]Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA.

[3]Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA 22908, USA.

[4]Department of Computer Engineering, University of Virginia, Charlottesville, VA 22908, USA.

[5]Department of Public Health Sciences, University of Virginia, Charlottesville, VA 22908, USA.

*corresponding author

Correspondence:

Clint L. Miller, PhD. (clintm@virginia.edu)

University of Virginia

Center for Public Health Genomics

PO Box 800717

Charlottesville, VA 22908 USA

23 **Abstract**

24 **Background and Aims**

25 The atherosclerotic plaque microenvironment is highly complex, and selective agents that
26 modulate plaque stability or other plaque phenotypes are not yet available. We sought to
27 investigate the human atherosclerotic cellular environment using scRNA-seq to uncover
28 potential therapeutic approaches. We aimed to make our workflow user-friendly, reproducible,
29 and applicable to other disease-specific scRNA-seq datasets.

30 **Methods**

31 Here we incorporate automated cell labeling, pseudotemporal ordering, ligand-receptor
32 evaluation, and drug-gene interaction analysis into an enhanced and reproducible scRNA-seq
33 analysis workflow. Notably, we also developed an R Shiny based interactive web application to
34 enable further exploration and analysis of the scRNA dataset.

35 **Results**

36 We applied this analysis workflow to a human coronary artery scRNA dataset and revealed
37 distinct derivations of chondrocyte-like and fibroblast-like cells from smooth muscle cells
38 (SMCs), and show the key changes in gene expression along their de-differentiation path. We
39 highlighted several key ligand-receptor interactions within the atherosclerotic environment
40 through functional expression profiling and revealed several attractive avenues for future
41 pharmacological repurposing in precision medicine. Further, our interactive web application,
42 *PlaqView* (www.plaqview.com), allows other researchers to easily explore this dataset and
43 benchmark applicable scRNA-seq analysis tools without prior coding knowledge.

44 **Conclusions**

45 These results suggest novel effects of chemotherapeutics on the atherosclerotic cellular
46 environment and provide future avenues of studies in precision medicine. This publicly available
47 workflow will also allow for more systematic and user-friendly analysis of scRNA datasets in

other disease and developmental systems. *PlaqView* allows for rapid visualization and analysis of atherosclerosis scRNA-seq datasets without the need of prior coding experience. Future releases of *PlaqView* will feature additional larger scRNA-seq and scATAC-seq atherosclerosis-related datasets, thus providing a critical resource for the field by promoting data harmonization and biological interpretation.

Keywords

Atherosclerosis, scRNA, analysis, pseudotemporal, drug-target, pipeline, web application

57 **Abbreviations**

58 C7- complement component C7

59 CAD- coronary artery disease

60 CH- chondrocytes

61 CMP- common myeloid progenitor cells

62 DCN- decorin

63 DGIdb- drug-gene interaction database

64 EC- endothelial cells

65 EGFR- epidermal growth factor receptor

66 FB- fibroblasts

67 FBLN1- fibulin 1

68 GMP- granulocyte-monocyte progenitor cells.

69 Mø- macrophages

70 MYH11- myosin heavy chain 11

71 SC- stem cells

72 sc/snATAC-seq - single cell/single nucleus assay for transposase-accessible chromatin

73 sequencing

74 sc/snRNA-seq- single cell/single nucleus RNA sequencing

75 TI- trajectory inference

76 SMC- smooth muscle cells

77 UMAP- uniform manifold approximation and projection

78

Background

Atherosclerosis is a complex process involving chronic inflammation and hardening of the vessel wall and represents one of the major causes of coronary artery disease (CAD), peripheral artery disease, and stroke [1]. Rupture of an unstable atherosclerotic lesion can lead to the formation of a thrombus, causing complete or partial occlusion of a coronary artery [2]. The contribution of smooth muscle cells (SMCs) to both lesion stability and progression has recently been established by numerous groups. However, the exact mechanisms by which SMCs modulate the atherosclerotic microenvironment and whether pharmacological agents can be used to selectively counter SMC-related deleterious mechanisms are still under investigation [3–5]

Recent advances in single-cell RNA-sequencing (scRNA-seq) have enabled ultra-fine gene expression profiling of many diseases at the cellular level, including atherosclerotic coronary artery disease [5]. As sequencing costs continue to decline, there has also been a consistent growth in scRNA datasets, data analysis tools and applications [6]. Currently, a major challenge with scRNA-seq analysis is the inherent bias introduced during manual cell labeling, in which cells are grouped by clusters and their identities called collectively based on their overall differential gene expression profiles [7]. Another draw-back inherent to commonly used scRNA-seq protocols is their destructive nature to the cells, making time-series analyses of the same cells impossible. Instead, these studies must rely on time-points from separate libraries to monitor processes such as clonal expansion and cell differentiation [8,9].

Recently, new approaches have been developed to compensate for both of these shortcomings, namely automatic cell labeling, pseudotemporal analysis, and trajectory inference. Tools such

103 as ‘SingleR’, ‘scCATCH’, and ‘Garnett’ have been used to assign unbiased identities to
104 individual cells using reference-based and machine learning algorithms, respectively [7,10,11].
105 Moreover, tools such as ‘Monocle3’ and ‘scVelo’ align and project cells onto a pseudotemporal
106 space where each cell becomes a snapshot within the single-cell time continuum [12,13]. In
107 essence, the single scRNA-seq dataset is transformed into a time series [12–14]. Although the
108 pseudotemporal scale does not reflect the actual time scale, it is a reliable approximation to
109 characterize cell fate and differentiation events, e.g., during organogenesis, disease states, or in
110 response to SARS-CoV-2 infections [13,15]
111
112 In this study, we present the application of an enhanced, scalable, and user-friendly scRNA-seq
113 analysis workflow on an existing human coronary artery scRNA-seq dataset. We performed
114 unbiased automatic cell identification at the single-cell level, pseudotemporal analysis, ligand-
115 receptor expression profiling, and drug repurposing analysis. Our results demonstrate potential
116 new mechanisms by which SMCs contribute to the atherosclerotic phenotype and signaling
117 within the lesion microenvironment. More importantly, we revealed attractive candidate avenues
118 for future pharmacological interventional studies. We also developed an interactive web
119 application to allow other users to explore this dataset. This reproducible analysis pipeline and
120 application can also be easily modified to incorporate different tissue data sources and single-
121 cell modalities such as scATAC-seq [16] or CITE-seq [17], and could serve as a template to
122 analyze and visualize single-cell datasets in other disease models.
123
124 **Results and Discussion**
125 **Unbiased automatic cell labeling reveals abundant cells with chondrocyte and fibroblast**
126 **characteristics.** Recently, automatic cell identification tools have been introduced to

127 compensate for the shortcomings of manual, cluster-based cell labeling [7]. For example,
128 'SingleR' and 'Garnett' use reference-based and machine learning algorithms, respectively, to
129 call individual cell identities [7,10,18]. Using 'SingleR,' which uses known purified cell
130 expression data as references, we found that endothelial cells (ECs) make up the highest
131 proportion of cells in this dataset (16.21%, Figure 1A-B), followed by smooth muscle cells
132 (SMCs, 13.8%) and stem cells (SCs, 14.06%), where the latter could be so-called
133 "atherosclerotic stem cells" or normal stromal stem cells but cannot be distinguished until
134 specific expression profile references are developed in the future [8]. Consistent with recent
135 scRNA-seq studies in atherosclerotic models, we identified abundant fibroblast (FB) and
136 chondrocyte-like (CH) cells, as well as cells with an osteoblast-like (OS) expression profile
137 (Figure 1B) [4]. In the UMAP clusters reflecting single-cell identities, there was a substantial
138 presence of SMC and FB cells in the OS and SC cluster. Such heterogeneity in cell clusters
139 would have been overlooked in manual cluster-based cell labeling. We also applied another
140 reference-based cell calling tool 'scCATCH,' and found that it underperforms relative to
141 'SingleR' and fails to provide consistent cell type assignment when provided with similar tissue
142 priors (Supplementary Table 1).
143
144 **Pseudotemporal ordering identifies distinct chondrocyte and fibroblast-like cell**
145 **differentiation states from smooth muscle cells.** To evaluate putative cell fate decisions or
146 differentiation events (e.g., SMC phenotypic transition states), we performed pseudotemporal
147 analysis and ultra-fine clustering using 'Monocle3', a method previously applied to normal and
148 diseased states, e.g., embryo organogenesis and response to coronavirus infection,
149 respectively [13,15]. We compared over 60 other trajectory inference (TI) methods including
150 Slingshot [19], PAGA [20], and SCORPIUS using the 'Dynverse' package [21]. However, we

