Cysteine-Rich Intestinal Protein 1 is a Novel Surface Marker for Myometrial Stem/Progenitor Cells.

Emmanuel N. Paul[1], Tyler J. Carpenter[1], Sarah Fitch[1,2], Rachael Sheridan[3], Kin H. Lau[3], Ripla Arora[1,2], Jose M. Teixeira[1*]

[1]Department of Obstetrics, Gynecology and Reproductive Biology, College of Human Medicine, Michigan State University, Grand Rapids, MI 48824, USA
[2]Institute for Quantitative Health Science and Engineering, East Lansing, MI 48824, USA
[3]Van Andel Research Institute, Grand Rapids, MI, USA.

***Corresponding Author:**
Jose M. Teixeira, PhD, Michigan State University, 400 Monroe Ave NW, Grand Rapids, MI 49503. Email: teixei15@msu.edu

Acknowledgments:

The authors have no conflicts to declare.

Summary

Uterine fibroids are benign tumors that develop in the myometria of most women with unknown etiology. Myometrial stem/progenitor cells (MyoSPCs) have been proposed as the cells of origin for fibroids because uterine fibroids are almost always clonal. We previously identified SUSD2 as a possible MyoSPC marker, but the relatively poor enrichment in stem cell characteristics of SUSD2+ over SUSD2- cells compelled us to find better discerning markers for more rigorous analysis. We combined bulk RNA-seq of SUSD2+/- cells with single cell RNA-seq to identify markers capable of further enriching for MyoSPCs. We observed seven distinct cell clusters within the myometrium, with the vascular myocyte cluster most highly enriched for MyoSPC characteristics and markers, including *SUSD2*. *CRIP1* expression was found highly upregulated in both techniques and was used as a marker to sort CRIP1^{+}/PECAM1^{-} cells that were enriched for colony forming units and were able to differentiate into mesenchymal lineages, suggesting that CRIP1^{+}/PECAM1^{-} cells could be used to better understand the etiology of uterine fibroids.

36 **Introduction**

37 Uterine fibroids, also known as leiomyomas, are benign tumors found in the
38 smooth muscle layer of the uterus, the myometrium. Although benign, uterine fibroids
39 develop in up to 80% of women during their reproductive years and are often associated
40 with debilitating symptoms such as menorrhagia, anemia, dysmenorrhea, fatigue, pelvic
41 pain, urinary incontinence, and obstruction of the surrounding pelvic structures (1,2).
42 Hormonal therapies, mainly used for alleviating fibroid symptoms, are generally short-
43 term treatments due to long-term side effects or induced infertility (3), and hysterectomy,
44 the most common and effective treatment for uterine fibroids, results in permanent
45 infertility (4). Therefore, there is a clear need for effective medical therapies that preserve
46 fertility and avoid invasive surgery. Understanding the disease etiology and pathogenesis
47 could lead to alternative and less invasive therapies.

48 A dysregulated myometrial stem/progenitor cell (MyoSPC) has been proposed as
49 the cell of origin for uterine fibroids. After embryonic development, tissue-specific stem
50 cells remain throughout the body and play important roles in tissue homeostasis, including
51 replacing dying cells and participating in tissue remodeling (5). The dramatic remodeling
52 that occurs during pregnancy and following parturition in the uterus suggest a need for
53 and the existence of myometrial stem cells (6). Uterine fibroids are thought to be a clonal
54 disease (7-9), and since most clonal diseases have a single cell origin (10), we and others
55 (11,12) have hypothesized that a mutated MyoSPC could be the cell of origin for uterine
56 fibroids (13). The presence of cells with stem cell properties has been demonstrated using
57 the label-retaining cells in a mouse model (14,15) and using the side population (SP)
58 method in human myometrium (16,17). Putative MyoSPCs have been isolated and

59 studied by using a combination of cell surface markers, including CD44/Stro-1 (11) and
60 CD34/CD49f/b (18), or using the side population (SP) discrimination assay (16,17).
61 However, cell surface markers in these studies have been selected using stem cell
62 markers from other tissues or using the SP technique, which has multiple pitfalls. Indeed,
63 the SP assay is fallible to variation in staining conditions, as well as a heterogeneity of
64 the isolated cells, indicating that the SP phenotype is not exclusive to stem cells, and
65 more importantly, the dye used in SP assay has profound effects on cell survival (19).
66 Because the endometrial stroma and the myometrium originate from the same embryonic
67 tissue, the Müllerian duct mesenchyme (6), we recently proposed that SUSD2, an
68 endometrium stem cell marker (20) also enriches for MyoSPCs (21). While SUSD2+ cells
69 do have mesenchymal stem cell characteristics, SUSD2+ cells represent between 25-
70 40% of total myometrial cells. Additionally, colony formation is only increased 2-3 fold
71 increase in SUSD2^{+} cells compared to the rest of the myometrium cells, suggesting that
72 further enrichment might be possible. The objective of the present study was to integrate
73 next-generation sequencing, including single cell RNA-seq and bulk RNA-seq, to identify
74 a more specific marker to significantly enrich for MyoSPCs from human myometrium,
75 which can then be used to better understand the molecular mechanisms underlying fibroid
76 etiology.
77

Results

SUSD2$^+$ are enriched for characteristic MSC genes compared with SUSD2$^-$ cells.

To determine how best to enrich for stem cell activities in the SUSD2+ MyoSPC population, we used SUSD2 to enrich for myometrium stem cells followed by RNA-seq to discover new cell surface markers for MyoSPCs in the human myometrium. Myometrial cells from non-fibroid patients (n = 5) were isolated and live SUSD2$^+$ and SUSD2$^-$ were sorted by flow cytometry. As with our previous results (21), 30-50% of the myometrium cells were SUSD2$^+$ (**Fig 1A**). Total RNA was isolated from the two cell populations and sequenced for differential gene expression analyses. Principal Component Analysis (PCA) plot showed that SUSD2$^+$ and SUSD2$^-$ cells were separated by principal component 1 with a variance of 39%, indicating a strong divergence in the transcriptomic profiles of these two cell populations (**Fig 1B**). A total of 6777 significant differentially expressed genes (DEGs) were detected between SUSD2$^+$ and SUSD2$^-$ myometrial cells with a false discovery rate (FDR) <0.05) (**Fig 1C and Table S1**). 3527 genes were down-regulated and 3250 were up-regulated in the SUSD2$^+$ population compared to the SUSD2$^-$ population. We confirmed that *SUSD2* was up-regulated in the SUSD2$^+$ sorted cells and were also enriched in other MSC markers such as *MCAM*, *PDGFRβ* and *CSPG4* (**Fig 1C and 1D**). A heatmap of the top 300 most DEGs from the SUSD2$^+$ compared to SUSD2$^-$ cells showed a good separation between cell types and included *SUSD2*, *MCAM*, *PDGFRβ* and *CSPG4* (**Fig 1E**).

Myometrium side population cells are not enriched in MSC markers.

100 The side population (SP) phenotype is another often used method to isolate cells with
101 stem cell characteristics that exploits the ability of some stem cells to efflux the DNA-
102 binding dye Hoechst 33342 via the ATP-binding cassette (ABC) transporters
103 (11,16,17,22,23). An average of 1.7% of the total myometrium cells were SP^+ (**Fig 2A**).
104 Addition of verapamil, a calcium channel blocker used as a negative control to validate
105 the SP, severely decreased of the number of the myometrium SP^+ cells (**Fig 2B**). SP^+
106 and SP^- myometrium cells were sorted for total RNA sequencing and analyzed by PCA
107 plot, which showed that matched SP^+ and SP^- cells segregated by the principal
108 component 2, accounting for 25% of the variance (**Fig 2C**). A total of 828 significant (FDR
109 <0.05) DEGs, including 478 upregulated genes and 350 downregulated genes, were
110 detected between the SP^+ and SP^- myometrium cells (**Table S2**). The top 10 DEGs
111 enriched in the SP^+ compared to the SP^- were associated with immune response (*XCL2*
112 *(24)*, *CD69 (25)*, *IL7R (26)*, *KLRD1 (27)*, and *IL18R1 (28)*) apoptosis (*TNFRSF10A (29)*),
113 extracellular matrix (*SPOCK2 (30)*) and hematopoietic stem cell (*SELE (31)*, *GATA3 (32)*,
114 *CD69 (33)*, and *VCAM1(34)*) (**Fig. 2D**). We confirmed an increase in expression of two
115 major ABC transporters, including *ABCB1*, and *ABCG2*, and a decrease in *PGR*, another
116 marker previously shown downregulated in the SP^+ compared to the SP^- of human
117 myometrium cells (16) (**Fig 2E**). Surprisingly, SP^+ cells did not show increased expression
118 of the MSC markers, *SUSD2*, *MCAM*, *PDGFRβ*, *CSPG4*, *CD44*, *CD34* and *ITGA6* (also
119 known as *CD49f*) compared to the SP^- (**Fig 2F**).
120