151 found that most TI algorithms were unsuitable for complex tissue environments such as the
152 atherosclerotic plaque due to their inability to distinguish disconnected topologies
153 (Supplementary Table 2). In particular, we found evidence of SMCs giving rise to both
154 chondrocyte (CH) and fibroblast (FB)-like cells (Figure 1C). This corroborates earlier findings
155 showing that SMCs may transition or de-differentiate into 'fibromyocytes'—SMCs that have
156 undergone a phenotypic modulation to an extracellular matrix producing cell type within the
157 atherosclerotic lesions [5,8]. Genes associated with healthy SMC phenotypes, such as *MYH11*
158 (a canonical marker of SMC), *IGFBP2* (associated with decreased visceral fat), and *PPP1R14A*
159 (which enhances smooth muscle contraction), are decreased by approximately 50-75% along
160 the SMC trajectory as these cells become more FB-like (Table 1, Figure 1D, $p < 0.1E-297$)
161 [3,22]. Similar results were found by another group using mouse lineage-traced models where
162 MYH11 expression was decreased in SMC-derived modulated "intermediate cell states" [4].
163
164 More importantly, specific inflammatory markers and proteins associated with thrombotic events
165 during CAD, including complement proteins *C7* and *C3, FBLN1*, and *CXCL12* are increased
166 along the same trajectory (Figure 1D) [2,23]. Recent evidence suggests that CAD-associated
167 *CXCL12* secreted from endothelial cells may promote atherosclerosis[24]. Our results point to a
168 potentially new source of *CXCL12* that could be targeted to inhibit SMC-to-FB dedifferentiation.
169 Together, and in corroboration of recent studies, our pseudotemporal analysis demonstrates
170 that SMCs could be a source of both FB and CH-like cells, associated with intermediate and
171 advanced atherosclerotic phenotypes, respectively [4,5]. This is further supported by a recent
172 study, in which blocking of SMC-derived intermediate cells coincides with less severe
173 atherosclerotic lesions [4]. Precisely how these cells might influence the overall stability of

174 atherosclerotic lesions and clinical outcomes requires additional longitudinal studies using
175 genetic models and deep phenotyping of human tissues [3,4,8].
176
177 **Comprehensive ligand-receptor analysis shows complex intercellular communications in**
178 **the human coronary micro-environment and reveals potential drug targets.** To examine
179 the potential cross-talk between different cell types using scRNA-seq data, we compared the
180 ligand and receptor expression profiles of each cell type with experimentally-validated
181 interactions using 'scTalk' [25]. We found that there is an intricate network of signaling
182 pathways connecting different cell types; some cell types, such as OS, have stronger and more
183 frequent outgoing signals, whereas other cell types such as Macrophages (Mø) have fewer and
184 weaker incoming and outgoing signals (Figure 2A). SMCs, OSs, and neurons also exhibit a high
185 degree of autocrine signaling profiles (Figure 2A). Specifically, SMCs are shown to have the
186 highest number of outgoing signals and are among those with the least number of incoming
187 signal weights (Figure 2B). This suggests that SMCs play an important role in regulating the
188 coronary microenvironment by transducing signals to neighboring cells in the lesion.
189
190 Specifically, ten signaling pathways were identified between SMC, FB, CH, and OS
191 (Supplemental Figure 1). These pathways involve C3 complement, fibulin-1 *(FBLN1)*,
192 apolipoprotein D *(APOD)*, decorin *(DCN)*, and matrix metalloprotease 2 (*MMP2*, Figure 2C,
193 Supplemental Figure 1). We searched for potentially druggable targets to interrupt SMC-FB
194 communication by performing an integrative analysis of the identified ligand-receptor
195 interactions and the druggable genome using several drug-gene interaction databases,
196 including DGIdb 3.0 and Pharos [26]. We found that SMCs signals to FBs through the
197 complement protein *C3* and syndecan-2 *(SDC2)* via *MMP2,* and these two pathways can be

198 disrupted by drugs such as compstatin and tanomastat, respectively (Figure 2C). Multiple
199 studies have linked *C3* and the complement system to atherosclerotic lesion maturation in
200 mouse models, and a recent case study showed that the C3 targeted inhibitor, compstatin
201 (AMY-101), may prevent cardiovascular complications in patients with severe COVID-19
202 pneumonia [27–30]. Our results provide a potential mechanistic explanation by which SMCs can
203 modulate the inflammatory environment and plaque formation. Further, a recent study
204 demonstrated that *microRNA-9* repression of *SDC2* impedes atherosclerosis formation [31],
205 while MMP2 alteration also contributes to atherosclerosis in mouse models [32]. Here, we show
206 that SMCs signal to FBs within the atherosclerotic environment via *SDC2-MMP2*, and reveal
207 additional upstream candidate drug therapies that may influence atherosclerosis progression.
208
209 Interestingly, anti-*EGFR* (epidermal growth factor receptor)-based cancer treatments such as
210 erlotinib, cetuximab, and gefitinib were identified as potential key mediators of signaling
211 pathways between SMCs and FBs via decorin (*DCN*) and *EGFR* (Figure 2C). It has been shown
212 that *DCN* overexpression increases SMC aggregation and SMC-induced calcification within
213 atherosclerotic plaque [33]. Although the overlap between CAD and cancer etiology has been
214 previously noted, the long-term efficacy and cardiovascular impact of chemotherapy drugs, such
215 as erlotinib, requires further translational studies to investigate their potential use in cancer
216 patients to treat CAD [34–36].
217
218 Furthermore, integrative analysis of gene expression variation along the SMC-to-FB RNA
219 trajectory revealed similar results; as shown earlier in Figure 1D, the expression of complement
220 genes such as *C3* and *C7,* and chemokine *CXCL12* are increased as SMCs become more FB-
221 like. Although CXCL12 derived from endothelial cells have been recognized to promote CAD in

222 mouse models [24], here we provide a potentially new source of CXCL12 using human data and
223 found several pharmacological agents such as tinzaparin, an FDA-approved anticoagulant, to
224 investigate in future interventional studies. Together, this combination of independent analyses
225 of a scRNA-seq dataset reaffirms the druggable potential of these target genes.
226
227 ***PlaqView* is a user-friendly web application to explore atherosclerosis-related datasets.**
228 To enable other researchers to explore the transcriptomic landscape of the atherosclerotic
229 environment, we developed a web interface called *PlaqView* (www.plaqview.com, Figure 3A).
230 This interactive, R Shiny-based tool allows for multiple gene queries and comparisons of gene
231 expression, cell-labeling methods, RNA-trajectory tools, integrative drug-gene analysis, and
232 outputs high quality graphs and detailed tables. To our knowledge, there are no publicly
233 available tools to visualize atherosclerosis-related single-cell datasets without prior coding
234 knowledge. Further, *PlaqView* is under active development and will be releasing new datasets
235 coincidently with future atherosclerosis-related publications. As the database in *PlaqView*
236 expands along with the growing number of single-cell datasets, we anticipate that it will become
237 an essential tool for the atherosclerosis research field.
238
239 **Limitations**
240 Despite the advances presented in this workflow, we are still working to improve several
241 limitations. For instance, the default reference in 'SingleR' cannot identify more recently
242 discovered cellular phenotypes such as fibromyocytes [5,7,37], which may require a
243 combination of manual and automated labeling methods for accurate identification. Factors such
244 as the intrinsic heterogeneity of the tissue sample, disease stage, and tissue processing
245 artifacts are difficult to isolate computationally and could influence automated labeling methods.
246 Nonetheless, interactive viewers and tools such as *PlaqView* which incorporate multiple

247 methods and visualizations in one location, could help users separate out the technical and
248 biological variation from various single-cell datasets. Additionally, the modular nature of
249 *PlaqView* will also allow for future improvement of labeling methods as more precise reference
250 datasets are made available. Lastly, the true efficacy of proposed drugs cannot be verified
251 without extensive pre-clinical testing and clinical trials. Nonetheless, these findings may catalyze
252 future investigative efforts to develop more targeted therapies.
253

254 **Conclusions**

255 Our findings show that an enhanced, reproducible pipeline for scRNA-seq analysis has the
256 potential to improve upon current standard scRNA-seq bioinformatics protocols. For instance,
257 we provide new insights into intricate vascular cell differentiation and communication pathways
258 while providing actionable and testable targets for future experimental studies (Figure 3B). In
259 our combined analysis, we found that SMCs give rise to substantial proportions of CH and FB,
260 with the latter associated with worse prognostic markers [2,4,22,23]. SMCs signal to FBs via a
261 series of pathways involving C3 complement and MMP2, whose expression coincidentally
262 increases along the SMC-to-FB trajectory. We revealed possible therapeutic avenues that may
263 disrupt these cell communication pathways and alter the atherosclerotic pathology.
264 Furthermore, several FDA-approved drugs (e.g., erlotinib, cetuximab, and gefitinib) were shown
265 as potential effectors of SMC signaling to FB, and may be used to treat CAD in cancer patients
266 to simplify or augment drug regiments [34]. This is consistent with recent reports showing
267 beneficial effects of the acute promyelocytic drug all-trans-retinoic acid (ATRA) in
268 atherosclerosis mouse models [4]. Further investment in scRNA-seq may also help resolve the
269 balance of anti-tumor efficacy and atheroprotection for immune checkpoint inhibitors as well as
270 immunomodulators at the interface of cardio-immuno-oncology [38].

271

272 Although the utilization of this workflow can compensate for many of the shortcomings of current
273 scRNA-seq analyses, we are still unable to perform cell-lineage tracing that reflects actual
274 timescales without additional gene engineering experiments *in vivo* [39]. However, leveraging
275 mitochondrial DNA variants in snATAC-seq data has enabled lineage tracing analysis in human
276 cells [40,41]. Likewise, these analyses can ultimately be extended to integrate spatial omics and
277 other multi-modal data [42]. In the future as spatial transcriptomics, scATAC-seq, and/or CITE-
278 seq data become more widely available, this workflow can be modified to discover signaling
279 pathways or differentiation events at specific tissue locations and timepoints, allowing for more
280 disease-relevant drug-gene interaction analyses (Figure 3B). Nonetheless, this pipeline can be
281 applied immediately to datasets from other tissues/diseases to generate informative directions
282 for follow-up studies, and is more user-friendly and reproducible compared to standard scRNA
283 analyses. Lastly, building web applications that democratize the access and analysis of single-
284 cell data will promote collaboration and innovation across disciplines [43]. As *PlaqView*
285 incorporates additional relevant single-cell datasets, we anticipate that this application will
286 become an indispensable resource for the community.

287

288 **Methods**

289 **Data retrieval and pre-processing**

290 Human coronary artery scRNA data read count matrix was retrieved from the Gene Expression
291 Omnibus (GEO) using #GSE131780 and loaded into R 4.0, and was preprocessed using
292 standard parameters of the R packages 'Seurat' v.3, and 'Monocle3' as required [5,44–46].
293 Uniform manifold approximation projections (UMAP) clusters from 'Seurat' were imported into
294 'Monocle3' before pseudotemporal analysis.

295

296 **Automatic cell Identification and pseudotemporal ordering**

297 scRNA read matrices were read into SingleR as previously described for cell labeling [7].
298 SingleR compares each cell's gene expression profile with known human primary cell atlas data
299 and gives the most likely cell identity independently. SingleR first corrects for batch effects, then
300 calculates the expression correlation scores for each test cell to each cell type in the reference,
301 and the cell identity is called based on reference cell type exhibiting the highest correlation.
302 Then, pseudotemporal analyses were performed as previously described in the analysis of
303 embryo organogenesis [13,45]. Briefly, the UMAP clusters were passed into Monocle3 and then
304 the 'learn_graph()' and 'order_cells()' functions. The SMCs and related clusters were then
305 subsetted for detailed subclustering and analysis. For each cluster, Moran's I statistics were
306 calculated, which identify genes that are differentially expressed along their trajectories.
307 Additionally, we explored 60+ other trajectory inference methods using the 'Dynverse' package,
308 where we simulated mock dataset of 5000 cells with known disconnected graph topologies [21].
309 Detailed codes to reproduce the figures in this publication can be found at the Miller Lab Github
310 (see availability of data and materials).

311

312 **Ligand-receptor cell communication analysis**

313 We analyzed candidate ligand-receptor interactions to infer cell communication using the R
314 package 'scTalk', as previously described in the analysis of glial cells [25]. This method is
315 based on permutation testing of random networks, where ligand-receptor interactions are
316 derived from experimentally derived interactions from the STRING database. We exported
317 statistically significant differentially expressed genes from 'Seurat' using the 'FindMarkers()'
318 function and imported the preprocessed data into 'scTalk.' Then, overall edges of the cellular

communication network were calculated using the 'GenerateNetworkPaths()' function, which reflects the overall ligand-receptor interaction strength between each cell type. Then, the cell types of interest were specified and treeplots were generated using the 'NetworkTreePlot()' function.

Gene-drug interaction analysis

The above identified ligand and receptor interaction pairs were fed into the Drug-Gene Interaction database (DGIdb 3.0) to reveal candidate drug-gene interactions [26]. Ligands and receptors that were deemed significant from 'scTalk' were evaluated using the 'queryDGIdb()' function of the 'rDGIdb' R package [26]. Additionally, we queried CTD2, OMIM, ClinVar, Pharos, GnomAD, and the ExAC databases using the docker-based tool 'DrugThatGene' [47]. We included all top FDA-approved drugs produced with verified inhibitory or antagonistic activities, as well as drugs that may influence or influenced by changes in the receptor. Figures 2C, 2D, and 3B were modified using BioRender for clarity.

Development of *PlaqView*

PlaqView is written in R and Shiny, and is hosted on the web at www.plaqview.com using dedicated shinyapps.io servers, and can be run locally through RStudio. The raw data was first processed as previously described, but packaged into .rds objects. These objects were then written into a Shiny script to allow for interactive display. This application is open-sourced and its code and data is available at github.com/MillerLab-CPHG/PlaqView. Further, *PlaqView* is actively recruiting available datasets on the github page.

Declarations

No conflicts of interest to disclose.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All data are publicly available on GEO, accession number GSE131778, as previously described [5]. All codes and analysis pipelines can be viewed at github.com/MillerLab-CPHG/Ma_2020 and repurposed for other scRNA datasets.

Competing interests

The authors declare no competing interests.