121 **A putative MyoSPC cluster is determined by single cell RNA-seq.**

122 Uniform Manifold Approximation and Projection (UMAP) of myometrial (n = 5) single cell
123 RNA-seq (scRNA-seq) revealed 7 main cell clusters (**Fig 3A**) with an equal cell
124 distribution across the five myometrium samples (**Fig S1**). Cluster identities were
125 assigned using the expression profiles of canonical markers for cell populations expected
126 to be found in the myometrium (**Fig 3B**) (35,36), including 4 different smooth muscle cell
127 types, vascular myocytes, myocytes, myofibroblasts, and fibroblast. The 4 MSC markers,
128 *SUSD2*, *MCAM*, *PDGFRβ* and *CSPG4*, were found highly expressed in the vascular
129 myocyte cluster, a known stem cell niche (**Fig 3C and 3D**). Known MSC characteristics
130 such as quiescent cell cycle state (G0) and the low gene regulation dynamics (37-39)
131 were determined by the cell cycle score and the velocity of the scRNA-seq data,
132 respectively. We identified a small group of cells within the vascular myocyte cluster in
133 the G1/G0 phase (**Fig 4A**) using a computational assignment of cell-cycle stage, as
134 described by Scialdone *et al.* (40). Cell velocity, which predicts the future state of
135 individual cells using the RNA splicing information from each cell (41), showed that the
136 same group of cells in G1/G0 phase in the vascular myocyte cluster are depicted with low
137 velocity vectors (**Fig 4B**) indicating low levels of transcriptional changes, another
138 characteristic of stem cells.

139

140 **Combining bulk SUSD2 RNA-seq and myometrial scRNA-seq reveals a new**
141 **MyoSPC marker (CRIP1).**

142 We defined cells within the vascular myocyte cluster presenting with high expression of
143 MSC markers in a G1/G0 phase and with low velocity as the “MyoSPC” cluster. A total of

144 6700 DEGs were found in the MyoSPC cluster compared to the rest of the myometrial
145 cells (**Table S3**). 54% (1929 DEGs) of the MyoSPC DEGs overlapped with the bulk RNA-
146 seq of SUSD2^{+} compared to SUSD2^{-} (**Fig 5A and Table S4**). Correlation analysis of the
147 log_2FC gene expression in the scRNA-seq analysis of the MyoSPC cluster compared to
148 the rest of the myometrial cells and the bulk RNA-seq of SUSD2+ compared with SUSD2-
149 cells by a scatter plot confirmed that the MSC markers, *SUSD2*, *MCAM*, *PDGFRβ* and
150 *CSPG4* were upregulated in both. The most highly upregulated gene in the MyoSPC
151 cluster, Cysteine-Rich Intestinal Protein 1 (*CRIP1*), is also significantly upregulated in the
152 SUSD2+ cells (**Fig 5B**). UMAP plot showed that CRIP1 was highly expressed in the
153 vascular myocyte cluster, which includes MyoSPC cluster (**Fig 5C**), more particularly in
154 the defined MyoSPC cluster (**Fig 5D**). *CRIP1* expression wasn't differentially expressed
155 (log_2FC= -0.2, FDR p= 9.9×10^{-1}) in the RNA-seq results of the SP assay (**Fig S2**).
156

157 **CRIP1+ cells have common stem/progenitor cell properties**

158 We next investigated the CRIP1+ cells to establish their stem cell bona fides.
159 Immunofluorescence analysis using 3D imaging of the myometrial layer showed that
160 CRIP1+ cells are located surrounding the vascular endothelial cells (PECAM1^{+}, also
161 known as CD31), a known mesenchymal stem cell niche (42-44) (**Fig 6A**). Interestingly,
162 CRIP1^{+} cell immunofluorescence was localized mostly near the larger blood vessels and
163 within a subset of SUSD2^{+} cells. Flow cytometry revealed that CRIP1^{+} cells represented
164 between 2 to 5% of the total myometrium cells (**Fig 6B**), a stem cell ratio expected to be
165 found in adult tissues (45). PECAM1 was used for negative selection of the smaller
166 population of endothelial cells that co-expressed CRIP1. CRIP1^{+}/PECAM1^{-} cells and the

167 depleted cell population were sorted, and typical downstream stem cell assays were
168 performed to determine if the CRIP1^{+}/PECAM1^{-} cells have stem/progenitor cell
169 proprieties. Colony formation assays indicated that CRIP1^{+}/PECAM1^{-} sorted cells have a
170 greater self-renewal capacity compared to the depleted sorted population (**Fig 6C**), with
171 a significant increase of 4.5-fold greater number of colonies formed (**Fig 6D and 6E**), as
172 well as a significant increase of the size of the colonies (**Fig 6F**). After 5 days in smooth
173 muscle differentiation media, CRIP1^{+}/PECAM1^{-} cells were positive for ACTA2, indicating
174 that they differentiated into smooth muscle cells (**Fig 6G**). What about depleted cells?
175 Similarly, CRIP1^{+}/PECAM1^{-} cells were positive for the Oil Red O staining (**Fig 6H**), and
176 the alkaline phosphatase activity (**Fig 6I**) when grown in either in adipogenic or
177 osteogenic differentiation media, respectively, compared to CRIP1^{+}/PECAM1^{-} cells
178 grown in control media, indicating that these putative MyoSPC cells have the capacity to
179 differentiate into adipocytes and osteocytes.

180 **Discussion**

181 We have identified CRIP1 as a novel cell surface marker that enriches for a
182 possible MyoSPC by combining the analyses of two next generation sequencing
183 techniques, bulk RNA-seq from SUSD2^{+} and SUSD2^{-} cells and scRNA-seq of total
184 myometrial samples. Bulk RNA-seq, enriched for known MSC markers, but the number
185 of DEGs were many, making it difficult to choose, identify and validate any possible
186 myometrium stem cell marker for further study. To reduce the number of candidate
187 markers for validation and follow up studies, we used the scRNA-seq to identify possible
188 stem cells based on MSC markers. Subsequent stem cell assays confirmed that CRIP1+
189 cells have MSC properties, and further studies are underway to determine whether these
190 cells could be a cell of origin for uterine fibroids.

191 *CRIP1* was a highly upregulated gene in both SUSD2+ cells and the MyoSPC cluster we
192 identified by scRNA-seq. *CRIP1*, Cysteine-rich intestinal protein 1, is a member of the
193 LIM/double zinc-finger proteins that is predicted to be a novel biomarker in multiple
194 cancers and can promote several biological processes, including cell migration, invasion
195 and epithelial-mesenchymal transition by activating the Wnt/β-catenin signaling an
196 essential pathway in stem cell homeostasis (46-48). *CRIP1* was upregulated in the SP of
197 fibroid tissues along with 153 other DEGs by microarray analysis (49). In that study, the
198 authors used one of the stem cells characteristics, the increased expression ad activity
199 of ABC transporters, to isolate SP^{+} cells from fibroid tissues. They demonstrated that the
200 SP^{+} cells from fibroid tissues have stem cell characteristics, including self-renewal, and
201 differentiation into adipose and osteocyte cells. We performed similar experiments using
202 myometrial samples from non-fibroid patients (MyoN) and also found an increase in the

203 expression of major ABC transporters in the SP^+ compared to the SP^- of human
204 myometrium cells We used MyoN samples because we recently reported that
205 myometrium from fibroid patients (MyoF) have a different transcriptomic profile compared
206 to MyoN samples, including an enrichment of DEGs in a leiomyoma disease ontology
207 panel (50). Here we have reported that *CRIP1* expression was not differentially expressed
208 and that the expression of the MSC markers, *SUSD2*, *MCAM*, *PDGFRβ*, and *CSPG4*
209 were decreased in SP^+ compared to the SP^- from MyoN cells. The discordant results from
210 these studies could arise from the tissue type, that is, fibroid tumor versus MyoN, or
211 because the SP technique could be more applicable to other tissues (19) or hematopoietic
212 stem cells (HSC). It is worth noting that HSC exhibit a specific ABC transporter gene
213 expression profile distinct from other stem cells, including MSC (51). HSC expressed
214 higher level of most of the ABC transporters including *ABCB1*, *ABCC1* and *ABCG2*
215 compared to other stem cells. Indeed, the SP+ MyoN cells were enriched for ABC
216 transporters and HSC-associated genes. Additionally, the SP technique relies on an intact
217 cell metabolism and considerable variation in results is observed (19).