Funding

Funding support was provided by grants from the National Institutes of Health (NIH): R00HL125912 (CLM), R01HL148239 (CLM) and T32HL007284 (CJH and DW); American Heart Association (AHA): POST35120545 (AWT) and Leducq Foundation Transatlantic Network of Excellence (PlaqOmics) (CLM and AWT).

Authors' contributions

366 WFM designed and performed the statistical analysis. CJH, AWT, DW and YS refined the
367 methodology and edited the manuscript. JVM helped with scripting. CLM conceived and refined
368 the project and edited the manuscript.
369

370 **Acknowledgements**

371 Not applicable.

372

373 **Works cited**

374

375 [1] G.L. Basatemur, H.F. Jørgensen, M.C.H. Clarke, M.R. Bennett, Z. Mallat, Vascular smooth
376 muscle cells in atherosclerosis, Nat Rev Cardiol. 16 (2019) 727–744.
377 https://doi.org/10.1038/s41569-019-0227-9.

378 [2] W.S. Argraves, A. Tanaka, E.P. Smith, W.O. Twal, K.M. Argraves, D. Fan, C.C.
379 Haudenschild, Fibulin-1 and fibrinogen in human atherosclerotic lesions, Histochem Cell Biol.
380 132 (2009) 559–565. https://doi.org/10.1007/s00418-009-0628-7.

381 [3] M.R. Bennett, S. Sinha, G.K. Owens, Vascular Smooth Muscle Cells in Atherosclerosis., Circ
382 Res. 118 (2016) 692-- 702. https://doi.org/10.1161/circresaha.115.306361.

383 [4] H. Pan, C. Xue, B.J. Auerbach, J. Fan, A.C. Bashore, J. Cui, D.Y. Yang, S.B. Trignano, W.
384 Liu, J. Shi, C.O. Ihuegbu, E.C. Bush, J. Worley, L. Vlahos, P. Laise, R.A. Solomon, E.S.
385 Connolly, A. Califano, P.A. Sims, H. Zhang, M. Li, M.P. Reilly, Single-Cell Genomics Reveals a
386 Novel Cell State During Smooth Muscle Cell Phenotypic Switching and Potential Therapeutic
387 Targets for Atherosclerosis in Mouse and Human, Circulation. (2020).
388 https://doi.org/10.1161/circulationaha.120.048378.

389 [5] R.C. Wirka, D. Wagh, D.T. Paik, M. Pjanic, T. Nguyen, C.L. Miller, R. Kundu, M. Nagao, J.
390 Coller, T.K. Koyano, R. Fong, Y.J. Woo, B. Liu, S.B. Montgomery, J.C. Wu, K. Zhu, R. Chang,
391 M. Alamprese, M.D. Tallquist, J.B. Kim, T. Quertermous, Atheroprotective roles of smooth
392 muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell
393 analysis, Nat Med. 25 (2019) 1280--1289. https://doi.org/10.1038/s41591-019-0512-5.

394 [6] C. Yifan, Y. Fan, P. Jun, Visualization of cardiovascular development, physiology and
395 disease at the single-cell level: Opportunities and future challenges, J Mol Cell Cardiol. 142
396 (2020) 80–92. https://doi.org/10.1016/j.yjmcc.2020.03.005.

397 [7] D. Aran, A.P. Looney, L. Liu, E. Wu, V. Fong, A. Hsu, S. Chak, R.P. Naikawadi, P.J. Wolters,
398 A.R. Abate, A.J. Butte, M. Bhattacharya, Reference-based analysis of lung single-cell

399 sequencing reveals a transitional profibrotic macrophage, Nat Immunol. 20 (2019) 163–172.
400 https://doi.org/10.1038/s41590-018-0276-y.

401 [8] Y. Wang, V. Nanda, D. Direnzo, J. Ye, S. Xiao, Y. Kojima, K.L. Howe, K.-U. Jarr, A.M.
402 Flores, P. Tsantilas, N. Tsao, A. Rao, A.A.C. Newman, A.V. Eberhard, J.R. Priest, A.
403 Ruusalepp, G. Pasterkamp, L. Maegdefessel, C.L. Miller, L. Lind, S. Koplev, J.L.M. Björkegren,
404 G.K. Owens, E. Ingelsson, I.L. Weissman, N.J. Leeper, Clonally expanding smooth muscle cells
405 promote atherosclerosis by escaping efferocytosis and activating the complement cascade., P
406 Natl Acad Sci Usa. 117 (2020) 15818–15826. https://doi.org/10.1073/pnas.2006348117.

407 [9] D. DiRenzo, G.K. Owens, N.J. Leeper, "Attack of the Clones," Circ Res. 120 (2017) 624--
408 626. https://doi.org/10.1161/circresaha.116.310091.

409 [10] H.A. Pliner, J. Shendure, C. Trapnell, Supervised classification enables rapid annotation of
410 cell atlases, Nat Methods. 16 (2019) 983–986. https://doi.org/10.1038/s41592-019-0535-3.

411 [11] X. Shao, J. Liao, X. Lu, R. Xue, N. Ai, X. Fan, scCATCH: automatic annotation on cell-types
412 of clusters from single-cell RNA-seq data, Iscience. 23 (2020) 100882.
413 https://doi.org/10.1016/j.isci.2020.100882.

414 [12] V. Bergen, M. Lange, S. Peidli, F.A. Wolf, F.J. Theis, Generalizing RNA velocity to transient
415 cell states through dynamical modeling, Biorxiv. (2019) 820936. https://doi.org/10.1101/820936.

416 [13] J. Cao, M. Spielmann, X. Qiu, X. Huang, D.M. Ibrahim, A.J. Hill, F. Zhang, S. Mundlos, L.
417 Christiansen, F.J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape
418 of mammalian organogenesis, Nature. 566 (2019) 496–502. https://doi.org/10.1038/s41586-
419 019-0969-x.

420 [14] G.L. Manno, R. Soldatov, A. Zeisel, E. Braun, H. Hochgerner, V. Petukhov, K. Lidschreiber,
421 M.E. Kastriti, P. Lönnerberg, A. Furlan, J. Fan, L.E. Borm, Z. Liu, D. van Bruggen, J. Guo, X.
422 He, R. Barker, E. Sundström, G. Castelo-Branco, P. Cramer, I. Adameyko, S. Linnarsson, P.V.
423 Kharchenko, RNA velocity of single cells, Nature. 560 (2018) 494--498.
424 https://doi.org/10.1038/s41586-018-0414-6.

425 [15] R.L. Chua, S. Lukassen, S. Trump, B.P. Hennig, D. Wendisch, F. Pott, O. Debnath, L.
426 Thürmann, F. Kurth, M.T. Völker, J. Kazmierski, B. Timmermann, S. Twardziok, S. Schneider,
427 F. Machleidt, H. Müller-Redetzky, M. Maier, A. Krannich, S. Schmidt, F. Balzer, J. Liebig, J.
428 Loske, N. Suttorp, J. Eils, N. Ishaque, U.G. Liebert, C. von Kalle, A. Hocke, M. Witzenrath, C.
429 Goffinet, C. Drosten, S. Laudi, I. Lehmann, C. Conrad, L.-E. Sander, R. Eils, COVID-19 severity
430 correlates with airway epithelium–immune cell interactions identified by single-cell analysis, Nat
431 Biotechnol. (2020) 1–10. https://doi.org/10.1038/s41587-020-0602-4.

432 [16] J.P. Smith, N.C. Sheffield, Analytical Approaches for ATAC-seq Data Analysis, Curr Protoc
433 Hum Genetics. 106 (2020). https://doi.org/10.1002/cphg.101.

434 [17] M. Stoeckius, C. Hafemeister, W. Stephenson, B. Houck-Loomis, P.K. Chattopadhyay, H.
435 Swerdlow, R. Satija, P. Smibert, Simultaneous epitope and transcriptome measurement in
436 single cells, Nat Methods. 14 (2017) 865–868. https://doi.org/10.1038/nmeth.4380.

[18] M. Slyper, C.B.M. Porter, O. Ashenberg, J. Waldman, E. Drokhlyansky, I. Wakiro, C. Smillie, G. Smith-Rosario, J. Wu, D. Dionne, S. Vigneau, J. Jané-Valbuena, T.L. Tickle, S. Napolitano, M.-J. Su, A.G. Patel, A. Karlstrom, S. Gritsch, M. Nomura, A. Waghray, S.H. Gohil, A.M. Tsankov, L. Jerby-Arnon, O. Cohen, J. Klughammer, Y. Rosen, J. Gould, L. Nguyen, M. Hofree, P.J. Tramontozzi, B. Li, C.J. Wu, B. Izar, R. Haq, F.S. Hodi, C.H. Yoon, A.N. Hata, S.J. Baker, M.L. Suvà, R. Bueno, E.H. Stover, M.R. Clay, M.A. Dyer, N.B. Collins, U.A. Matulonis, N. Wagle, B.E. Johnson, A. Rotem, O. Rozenblatt-Rosen, A. Regev, A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors, Nat Med. 26 (2020) 792--802. https://doi.org/10.1038/s41591-020-0844-1.

[19] K. Street, D. Risso, R.B. Fletcher, D. Das, J. Ngai, N. Yosef, E. Purdom, S. Dudoit, Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics, Bmc Genomics. 19 (2018) 477. https://doi.org/10.1186/s12864-018-4772-0.

[20] F.A. Wolf, F.K. Hamey, M. Plass, J. Solana, J.S. Dahlin, B. Göttgens, N. Rajewsky, L. Simon, F.J. Theis, PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells, Genome Biol. 20 (2019) 59. https://doi.org/10.1186/s13059-019-1663-x.

[21] W. Saelens, R. Cannoodt, H. Todorov, Y. Saeys, A comparison of single-cell trajectory inference methods, Nat Biotechnol. 37 (2019) 547–554. https://doi.org/10.1038/s41587-019-0071-9.

[22] S. Carter, I. Lemieux, Z. Li, N. Alméras, A. Tremblay, J. Bergeron, P. Poirier, J.-P. Després, F. Picard, Changes in IGFBP-2 levels following a one-year lifestyle modification program are independently related to improvements in plasma apo B and LDL apo B levels, Atherosclerosis. 281 (2018) 89–97. https://doi.org/10.1016/j.atherosclerosis.2018.12.016.

[23] A.M. Carter, Complement Activation: An Emerging Player in the Pathogenesis of Cardiovascular Disease, Sci. 2012 (2012) 1–14. https://doi.org/10.6064/2012/402783.

[24] Y. Döring, E.P.C. van der Vorst, J. Duchene, Y. Jansen, S. Gencer, K. Bidzhekov, D. Atzler, D. Santovito, D.J. Rader, D. Saleheen, C. Weber, CXCL12 Derived From Endothelial Cells Promotes Atherosclerosis to Drive Coronary Artery Disease, Circulation. 139 (2019) 1338–1340. https://doi.org/10.1161/circulationaha.118.037953.

[25] N. Farbehi, R. Patrick, A. Dorison, M. Xaymardan, V. Janbandhu, K. Wystub-Lis, J.W. Ho, R.E. Nordon, R.P. Harvey, Single-cell expression profiling reveals dynamic flux of cardiac stromal, vascular and immune cells in health and injury, Elife. 8 (2019) e43882. https://doi.org/10.7554/elife.43882.

[26] K.C. Cotto, A.H. Wagner, Y.-Y. Feng, S. Kiwala, A.C. Coffman, G. Spies, A. Wollam, N.C. Spies, O.L. Griffith, M. Griffith, DGIdb 3.0: a redesign and expansion of the drug–gene interaction database, Nucleic Acids Res. 46 (2017) D1068–D1073. https://doi.org/10.1093/nar/gkx1143.

[27] C. Buono, C.E. Come, J.L. Witztum, G.F. Maguire, P.W. Connelly, M. Carroll, A.H. Lichtman, Influence of C3 Deficiency on Atherosclerosis, Circulation. 105 (2002) 3025–3031. https://doi.org/10.1161/01.cir.0000019584.04929.83.

[28] S. Mastaglio, A. Ruggeri, A.M. Risitano, P. Angelillo, D. Yancopoulou, D.C. Mastellos, M. Huber-Lang, S. Piemontese, A. Assanelli, C. Garlanda, J.D. Lambris, F. Ciceri, The first case of COVID-19 treated with the complement C3 inhibitor AMY-101, Clin Immunol. 215 (2020) 108450. https://doi.org/10.1016/j.clim.2020.108450.

[29] A. Java, A.J. Apicelli, M.K. Liszewski, A. Coler-Reilly, J.P. Atkinson, A.H.J. Kim, H.S. Kulkarni, The complement system in COVID-19: friend and foe?, Jci Insight. 5 (2020). https://doi.org/10.1172/jci.insight.140711.

[30] D. Pellegrini, R. Kawakami, G. Guagliumi, A. Sakamoto, K. Kawai, A. Gianatti, A. Nasr, R. Kutys, L. Guo, A. Cornelissen, L. Faggi, M. Mori, Y. Sato, I. Pescetelli, M. Brivio, M. Romero, R. Virmani, A.V. Finn, Microthrombi As A Major Cause of Cardiac Injury in COVID-19: A Pathologic Study, Circulation. (2021). https://doi.org/10.1161/circulationaha.120.051828.

[31] R. Zhang, B. Song, X. Hong, Z. Shen, L. Sui, S. Wang, microRNA-9 Inhibits Vulnerable Plaque Formation and Vascular Remodeling via Suppression of the SDC2-Dependent FAK/ERK Signaling Pathway in Mice With Atherosclerosis, Front Physiol. 11 (2020) 804. https://doi.org/10.3389/fphys.2020.00804.

[32] M. Kuzuya, K. Nakamura, T. Sasaki, X.W. Cheng, S. Itohara, A. Iguchi, Effect of MMP-2 Deficiency on Atherosclerotic Lesion Formation in ApoE-Deficient Mice, Arteriosclerosis Thrombosis Vasc Biology. 26 (2006) 1120–1125. https://doi.org/10.1161/01.atv.0000218496.60097.e0.

[33] J.W. Fischer, S.A. Steitz, P.Y. Johnson, A. Burke, F. Kolodgie, R. Virmani, C. Giachelli, T.N. Wight, Decorin Promotes Aortic Smooth Muscle Cell Calcification and Colocalizes to Calcified Regions in Human Atherosclerotic Lesions, Arteriosclerosis Thrombosis Vasc Biology. 24 (2004) 2391–2396. https://doi.org/10.1161/01.atv.0000147029.63303.28.

[34] C. Camaré, M. Pucelle, A. Nègre-Salvayre, R. Salvayre, Angiogenesis in the atherosclerotic plaque., Redox Biol. 12 (2017) 18-- 34. https://doi.org/10.1016/j.redox.2017.01.007.

[35] J.V. Tapia-Vieyra, B. Delgado-Coello, J. Mas-Oliva, Atherosclerosis and Cancer; A Resemblance with Far-reaching Implications, Arch Med Res. 48 (2017) 12–26. https://doi.org/10.1016/j.arcmed.2017.03.005.

[36] P. Libby, G.K. Hansson, Taming Immune and Inflammatory Responses to Treat Atherosclerosis ∗, J Am Coll Cardiol. 71 (2018) 173–176. https://doi.org/10.1016/j.jacc.2017.10.081.