218 Our scRNA-seq results suggested that human myometrium has at least 7 different
219 cell types, including different types of smooth muscle cells, endothelial cells, and immune
220 cells. Similar clusters were reported in a scRNA-seq comparison of fibroids and
221 myometrium (36). The depth of sequencing for each cell was close to saturation allowing
222 us to identify a small cell population with stem cell characteristics, including the
223 expression of MSC markers *SUSD2*, *MCAM*, *PDGFRβ*, and *CSPG4*, a quiescent (G0)
224 cell cycle state (37) and low transcriptomic activity/low RNA velocity (37-39).
225 $CRIP1^+$/$PECAM1^-$ cells were primarily located in the perivascular region, a known stem

226 cell niche (42-44), particularly by the larger myometrial blood vessel. These results were
227 consistent with our scRNA-seq results showing that *CRIP1* expression was highly
228 expressed in the vascular myocyte cluster. Moreover, immunofluorescence staining
229 showed that CRIP1^{+} cells were a subset the SUSD2^{+} cells and that CRIP1^{+}/PECAM1^{-}
230 cells account for only 2 to 5% of the total human myometrium cells, a stem cell proportion
231 expected to be in an adult tissue (45). Interestingly, the depleted population was able to
232 form a few colonies, an indication that some cells in the depleted population still have a
233 self-renewal properties. Similar results were observed by our group using other MSC
234 markers to isolate myometrium stem/progenitor cells, including SUSD2, MCAM, or
235 PDGFRβ (21). This finding suggests that further enrichment of the MyoSPCs with some
236 of the other MSC markers could be done.
237

238 **Study approval**

239 The use of human tissue specimens was approved by the Spectrum Health Systems and
240 Michigan State University Institutional Review Boards (MSU IRB Study ID:
241 STUDY00003101, SR IRB #2017-198) as secondary use of biobank materials.
242

243 **Author contributions**

244 Experimental design (E.N.P, J.M.T) collected data and performed experiments (E.N.P,
245 T.J.C, S.F), analyzed data (E.N.P, R.S, K.H.L, R.A, J.M.T), wrote/reviewed manuscript
246 (E.N.P, T.J.C, S.F, R.S, K.H.L, R.A, J.M.T).
247

248 **Acknowledgments**

249 We would like to thank the patients who consented for the study, the Spectrum Health
250 Systems Universal Biorepository staff, and the Van Andel Institute Genomics Core
251 (RRID:SCR_022913), especially Marie Adams, Rebecca Siwicki and Julie Koeman, for
252 their assistance with the construction and sequencing of 10X libraries for the single cell
253 RNA-seq and bulk RNA-seq.This work was supported by grants HD096259 and
254 HD100959 from the Eunice Kennedy Shriver National Institute of Child Health and Human
255 Development (to J.M.T) and the SRI/Bayer Discovery/Innovation Grant (to E.N.P).
256

257 **Methods**
258
259 **Sample collection and cell isolation**

260 The use of human tissue specimens was approved by the Spectrum Health Systems
261 Institutional Review Board as secondary use of biobank materials. Myometrial samples
262 from non-fibroid patients (MyoN) were obtained following total hysterectomy from pre-
263 menopausal, self-identified Caucasian women. No fibroids were detected by ultrasound
264 prior to surgery. All patients who participated to the study gave consent to donate tissue
265 through the Spectrum Health Biorepository. Tissues were washed with phosphate-
266 buffered saline, dissected away from non-myometrial tissue, and minced. Cells were
267 isolated in baffled flask with digestion media (DMEM/F12, Thermo Fisher, #SH3027101)
268 containing 10% fetal bovine serum (FBS), collagenase type I (Sigma, #C0130), DNAse
269 type I (Sigma, #10104159001) and $MgCl_2$ and incubated with agitation at 37°C. The
270 resulting cell suspensions were strained through 100- and 40-µm cell strainers, washed
271 with warm media (DMEM/F12) containing 15% FBS, and centrifuged. Isolated cells were
272 then stored in freeze media (90% FBS, 10% DMSO, Thermo Fisher, #BP231-100) at -80
273 °C until needed.
274

275 **Cell staining for FACS**

276 Human primary myometrium cells were thawed and resuspended in 1% bovine serum
277 albumin blocking buffer for 20 min at room temperature (RT). Cells were then incubated
278 with the primary antibody for 45 min at RT; SUSD2-PE anti-human (Miltenyi Biotec, #130-
279 117-682), PECAM1-FITC anti-human (Thermo Fisher, #11-0319-42), and CRIP1 rabbit
280 anti-human (Thermo Fisher, #PA5-24643). For CRIP1/PECAM1 staining, cells were

281 incubated with a secondary antibody for 30 min at RT using an Alexa-647 anti-rabbit
282 (Thermo Fisher, #A21244). Stained myometrium cells were then wash with flow buffer,
283 and resuspend in 1 mL of flow buffer with 1 µg of 4′,6-diamidino-2-phenylindole (DAPI) or
284 Propidium Iodode (PI), depending on the experiment, for live dead discrimination. Cells
285 were sorted by the flow cytometry core at Van Andel Research Institute (VARI) using a
286 FACSymphony S6 cytometer (BD Biosciences) and analyzed with FlowJo Software (BD
287 Biosciences, version 10.8.1). Cells were sorted in media (DMEM/F12) at 4C.
288
289 The side population assay was conducted as described previously (17). Briefly, live cells
290 were labeled in a water bath at 37 °C for 90 min with 5 ug/mL of Hoechst 33342 dye
291 (Thermo Fisher). As a negative control, separate aliquots of cells from the same patients
292 were treated with 25 mg/ml of verapamil (Sigma) prior to addition of the Hoechst dye. PI
293 was added to stained cells with and without verapamil treatment and analyzed in a MoFlo
294 Astrios (Beckman Coulter) for side population gating by Hoechst red and blue filters and
295 sorting in media (DMEM/F12) at 4 °C.
296
297 **RNA Isolation, Library Preparation and Sequencing**
298 Sorted cells (SUSD2$^+$, SUSD2$^-$, SP$^+$ and SP$^-$) were lysed and isolated using RNeasy mini
299 kit (Qiagen, #74104). Isolated RNA was stored at −80 °C in nuclease-free water.
300 Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Fair-lawn, NJ, USA) and
301 Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) instruments were
302 used to measure RNA concentration and quality, according to the manufacturers'
303 protocols. RNA integrity values of ≥7.5 were required for library preparation and paired-

304 end (2 × 100 bp) RNA-sequencing on an Illumina NextSeq 6000 instrument (Illumina Inc.,
305 San Diego, CA, USA). Libraries were prepared using a Kapa RNA HyperPrep kit with
306 ribosomal reduction, pooled, and sequenced on flowcells to yield approximately 50–60
307 million reads/sample.
308
309 For single cell RNA-seq, dead cells from isolated myometrium cells were removed using
310 the Dead Cell Removal Kit (Miltenyi Biotec, #130-090-101) per manufacturer's
311 instructions. Live myometrium cells from 5 non-fibroid patients were then sequenced.
312 Libraries were generated and sequenced using the 10X Chromium Next GEM Single Cell
313 3' GEM kit (10X Genomics, v2) platform according to the manufacturer's instruction to
314 target an output of 7,000 cells per sample. Quality and quantity of the finished libraries
315 were assessed using a combination of Agilent DNA High Sensitivity chip (Agilent
316 Technologies, Inc.) and QuantiFluor® dsDNA System (Promega Corp., Madison, WI,
317 USA). 2 x 75 bp, paired end sequencing was performed on an Illumina NovaSeq 6000
318 sequencer using an S2, 100 cycle sequencing kit (v1.5) to a minimum depth of 50K reads
319 per cell (Illumina Inc., San Diego, CA, USA). Base calling was done by Illumina RTA3
320 and output was demultiplexed and converted to FastQ format with Cell Ranger (10X
321 Genomics, v3.1.0).
322
323 **RNA Seq Analysis**
324 For bulk RNA-seq, reads from 5 MyoN SUSD2+ and SUSD2+ cells and 3 MyoN SP^+ and
325 SP^- cells, were trimmed for quality and adapters using TrimGalore (version 0.6.5), and
326 quality trimmed reads were assessed with FastQC (version 0.11.7). Trimmed reads were