[37] N.A. Mabbott, J.K. Baillie, H. Brown, T.C. Freeman, D.A. Hume, An expression atlas of human primary cells: inference of gene function from coexpression networks, Bmc Genomics. 14 (2013) 632. https://doi.org/10.1186/1471-2164-14-632.

511 [38] V.G. Zaha, W.C. Meijers, J. Moslehi, Cardio-Immuno-Oncology, Circulation. 141 (2020) 87–
512 89. https://doi.org/10.1161/circulationaha.119.042276.

513 [39] M. Liu, D. Gomez, Smooth Muscle Cell Phenotypic Diversity: At the Crossroads of Lineage
514 Tracing and Single-Cell Transcriptomics, Arteriosclerosis Thrombosis Vasc Biology. 39 (2019)
515 1715–1723. https://doi.org/10.1161/atvbaha.119.312131.

516 [40] C.A. Lareau, L.S. Ludwig, C. Muus, S.H. Gohil, T. Zhao, Z. Chiang, K. Pelka, J.M. Verboon,
517 W. Luo, E. Christian, D. Rosebrock, G. Getz, G.M. Boland, F. Chen, J.D. Buenrostro, N.
518 Hacohen, C.J. Wu, M.J. Aryee, A. Regev, V.G. Sankaran, Massively parallel single-cell
519 mitochondrial DNA genotyping and chromatin profiling, Nat Biotechnol. (2020) 1–11.
520 https://doi.org/10.1038/s41587-020-0645-6.

521 [41] J. Xu, K. Nuno, U.M. Litzenburger, Y. Qi, M.R. Corces, R. Majeti, H.Y. Chang, Single-cell
522 lineage tracing by endogenous mutations enriched in transposase accessible mitochondrial
523 DNA, Elife. 8 (2019) e45105. https://doi.org/10.7554/elife.45105.

524 [42] T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W.M. Mauck, Y. Hao, M.
525 Stoeckius, P. Smibert, R. Satija, Comprehensive Integration of Single-Cell Data, Cell. 177
526 (2019) 1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

527 [43] J. Gao, B.A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S.O. Sumer, Y. Sun, A.
528 Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, N. Schultz, Integrative Analysis of
529 Complex Cancer Genomics and Clinical Profiles Using the cBioPortal, Sci Signal. 6 (2013) pl1–
530 pl1. https://doi.org/10.1126/scisignal.2004088.

531 [44] R. Satija, J.A. Farrell, D. Gennert, A.F. Schier, A. Regev, Spatial reconstruction of single-
532 cell gene expression data., Nat Biotechnol. 33 (2015) 495-- 502.
533 https://doi.org/10.1038/nbt.3192.

534 [45] C. Trapnell, D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N.J. Lennon, K.J.
535 Livak, T.S. Mikkelsen, J.L. Rinn, The dynamics and regulators of cell fate decisions are revealed
536 by pseudotemporal ordering of single cells, Nat Biotechnol. 32 (2014) 381--386.
537 https://doi.org/10.1038/nbt.2859.

538 [46] R.C. Team, R: A Language and Environment for Statistical Computing, R Foundation for
539 Statistical Computing, Vienna, Austria, 2020. https://www.R-project.org.

540 [47] M.C. Canver, D.E. Bauer, T. Maeda, L. Pinello, DrugThatGene: integrative analysis to
541 streamline the identification of druggable genes, pathways and protein complexes from CRISPR
542 screens, Bioinformatics. 35 (2018) 1981–1984. https://doi.org/10.1093/bioinformatics/bty913.

543

544 **Figure Legends**

545 **Figure 1. Unbiased automatic cell labeling of human coronary scRNA data using**
546 **'SingleR' reveals abundant cells with chondrocyte (CH) and fibroblast (FB) gene**
547 **expression patterns, and pseudotemporal and ultra-fine clustering reveals derivation of**
548 **CH and FB-like cells from SMCs. (A)** UMAP clustering of 9798 cells derived from human
549 coronary artery explants with labels based on singleR annotation (Top) and cluster-based
550 manual annotation based on Wirka et al., 2019 (Bottom). **(B)** population breakdown by
551 percentage. SMC: smooth muscle cells, EC: endothelial cells, CH: chondrocytes, FB:
552 fibroblasts, Mø: macrophages, SC: stem cells, CMP: common myeloid progenitor cells, GMP:
553 granulocyte-monocyte progenitor cells. **(C) Left,** RNA trajectory (line) shows path two direct
554 paths from the SMC starting nodes (grey circles). **Middle,** ultra-fine clustering shows the logical
555 transition stages from SMC to CH and FB. **Right,** pseudotemporal analysis confirms that the
556 cell and clusters existing along a logical single-cell continuum. **(D)** selected genes that were
557 shown to vary over pseudotime by Moran's I test were visualized.
558
559 **Figure 2. Comprehensive ligand-receptor analysis coupled to drug-gene interaction**
560 **analysis reveals potential drug therapies to interrupt SMC to CH and FB communication**
561 **and SMC trajectory (A)** circle plot representation of the inferred intercellular communications
562 within the coronary artery environment. **(B)** incoming and outgoing signal summed weights by
563 cell type. **(C)** 'scTalk' shows that SMCs interacts with FB through three pairs of experimentally
564 verified interactions. Of these three, complement and *DCN-EGFR* signaling is the most
565 druggable as revealed by DGIdb 3.0. **(D)** druggable genome analysis revealed agents that can
566 target genes such as complement signaling (C3), C-X-C Motif Chemokine Ligand 12 (CXCL12),
567 and matrix metalloprotease (MMP2), which increases along the SMC-to-FB phenotype
568 trajectory.
569

570 **Figure 3. (A)** screenshot of the PlaqView web application. **(B)** a roadmap of enhanced scRNA-
571 seq analysis process. Instead of cluster-based grouping, our pipeline uses automatic cell
572 labeling, coupled with pseudotime trajectory, cellular network interaction and drug targeting, and
573 provides a reproducible process for scRNA datasets. From this roadmap, it is easy to add
574 additional analysis tools and modify workflow as more tools and datasets become available.
575