327 mapped to Homo sapiens genome assembly GRCh38 (hg38) using STAR (version
328 2.7.9a) (72). Reads overlapping Ensembl annotations (version 99) were quantified with
329 STAR prior to model-based differential expression analysis using the edgeR-robust
330 method with paired samples. Genes with low counts per million (CPM) were removed
331 using the filterByExpr function from edgeR (72). Scatterplots of two selected principal
332 components was constructed with the pca function of the PCAtools R package (version
333 2.5.13) to verify sample separation prior to statistical testing. Generalized linear models
334 were used to determine if principal components were significantly associated with cell
335 type. Genes were considered differentially expressed if their respective edgeR-robust
336 false discovery rates (FDR) corrected p-values were less than 0.05. Differential
337 expression was calculated by comparing $SUSD2^+$ versus $SUSD2^-$ cells. Differentially
338 expressed genes (DEG) were visualized with volcano plots and heatmaps generated
339 using the EnhancedVolcano (version 1.8.0) and pheatmap (version 1.0.12) packages in
340 R. Box plots of the $\log_2$(CPM) values were generated using the R package ggplot2
341 (version 3.4.0).
342
343 A total of 9,775 total cells sent for sc-RNA-seq from 5 MyoN samples with ~512,000 reads
344 per cell passed the quality control. Demultiplexed raw sequencing reads were processed
345 and aligned to the Homo sapiens genome assembly GRCh38 (hg38) using STAR (version
346 2.7.9a) by the 10X Genomics Cell Ranger (version 3.1.0). Samples were merged using
347 the integration anchors function of the Seurat package (version 4.2.1) from R. Genes
348 expressed in fewer than three cells in a sample were excluded, as well as cells that
349 expressed fewer than 200 genes and mitochondrial gene content >5% of the total unique

350 molecular identifier count. Data were normalized using a global-scaling normalization
351 method that normalizes the feature expression measurements for each cell by the total
352 expression, multiplies this by a scale factor (10,000), and then log-transforms the results.
353 The top 2,000 most variable genes that were used for cell clustering were found using the
354 *FindVariableFeatures* function and were then normalized using the *ScaleData* function.
355 To determine the dimensionality of the dataset, an elbow plot was generated using the
356 *Elbowplot* function of Seurat. Based on the Elbow plot, we selected 15 principal
357 components (PC) for downstream analyses. Cell clusters were generated using
358 *FindNeighbors* and *FindClusters* functions. For visualization UMAPs were generated
359 using the *RunUMAP*, *FeaturePlot* and *DimPlot* fuctions. The *DotPlot* Seurat fuction was
360 used to generate dot plots to visualize gene expression for each assigned cluster. Cell
361 cycle score and velocity were determined using the functions *CellCycleScoring* from
362 Seurat and *RunVelocity* from SeuratWrappers (version 0.3.0), respectively. The "stem
363 cell" cluster was selected using the *CellSelector* function from Seurat. *RidgePlot* function
364 from the Seurat R package was use for the visualization of *CRIP1* gene expression in the
365 different myometrium clusters.
366
367 Venn diagrams of the overlapping DEG from the Bulk RNA-seq of SUSD2^{+} and SUSD2^{-}
368 cells and DEG of scRNA-seq of the "stem cell" cluster vs the rest of the myometrium cells
369 were constructed using the eulerr package (version 6.1.1). The scatter plot of overlapping
370 DEG from bulk- and scRNA-seq was generated with ggplot2 (version 3.4.0).
371
372 **Imaging**

373 Whole mount immunofluorescent staining (52) was performed for human myometrial
374 samples that were fixed in a 4:1 solution of Methanol:DMSO and stored at -20 °C. Tissue
375 was removed from the fixative and rehydrated in (1:1) methanol: PBST (PBS, 1% triton)
376 solution for 15min followed by a 15min wash in 100% PBST. Samples were incubated in
377 a blocking solution (PBS, 2% powdered milk, 1% triton) for 1 h at room temperature and
378 then stained with a 1:500 dilution for primary antibodies in blocking solution for 7 nights
379 at 4°C. Primary antibodies used were Rabbit anti-human CRIP1 (Thermofisher, #PA5-
380 24643), Mouse anti-human SUSD2 (Biolegend #327401) and Mouse anti-human
381 PECAM1 (Abcam, #ab9498). Samples were then incubated with secondary antibodies,
382 Donkey anti-Rabbit IgG Alexa Fluor 555 (Invitrogen, #A31572), Goat anti-Mouse IgG
383 Alexa Fluor 647 (Invitrogen, #A21235) at a dilution of 1:300 and Hoechst dye (Sigma-
384 Aldrich, #B2261), for three nights at 4°C. Samples were transferred to a Methanol:PBST
385 (1:1) solution for 15 minutes, washed in 100% Methanol for 30 minutes, and incubated at
386 4°C overnight in a 3% H_2O_2 solution (Sigma-Aldrich, #216763) diluted in Methanol.
387 Tissues were then washed in 100% Methanol for 30 minutes and then cleared in Benzyl
388 Alcohol:Benzyl Benzoate (1:2) (Sigma-Aldrich, #108006, #B6630) overnight. Imaging
389 was performed using a Leica SP8 TCS white light laser confocal microscope utilizing 10x
390 air objective and a 7.0 μm Z stack (53). Imaris v9.2.1 (Bitplane) commercial software was
391 used to analyze confocal image files and create 3D renderings. The image files were
392 imported into Imaris 3D surpass mode and 3D renderings were created using the Surface
393 plugin. Images were captured using the Snapshot plugin of Imaris and videos were
394 generated using the Animation plugin.
395
396 **Colony formation**

397 After isolation from human myometrium tissues of CRIP1^{+}/PECAM1^{-} cells and the
398 depleted population, cells were plated at colony forming density (50 cells/cm^{2}) in triplicate
399 in growth media (DMEM/F12, 10% FBS, 1× antibiotics/antimycotics) overnight and then
400 grown in MesenPro RS (Thermo Fisher, # 12746012) for 2 to 3 weeks. Cultures were
401 fixed in 4% paraformaldehyde (PFA) and stained with crystal violet for colony
402 visualization. Colonies were counted and total surface area was estimated using ImageJ
403 (version 1.53k), and the percent colony-forming units (CFUs) was calculated as (number
404 of colonies/number of cells plated) × 100 and averaged for triplicates. Matched
405 CRIP1^{+}/PECAM1^{-} and depleted populations were cultured in different wells of the same
406 plate, and therefore culture was stopped for both cell types on the same day. Images
407 were taken using a Nikon SMZ18 microscope and Ds-Ri1 camera (Nikon Instruments
408 Inc.).
409
410 **Mesenchymal lineage differentiation**
411 For osteogenic and adipogenic differentiation, CRIP1^{+}/PECAM1^{-} cells were plated in
412 triplicate at 50% to 80% confluency in growth media (DMEM/F12, 10% FBS, 1×
413 antibiotics/antimycotics) overnight and then cultured for 10 days in fresh StemPro
414 Adipogenesis Differentiation (Thermo Fisher Scientific, #A1007001) or StemPro
415 Osteogenesis Differentiation (Thermo Fisher Scientific, #A1007201) media according to
416 the manufacturer's instructions. Cells were cultured in regular growth media to serve as
417 differentiation controls. To assay adipogenic differentiation, cultures were fixed in 4%
418 PFA, stained using Oil Red O (Sigma, #01391) according to the manufacturer's
419 instructions. To assay osteogenic differentiation, cultures were stained for alkaline

420 phosphatase activity using the Alkaline Phosphatase (AP), Leukocyte kit (Sigma, #86R-
421 1KT) according to the manufacturer's instructions. For smooth muscle differentiation,
422 CRIP1^{+}/PECAM1^{-} cells were plated on 1 mg/ml dried rat tail collagen (Corning, #354 236)
423 8 well chamber slide with growth media (DMEM/F12, 10% FBS, 1×
424 antibiotics/antimycotics) overnight and then cultured in Medium 231 with a smooth muscle
425 differentiation supplement (Thermo Fisher, #M231500). Cells were fixed in 4% PFA at the
426 indicated times (D0: before adding the differentiation media, D5: 5 days in differentiation
427 media) and stained using αSMA-Cy3 (Sigma, #C6198). Images were taken using a Nikon
428 Eclipse Ni-U or Nikon SMZ18 microscope and Ds-Qi1MC or Ds-Ri1 camera (Nikon
429 Instruments Inc.).
430

431 **Statistical analyses**

432 Bioinformatic statistics were performed using the listed packages in R (version 4.0.2).
433 Differentially expressed genes of the bulk RNA-seq from SUSD2^{+} vs. SUSD2^{-} and the
434 SP^{+} vs. SP^{-} were identified as those having an Benjamini–Hochberg FDR corrected p
435 <0.05 (54). Data with unequal variances were log transformed, and homogeneity of
436 variances verified before completion of analyses. Differential expressed genes of the
437 single cell RNA-seq were calculated using the non-parametric Wilcoxon Rank Sum test.
438 Adjusted p-value, based on Bonferroni correction using all features in the dataset was
439 used to determine significance. Genes with adjusted p-value < 0.05 were consider as
440 significant. For the colony formation assay, comparison of two means was performed with
441 a two-sided student t test, and significance was determined at p <0.05 after confirming
442 normal distribution using Graphpad Prism (version 9.4.1).