576 **Tables and Table Legends**

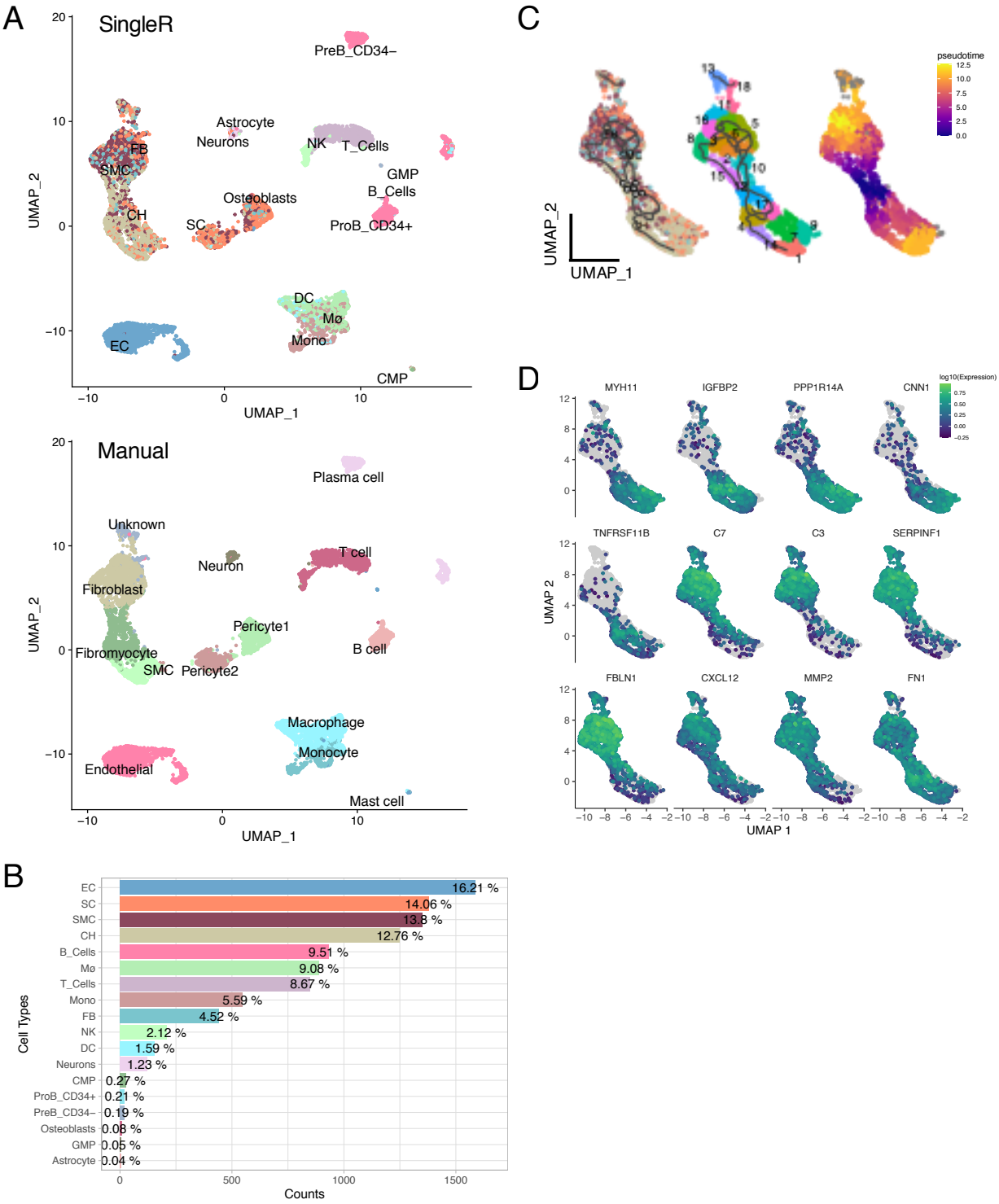
577 **Table 1.**

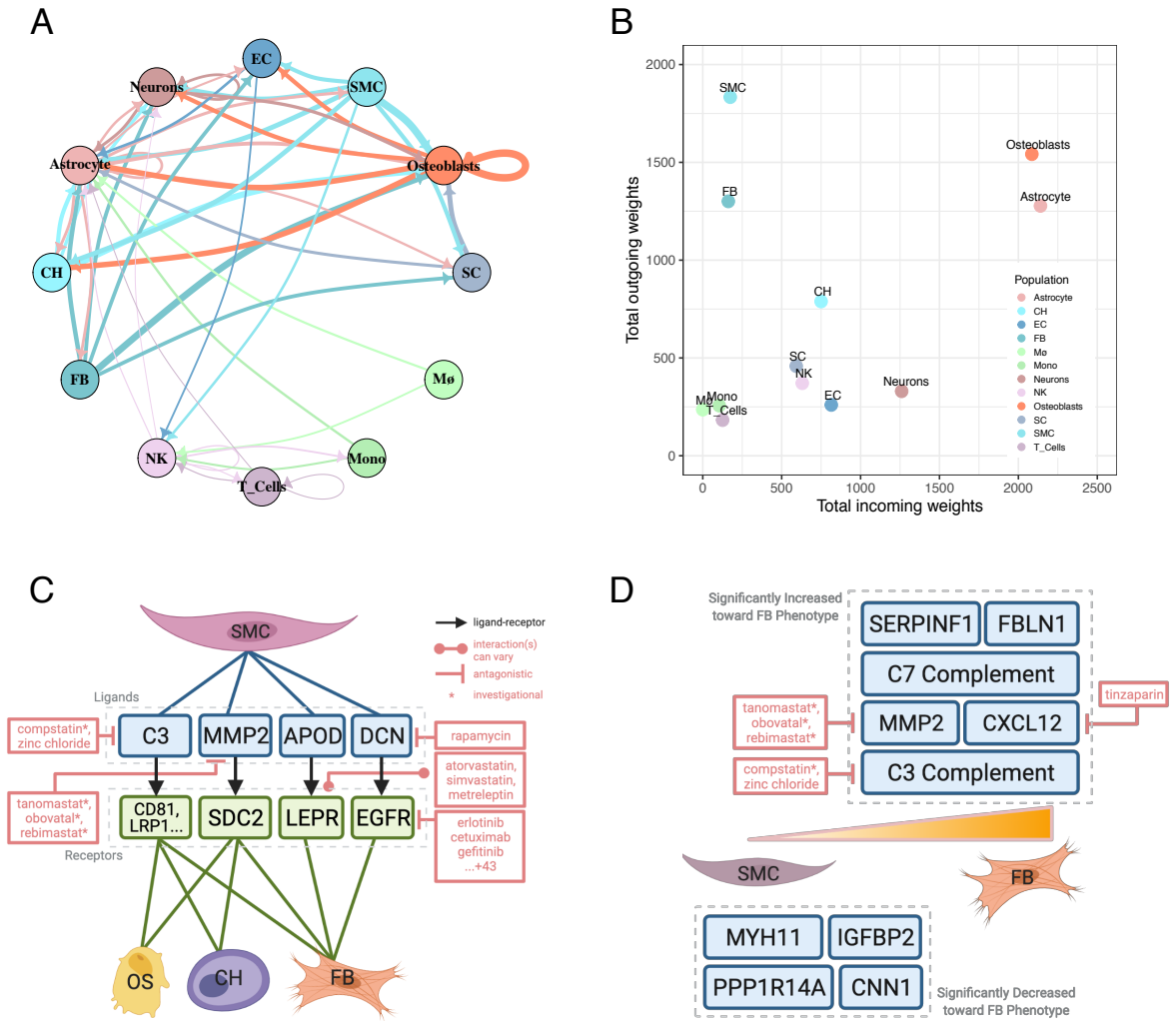
578

	p_value	morans_test_statistic	morans_I
MYH11	<1E-297	169.0285919	0.66505236
IGFBP2	<1E-297	129.2299117	0.50844307
PPP1R14A	<1E-297	169.4172149	0.66663148
CNN1	<1E-297	170.9288702	0.67251583
TNFRSF11B	<1E-297	97.66515083	0.38400709
C7	<1E-297	156.2248703	0.61472417
C3	<1E-297	103.4495	0.7252684
SERPINF1	<1E-297	157.9946249	0.62168635
FBLN1	<1E-297	154.3399713	0.60730135
CXCL12	<1E-298	48.92096	0.3428102
MMP2	<1E-299	62.57231	0.4385507
FN1	<1E-297	125.4354275	0.49354242

579 **Table 1. Selected Moran's I statistics for genes listed in Figure 1D.** Moran's I statistics was
580 used to identify spatial correlations within the single-cell trajectory. +1 indicates that nearby cells
581 are perfectly similar, 0 indicates no similarity or pattern, and -1 indicates total dissimilarity [13].