443 **References**

444 1. Zimmermann A, Bernuit D, Gerlinger C, Schaefers M, Geppert K. Prevalence, symptoms
445 and management of uterine fibroids: an international internet-based survey of 21,746
446 women. *BMC Womens Health*. 2012;12:6.
447 2. Gupta S, Jose J, Manyonda I. Clinical presentation of fibroids. *Best Pract Res Clin Obstet*
448 *Gynaecol*. 2008;22(4):615-626.
449 3. Farris M, Bastianelli C, Rosato E, Brosens I, Benagiano G. Uterine fibroids: an update on
450 current and emerging medical treatment options. *Ther Clin Risk Manag*. 2019;15:157-178.
451 4. Sohn GS, Cho S, Kim YM, Cho CH, Kim MR, Lee SR, Working Group of Society of Uterine L.
452 Current medical treatment of uterine fibroids. *Obstet Gynecol Sci*. 2018;61(2):192-201.
453 5. Ono M, Maruyama T. Stem Cells in Myometrial Physiology. *Semin Reprod Med*.
454 2015;33(5):350-356.
455 6. Teixeira J, Rueda BR, Pru JK. Uterine stem cells. StemBook. Cambridge (MA)2008.
456 7. Canevari RA, Pontes A, Rosa FE, Rainho CA, Rogatto SR. Independent clonal origin of
457 multiple uterine leiomyomas that was determined by X chromosome inactivation and
458 microsatellite analysis. *Am J Obstet Gynecol*. 2005;193(4):1395-1403.
459 8. Zhang P, Zhang C, Hao J, Sung CJ, Quddus MR, Steinhoff MM, Lawrence WD. Use of X-
460 chromosome inactivation pattern to determine the clonal origins of uterine leiomyoma
461 and leiomyosarcoma. *Hum Pathol*. 2006;37(10):1350-1356.
462 9. Holdsworth-Carson SJ, Zaitseva M, Vollenhoven BJ, Rogers PA. Clonality of smooth muscle
463 and fibroblast cell populations isolated from human fibroid and myometrial tissues. *Mol*
464 *Hum Reprod*. 2014;20(3):250-259.
465 10. Fialkow PJ. Clonal origin of human tumors. *Annu Rev Med*. 1979;30:135-143.
466 11. Mas A, Nair S, Laknaur A, Simon C, Diamond MP, Al-Hendy A. Stro-1/CD44 as putative
467 human myometrial and fibroid stem cell markers. *Fertil Steril*. 2015;104(1):225-234 e223.
468 12. Yin P, Ono M, Moravek MB, Coon JSt, Navarro A, Monsivais D, Dyson MT, Druschitz SA,
469 Malpani SS, Serna VA, Qiang W, Chakravarti D, Kim JJ, Bulun SE. Human uterine
470 leiomyoma stem/progenitor cells expressing CD34 and CD49b initiate tumors in vivo. *J*
471 *Clin Endocrinol Metab*. 2015;100(4):E601-606.
472 13. Bulun SE. Uterine fibroids. *N Engl J Med*. 2013;369(14):1344-1355.
473 14. Szotek PP, Chang HL, Zhang L, Preffer F, Dombkowski D, Donahoe PK, Teixeira J. Adult
474 mouse myometrial label-retaining cells divide in response to gonadotropin stimulation.
475 *Stem Cells*. 2007;25(5):1317-1325.
476 15. Patterson AL, George JW, Chatterjee A, Carpenter T, Wolfrum E, Pru JK, Teixeira JM. Label-
477 Retaining, Putative Mesenchymal Stem Cells Contribute to Myometrial Repair During
478 Uterine Involution. *Stem cells and development*. 2018.
479 16. Ono M, Maruyama T, Masuda H, Kajitani T, Nagashima T, Arase T, Ito M, Ohta K, Uchida
480 H, Asada H, Yoshimura Y, Okano H, Matsuzaki Y. Side population in human uterine
481 myometrium displays phenotypic and functional characteristics of myometrial stem cells.
482 *Proc Natl Acad Sci U S A*. 2007;104(47):18700-18705.
483 17. Chang HL, Senaratne TN, Zhang L, Szotek PP, Stewart E, Dombkowski D, Preffer F,
484 Donahoe PK, Teixeira J. Uterine leiomyomas exhibit fewer stem/progenitor cell
485 characteristics when compared with corresponding normal myometrium. *Reprod Sci*.
486 2010;17(2):158-167.

487 18. Ono M, Kajitani T, Uchida H, Arase T, Oda H, Uchida S, Ota K, Nagashima T, Masuda H,
488 Miyazaki K, Asada H, Hida N, Mabuchi Y, Morikawa S, Ito M, Bulun SE, Okano H, Matsuzaki
489 Y, Yoshimura Y, Maruyama T. CD34 and CD49f Double-Positive and Lineage Marker-
490 Negative Cells Isolated from Human Myometrium Exhibit Stem Cell-Like Properties
491 Involved in Pregnancy-Induced Uterine Remodeling. *Biol Reprod*. 2015;93(2):37.
492 19. Golebiewska A, Brons NH, Bjerkvig R, Niclou SP. Critical appraisal of the side population
493 assay in stem cell and cancer stem cell research. *Cell Stem Cell*. 2011;8(2):136-147.
494 20. Masuda H, Anwar SS, Buhring HJ, Rao JR, Gargett CE. A novel marker of human
495 endometrial mesenchymal stem-like cells. *Cell Transplant*. 2012;21(10):2201-2214.
496 21. Patterson AL, George JW, Chatterjee A, Carpenter TJ, Wolfrum E, Chesla DW, Teixeira JM.
497 Putative human myometrial and fibroid stem-like cells have mesenchymal stem cell and
498 endometrial stromal cell properties. *Human reproduction*. 2020;35(1):44-57.
499 22. Redvers RP, Li A, Kaur P. Side population in adult murine epidermis exhibits phenotypic
500 and functional characteristics of keratinocyte stem cells. *Proc Natl Acad Sci U S A*.
501 2006;103(35):13168-13173.
502 23. Schienda J, Engleka KA, Jun S, Hansen MS, Epstein JA, Tabin CJ, Kunkel LM, Kardon G.
503 Somitic origin of limb muscle satellite and side population cells. *Proc Natl Acad Sci U S A*.
504 2006;103(4):945-950.
505 24. Fox JC, Nakayama T, Tyler RC, Sander TL, Yoshie O, Volkman BF. Structural and agonist
506 properties of XCL2, the other member of the C-chemokine subfamily. *Cytokine*.
507 2015;71(2):302-311.
508 25. Lopez-Cabrera M, Santis AG, Fernandez-Ruiz E, Blacher R, Esch F, Sanchez-Mateos P,
509 Sanchez-Madrid F. Molecular cloning, expression, and chromosomal localization of the
510 human earliest lymphocyte activation antigen AIM/CD69, a new member of the C-type
511 animal lectin superfamily of signal-transmitting receptors. *J Exp Med*. 1993;178(2):537-
512 547.
513 26. Akashi K, Kondo M, von Freeden-Jeffry U, Murray R, Weissman IL. Bcl-2 rescues T
514 lymphopoiesis in interleukin-7 receptor-deficient mice. *Cell*. 1997;89(7):1033-1041.
515 27. Bongen E, Vallania F, Utz PJ, Khatri P. KLRD1-expressing natural killer cells predict
516 influenza susceptibility. *Genome Med*. 2018;10(1):45.
517 28. Garlanda C, Dinarello CA, Mantovani A. The interleukin-1 family: back to the future.
518 *Immunity*. 2013;39(6):1003-1018.
519 29. Passarelli C, Selvatici R, Carrieri A, Di Raimo FR, Falzarano MS, Fortunato F, Rossi R, Straub
520 V, Bushby K, Reza M, Zharaieva I, D'Amico A, Bertini E, Merlini L, Sabatelli P, Borgiani P,
521 Novelli G, Messina S, Pane M, Mercuri E, Claustres M, Tuffery-Giraud S, Aartsma-Rus A,
522 Spitali P, T'Hoen PAC, Lochmuller H, Strandberg K, Al-Khalili C, Kotelnikova E, Lebowitz M,
523 Schwartz E, Muntoni F, Scapoli C, Ferlini A. Tumor Necrosis Factor Receptor SF10A
524 (TNFRSF10A) SNPs Correlate With Corticosteroid Response in Duchenne Muscular
525 Dystrophy. *Front Genet*. 2020;11:605.
526 30. Liu G, Ren F, Song Y. Upregulation of SPOCK2 inhibits the invasion and migration of
527 prostate cancer cells by regulating the MT1-MMP/MMP2 pathway. *PeerJ*. 2019;7:e7163.
528 31. Nabors LK, Wang LD, Wagers AJ, Kansas GS. Overlapping roles for endothelial selectins in
529 murine hematopoietic stem/progenitor cell homing to bone marrow. *Exp Hematol*.
530 2013;41(7):588-596.