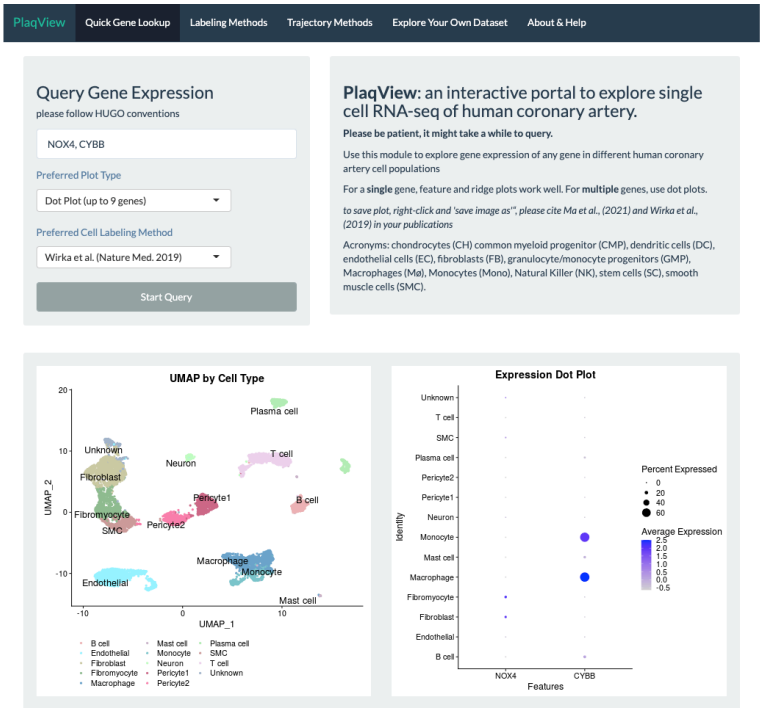
582


A
SingleR
PreB_CD34−
Astrocyte
Neurons
NK
T_Cells
FB
SMC
GMP
B_Cells
Osteoblasts
CH
SC
ProB_CD34+
DC
Mø
Mono
EC
CMP
UMAP_2
UMAP_1
Manual
Plasma cell
Unknown
T cell
Neuron
Fibroblast
Pericyte1
B cell
Fibromyocyte
SMC
Pericyte2
Macrophage
Monocyte
Endothelial
Mast cell
B
Cell Types
Counts
EC 16.21 %
SC 14.06 %
SMC 13.3 %
CH 12.76 %
B_Cells 9.51 %
Mø 9.08 %
T_Cells 8.67 %
Mono 5.59 %
FB 4.52 %
NK 2.12 %
DC 1.59 %
Neurons 1.23 %
CMP 0.27 %
ProB_CD34+ 0.21 %
PreB_CD34− 0.19 %
Osteoblasts 0.08 %
GMP 0.05 %
Astrocyte 0.04 %
C
pseudotime
UMAP_2
UMAP_1
D
log10(Expression)
MYH11
IGFBP2
PPP1R14A
CNN1
TNFRSF11B
C7
C3
SERPINF1
FBLN1
CXCL12
MMP2
FN1
UMAP 2
UMAP 1

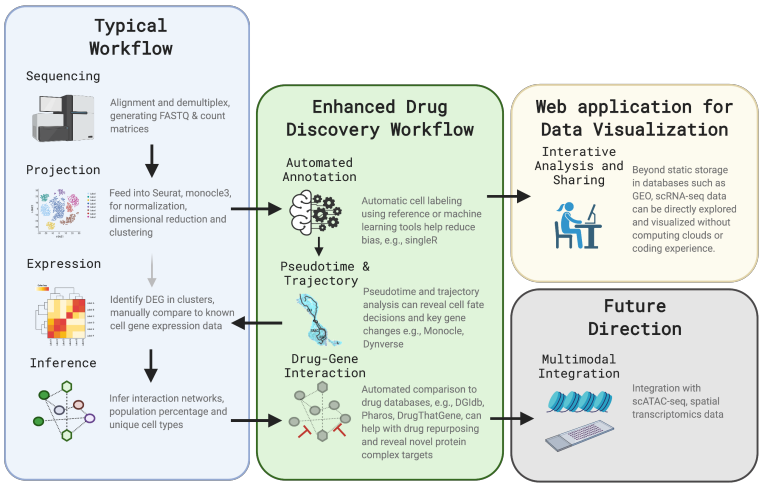

A
EC
Neurons
SMC
Astrocyte
Osteoblasts
CH
SC
FB
Mø
NK
Mono
T_Cells
B
Total outgoing weights
Total incoming weights
SMC
Osteoblasts
FB
Astrocyte
CH
SC
NK
Neurons
EC
Mø
Mono
T_Cells
Population
Astrocyte
CH
EC
FB
Mø
Mono
Neurons
NK
Osteoblasts
SC
SMC
T_Cells
C
SMC
ligand-receptor
interaction(s) can vary
antagonistic
investigational
Ligands
compstatin*, zinc chloride
C3
MMP2
APOD
DCN
rapamycin
atorvastatin, simvastatin, metreleptin
tanomastat*, obovatal*, rebimastat*
CD81, LRP1...
SDC2
LEPR
EGFR
erlotinib cetuximab gefitinib ...+43
Receptors
OS
CH
FB
D
Significantly Increased toward FB Phenotype
SERPINF1
FBLN1
C7 Complement
tinzaparin
tanomastat*, obovatal*, rebimastat*
MMP2
CXCL12
compstatin*, zinc chloride
C3 Complement
SMC
FB
MYH11
IGFBP2
PPP1R14A
CNN1
Significantly Decreased toward FB Phenotype

583

A


PlaqView
Quick Gene Lookup
Labeling Methods
Trajectory Methods
Explore Your Own Dataset
About & Help
Query Gene Expression
please follow HUGO conventions
NOX4, CYBB
Preferred Plot Type
Dot Plot (up to 9 genes)
Preferred Cell Labeling Method
Wirka et al. (Nature Med. 2019)
Start Query
PlaqView: an interactive portal to explore single cell RNA-seq of human coronary artery.
Please be patient, it might take a while to query.
Use this module to explore gene expression of any gene in different human coronary artery cell populations
For a single gene, feature and ridge plots work well. For multiple genes, use dot plots.
to save plot, right-click and 'save image as'', please cite Ma et al., (2021) and Wirka et al., (2019) in your publications
Acronyms: chondrocytes (CH) common myeloid progenitor (CMP), dendritic cells (DC), endothelial cells (EC), fibroblasts (FB), granulocyte/monocyte progenitors (GMP), Macrophages (Mø), Monocytes (Mono), Natural Killer (NK), stem cells (SC), smooth muscle cells (SMC).
UMAP by Cell Type
Plasma cell
Unknown
Neuron
T cell
Fibroblast
Pericyte1
B cell
Fibromyocyte
SMC
Pericyte2
Macrophage
Monocyte
Endothelial
Mast cell
UMAP_1
UMAP_2
Expression Dot Plot
Unknown
T cell
SMC
Plasma cell
Pericyte2
Pericyte1
Neuron
Monocyte
Mast cell
Macrophage
Fibromyocyte
Fibroblast
Endothelial
B cell
Identity
NOX4
CYBB
Features
Percent Expressed
Average Expression

B


Typical Workflow
Sequencing
Alignment and demultiplex, generating FASTQ & count matrices
Projection
Feed into Seurat, monocle3, for normalization, dimensional reduction and clustering
Expression
Identify DEG in clusters, manually compare to known cell gene expression data
Inference
Infer interaction networks, population percentage and unique cell types
Enhanced Drug Discovery Workflow
Automated Annotation
Automatic cell labeling using reference or machine learning tools help reduce bias, e.g., singleR
Pseudotime & Trajectory
Pseudotime and trajectory analysis can reveal cell fate decisions and key gene changes e.g., Monocle, Dynverse
Drug-Gene Interaction
Automated comparison to drug databases, e.g., DGIdb, Pharos, DrugThatGene, can help with drug repurposing and reveal novel protein complex targets
Web application for Data Visualization
Interative Analysis and Sharing
Beyond static storage in databases such as GEO, scRNA-seq data can be directly explored and visualized without computing clouds or coding experience.
Future Direction
Multimodal Integration
Integration with scATAC-seq, spatial transcriptomics data

584