531 32. Tufa DM, Yingst AM, Shank T, Shim S, Trahan GD, Lake J, Woods R, Jones KL, Verneris MR.
532 Transient Expression of GATA3 in Hematopoietic Stem Cells Facilitates Helper Innate
533 Lymphoid Cell Differentiation. *Front Immunol*. 2019;10:510.
534 33. Bujanover N, Goldstein O, Greenshpan Y, Turgeman H, Klainberger A, Scharff Y, Gazit R.
535 Identification of immune-activated hematopoietic stem cells. *Leukemia*.
536 2018;32(9):2016-2020.
537 34. Pinho S, Wei Q, Maryanovich M, Zhang D, Balandran JC, Pierce H, Nakahara F, Di Staulo
538 A, Bartholdy BA, Xu J, Borger DK, Verma A, Frenette PS. VCAM1 confers innate immune
539 tolerance on haematopoietic and leukaemic stem cells. *Nat Cell Biol*. 2022;24(3):290-298.
540 35. Pique-Regi R, Romero R, Garcia-Flores V, Peyvandipour A, Tarca AL, Pusod E, Galaz J,
541 Miller D, Bhatti G, Para R, Kanninen T, Hadaya O, Paredes C, Motomura K, Johnson JR,
542 Jung E, Hsu CD, Berry SM, Gomez-Lopez N. A single-cell atlas of the myometrium in human
543 parturition. *JCI Insight*. 2022;7(5).
544 36. Goad J, Rudolph J, Zandigohar M, Tae M, Dai Y, Wei JJ, Bulun SE, Chakravarti D, Rajkovic
545 A. Single-cell sequencing reveals novel cellular heterogeneity in uterine leiomyomas. *Hum*
546 *Reprod*. 2022;37(10):2334-2349.
547 37. O'Connor SA, Feldman HM, Arora S, Hoellerbauer P, Toledo CM, Corrin P, Carter L, Kufeld
548 M, Bolouri H, Basom R, Delrow J, McFaline-Figueroa JL, Trapnell C, Pollard SM, Patel A,
549 Paddison PJ, Plaisier CL. Neural G0: a quiescent-like state found in neuroepithelial-derived
550 cells and glioma. *Mol Syst Biol*. 2021;17(6):e9522.
551 38. Riba A, Oravecz A, Durik M, Jimenez S, Alunni V, Cerciat M, Jung M, Keime C, Keyes WM,
552 Molina N. Cell cycle gene regulation dynamics revealed by RNA velocity and deep-
553 learning. *Nat Commun*. 2022;13(1):2865.
554 39. Zywitza V, Misios A, Bunatyan L, Willnow TE, Rajewsky N. Single-Cell Transcriptomics
555 Characterizes Cell Types in the Subventricular Zone and Uncovers Molecular Defects
556 Impairing Adult Neurogenesis. *Cell Rep*. 2018;25(9):2457-2469 e2458.
557 40. Scialdone A, Natarajan KN, Saraiva LR, Proserpio V, Teichmann SA, Stegle O, Marioni JC,
558 Buettner F. Computational assignment of cell-cycle stage from single-cell transcriptome
559 data. *Methods*. 2015;85:54-61.
560 41. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K,
561 Kastriti ME, Lonnerberg P, Furlan A, Fan J, Borm LE, Liu Z, van Bruggen D, Guo J, He X,
562 Barker R, Sundstrom E, Castelo-Branco G, Cramer P, Adameyko I, Linnarsson S,
563 Kharchenko PV. RNA velocity of single cells. *Nature*. 2018;560(7719):494-498.
564 42. Bautch VL. Stem cells and the vasculature. *Nat Med*. 2011;17(11):1437-1443.
565 43. Crisan M, Yap S, Casteilla L, Chen CW, Corselli M, Park TS, Andriolo G, Sun B, Zheng B,
566 Zhang L, Norotte C, Teng PN, Traas J, Schugar R, Deasy BM, Badylak S, Buhring HJ,
567 Giacobino JP, Lazzari L, Huard J, Peault B. A perivascular origin for mesenchymal stem cells
568 in multiple human organs. *Cell Stem Cell*. 2008;3(3):301-313.
569 44. Lv FJ, Tuan RS, Cheung KM, Leung VY. Concise review: the surface markers and identity of
570 human mesenchymal stem cells. *Stem Cells*. 2014;32(6):1408-1419.
571 45. Bhartiya D. Adult tissue-resident stem cells-fact or fiction? *Stem Cell Res Ther*.
572 2021;12(1):73.

573 46. Zhang LZ, Huang LY, Huang AL, Liu JX, Yang F. CRIP1 promotes cell migration, invasion and
574 epithelial-mesenchymal transition of cervical cancer by activating the Wnt/beta-catenin
575 signaling pathway. *Life Sci*. 2018;207:420-427.
576 47. Fevr T, Robine S, Louvard D, Huelsken J. Wnt/beta-catenin is essential for intestinal
577 homeostasis and maintenance of intestinal stem cells. *Mol Cell Biol*. 2007;27(21):7551-
578 7559.
579 48. Nusse R. Wnt signaling and stem cell control. *Cell Res*. 2008;18(5):523-527.
580 49. Mas A, Cervello I, Gil-Sanchis C, Faus A, Ferro J, Pellicer A, Simon C. Identification and
581 characterization of the human leiomyoma side population as putative tumor-initiating
582 cells. *Fertil Steril*. 2012;98(3):741-751 e746.
583 50. Paul EN, Burns GW, Carpenter TJ, Grey JA, Fazleabas AT, Teixeira JM. Transcriptome
584 Analyses of Myometrium from Fibroid Patients Reveals Phenotypic Differences Compared
585 to Non-Diseased Myometrium. *Int J Mol Sci*. 2021;22(7).
586 51. Tang L, Bergevoet SM, Gilissen C, de Witte T, Jansen JH, van der Reijden BA, Raymakers
587 RA. Hematopoietic stem cells exhibit a specific ABC transporter gene expression profile
588 clearly distinct from other stem cells. *BMC Pharmacol*. 2010;10:12.
589 52. Arora R, Fries A, Oelerich K, Marchuk K, Sabeur K, Giudice LC, Laird DJ. Insights from
590 imaging the implanting embryo and the uterine environment in three dimensions.
591 *Development*. 2016;143(24):4749-4754.
592 53. Flores D, Madhavan M, Wright S, Arora R. Mechanical and signaling mechanisms that
593 guide pre-implantation embryo movement. *Development*. 2020;147(24).
594 54. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful
595 Approach to Multiple Testing. *J Roy Stat Soc B Met*. 1995;57(1):289-300.
596

597 **Figure Legends:**

598 **Figure 1. Bulk RNA-seq results of SUSD2$^-$ and SUSD2$^+$ cells from myometrial**
599 **samples.** (**A**) Representative flow cytometry scatter plot of single cell SUSD2$^-$ and
600 SUSD2$^+$ sort showing over 30% of the live myometrial cells positive for the SUSD2
601 marker. Boxed areas indicate gating strategy. (**B**) Principal component analysis (PCA)
602 plot of RNA-seq results from SUSD2$^-$ (in black) and SUSD2$^+$ (in red) cells, each label
603 represents one sample (n = 5), variance for each PC is indicated in percentage. (**C**)
604 Volcano plot showing up (n = 3527) and down (n = 3250) DEGs with a false discovery
605 rate (FDR) p-value < 0.05 in SUSD2$^+$ vs. SUSD2$^-$ cells depicted as red dots. Grey dots
606 represent genes with an FDR p-value > 0.05. (**D**) Boxplot of mesenchymal stem cells
607 markers, *SUSD2*, *MCAM*, *PDGFRβ* and *CSPG4* in the SUSD2$^-$ (in grey) and SUSD2$^+$ (in
608 red) cell population (n = 5). SUSD2$^+$ sorted cells are significantly enriched for MSC
609 markers including *SUSD2* ($\log_2$FC= 4.5, FDR p= 2.9x10^{-5}), *MCAM* ($\log_2$FC= 3.1, FDR p=
610 8.8x10^{-16}), *PDGFRβ* ($\log_2$FC= 2.8, FDR p= 1.5x10^{-11}) and *CSPG4* ($\log_2$FC= 3.4, FDR p=
611 1.1x10^{-11}). (**E**) Gene expression is shown as $\log_2$(CPM). Heatmap of the top 300 DEGs
612 from SUSD2$^+$ vs. SUSD2$^-$ cells comparison with unsupervised hierarchical clustering of
613 genes and samples (n = 5). Color gradient represents gene expression as z-score.

614

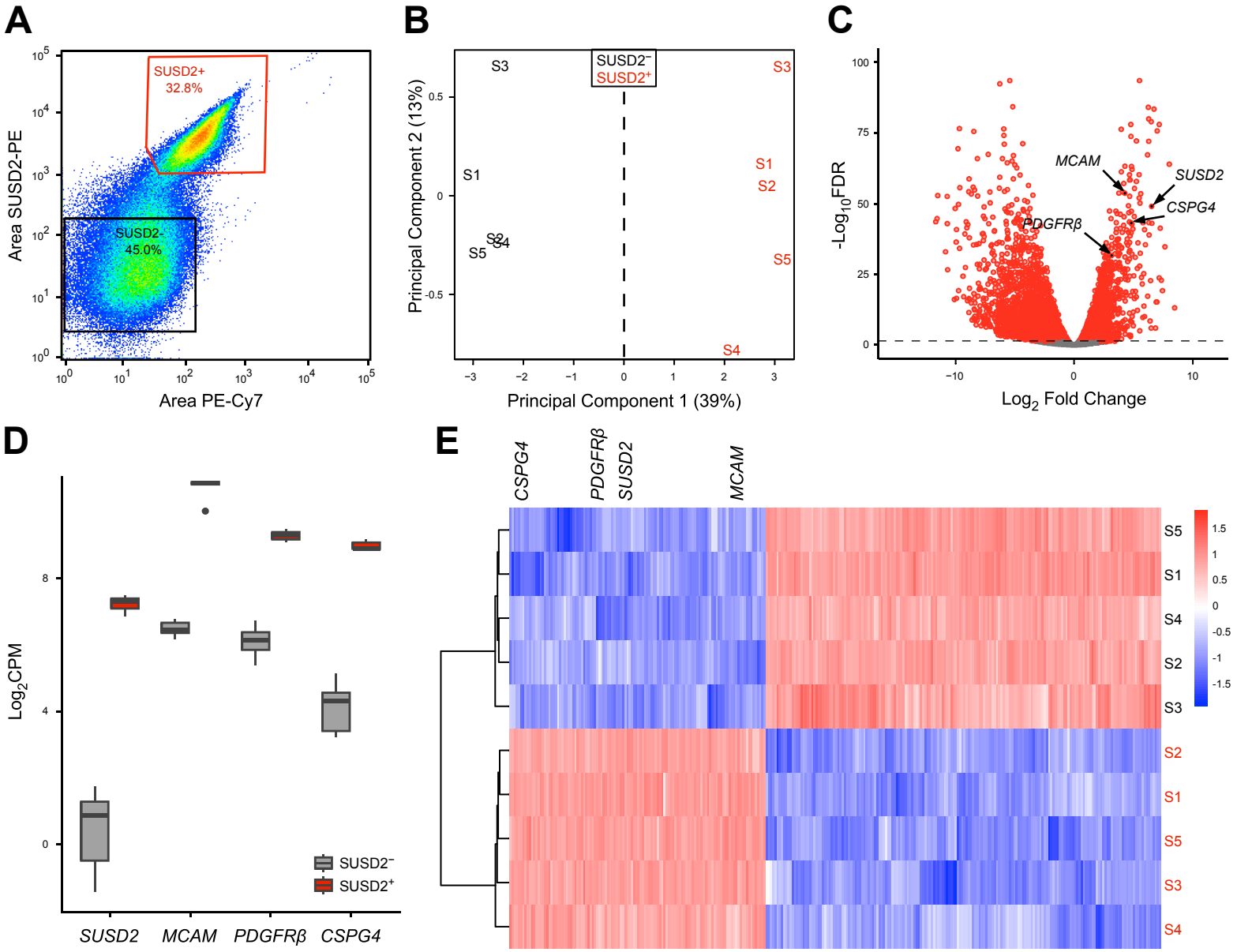
615 **Figure 2. Transcriptomic analysis of the myometrium side population.** (**A**) Scatter
616 plot of the gating strategy to sort the side population (SP$^+$) and the SP$^-$ cells from human
617 myometrium. (**B**) Verapamil pre-treatment of myometrium cells show a decrease of the
618 SP$^+$ cells from 1.99% to 0.1% of the total live single cells. (**C**) PCA plot of RNA-seq results
619 from SP$^-$ (in black) and SP$^+$ (in red) cells, each label represents one sample (n = 3),

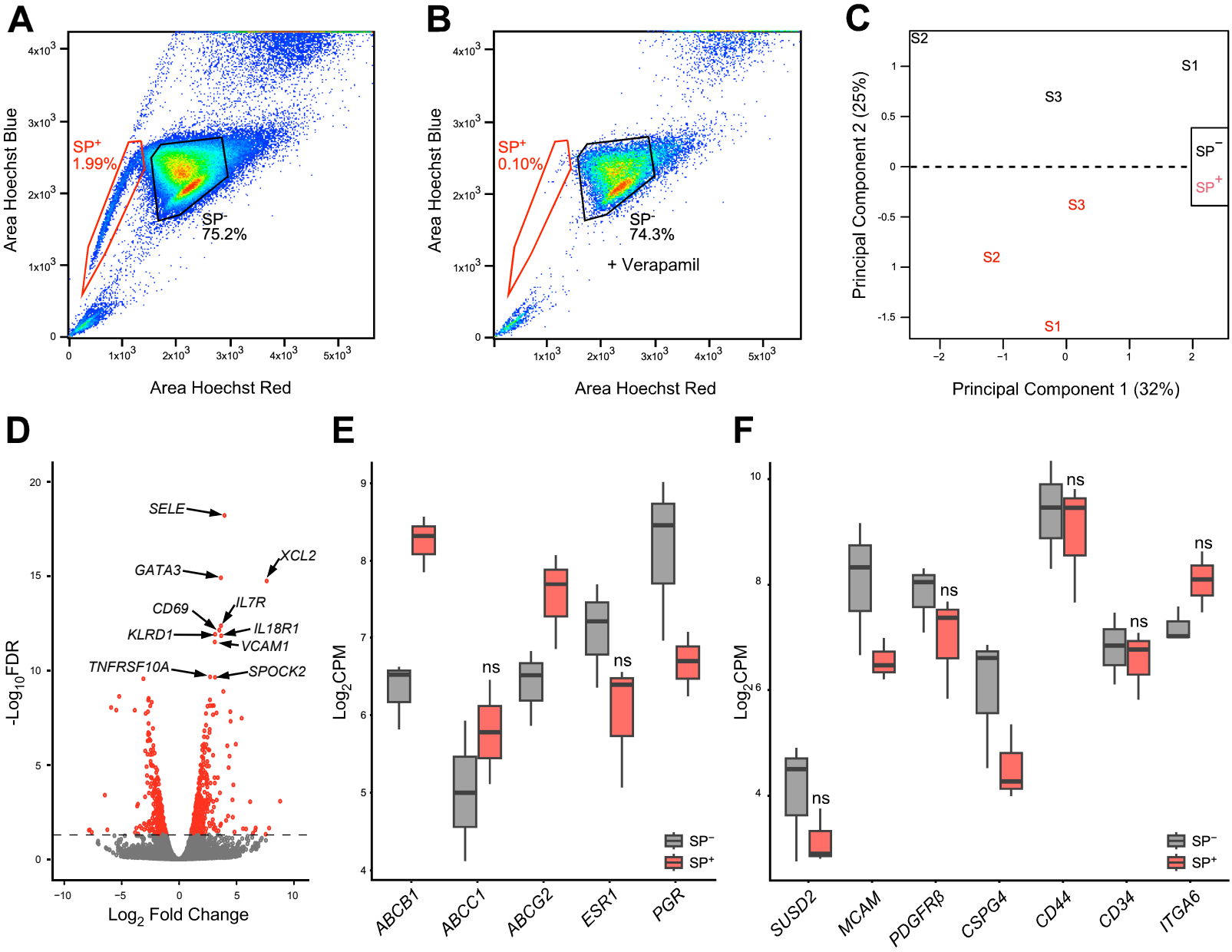
620 variance for each PC is indicated in percentage. (**D**) Volcano plot showing up (n = 478)
621 and down (n = 350) DEGs with a false discovery rate (FDR) p-value < 0.05 in SP^+ vs. SP^-
622 cells depicted as red dots, including *SELE* (log_2FC= 4.0, FDR p= 6.0 $x10^{-19}$), *GATA3*
623 (log_2FC= 3.7, FDR p= 2.8 $x10^{-15}$), *XCL2* (log_2FC= 7.7, FDR p= 1.7 $x10^{-15}$), *IL7R* (log_2FC=
624 3.7, FDR p= 4.5 $x10^{-13}$), *CD69* (log_2FC= 3.5, FDR p= 7.6 $x10^{-13}$), *KLRD1* (log_2FC= 3.2,
625 FDR p= 1.3 $x10^{-12}$), *IL18R1* (log_2FC= 3.7, FDR p= 1.5 $x10^{-12}$), *VCAM1* (log_2FC= 3.1, FDR
626 p= 3.3 $x10^{-12}$), *TNFRSF10A* (log_2FC= 2.7, FDR p= 2.3 $x10^{-10}$), and *SPOCK2* (log_2FC=
627 3.2, FDR p= 2.4 $x10^{-10}$). Grey dots represent genes with an FDR p-value > 0.05. (**E**)
628 Boxplots of myometrium SP associated genes, *ABCB1* (log_2FC= 1.9, FDR p= 2.8$x10^{-5}$),
629 *ABCC1* (log_2FC= 0.7, FDR p= 5.9 $x10^{-1}$), *ABCG2* (log_2FC= 1.1, FDR p= 9.5$x10^{-2}$), *ESR1*
630 (log_2FC= -1.1, FDR p= 2.3 $x10^{-1}$) and *PGR* (log_2FC= -1.5, FDR p= 6.4 $x10^{-2}$) and (**F**)
631 mesenchymal stem cells associated genes, *SUSD2* (log_2FC= -1, FDR p= 5.6$x10^{-1}$),
632 *MCAM* (log_2FC= -1.5, FDR p= 3.9$x10^{-2}$), *PDGFRβ* (log_2FC= -0.9, FDR p= 4.9$x10^{-1}$),
633 *CSPG4* (log_2FC= -1.5, FDR p= 3.9$x10^{-2}$), *CD44* (log_2FC= -0.5, FDR p= 8.3$x10^{-1}$), *CD34*
634 (log_2FC= -0.2, FDR p= 4.9$x10^{-1}$), and *ITGA6* (log_2FC= 0.9, FDR p= 2.1$x10^{-1}$) in the SP^-
635 (in grey) and SP^+ (in red) cell population (n = 3), genes are expressed in log_2(CPM). ns =
636 non-significant, with FDR > 0.05.
637
638 **Figure 3. Single cell RNA-seq analysis of isolated cells from human myometrial**
639 **samples.** (**A**) Uniform manifold approximation and projection (UMAP) visualization of
640 9775 isolated cells from human myometrial samples (n = 5). Each cluster (n = 7) represent
641 a cell population with a similar transcriptomic profile. (**B**) Dotplot for cluster identification
642 using specific markers for cell types found in the myometrium. Mesenchymal stem cell

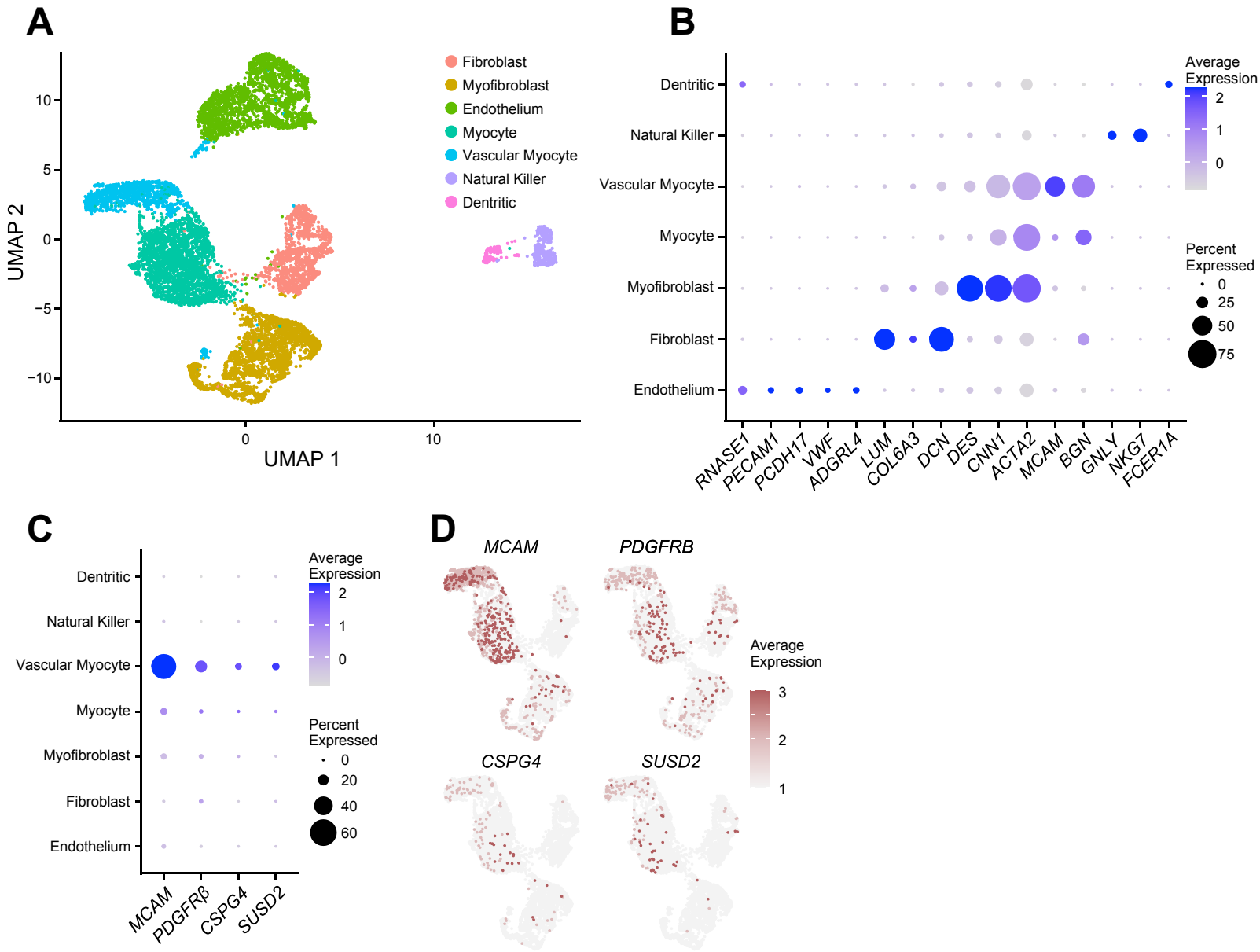
643 marker gene expression in the different myometrial cell clusters shown in a dotplot (**C**)
644 and by UMAP (**D**). Average gene expression and percentage of cells expressing the
645 specific gene in each cell cluster are shown by the color intensity and the diameter of the
646 dot, respectively, in B and C. Color gradient in the UMAP represents gene expression as
647 log_2(CPM).
648
649 **Figure 4. Identification of putative MyoSPCs from scRNA-seq**
650 (**A**) Cell cycle score for each myometrium cells visualized in the UMAP plot. Cells in
651 G1/G0, S, and G2/M phases are plotted with the corresponding color. Boxed area is
652 shown at higher magnification. (**B**) Cell velocity predicting the future state of individual
653 cells illustrated in a UMAP plotted with the clusters as in panel A showed that in the
654 vascular myocyte cluster, the same group of cells in G1/G0 phase in panel A exhibit low
655 velocity. Boxed area is shown at higher magnification. Putative MyoSPCs are encircle in
656 black dots.
657
658 **Figure 5. Overlapping analysis of the bulk RNA-seq of SUSD2+ vs SUSD2- and the**
659 **assigned stem cell cluster vs the rest of the myometrium cells from the single cell**
660 **RNA-seq.** Venn diagrams illustrate the overlap DEG between the bulk RNA-seq of
661 SUSD2+ vs SUSD2- and the assigned stem cell cluster vs the rest of the myometrium
662 cells from the single cell RNA-seq analysis **(A)**. Scatter plot of Log_2(fold change) genes
663 from bulk RNA-seq (x axis) and single cell RNA-seq (y axis) **(B)**. Non-significant genes in
664 both analyses were represented in purple dots, DEG in the bulk RNA-seq only were
665 represented in blue dots, DEG in the scRNA-seq only were represented in green dots,

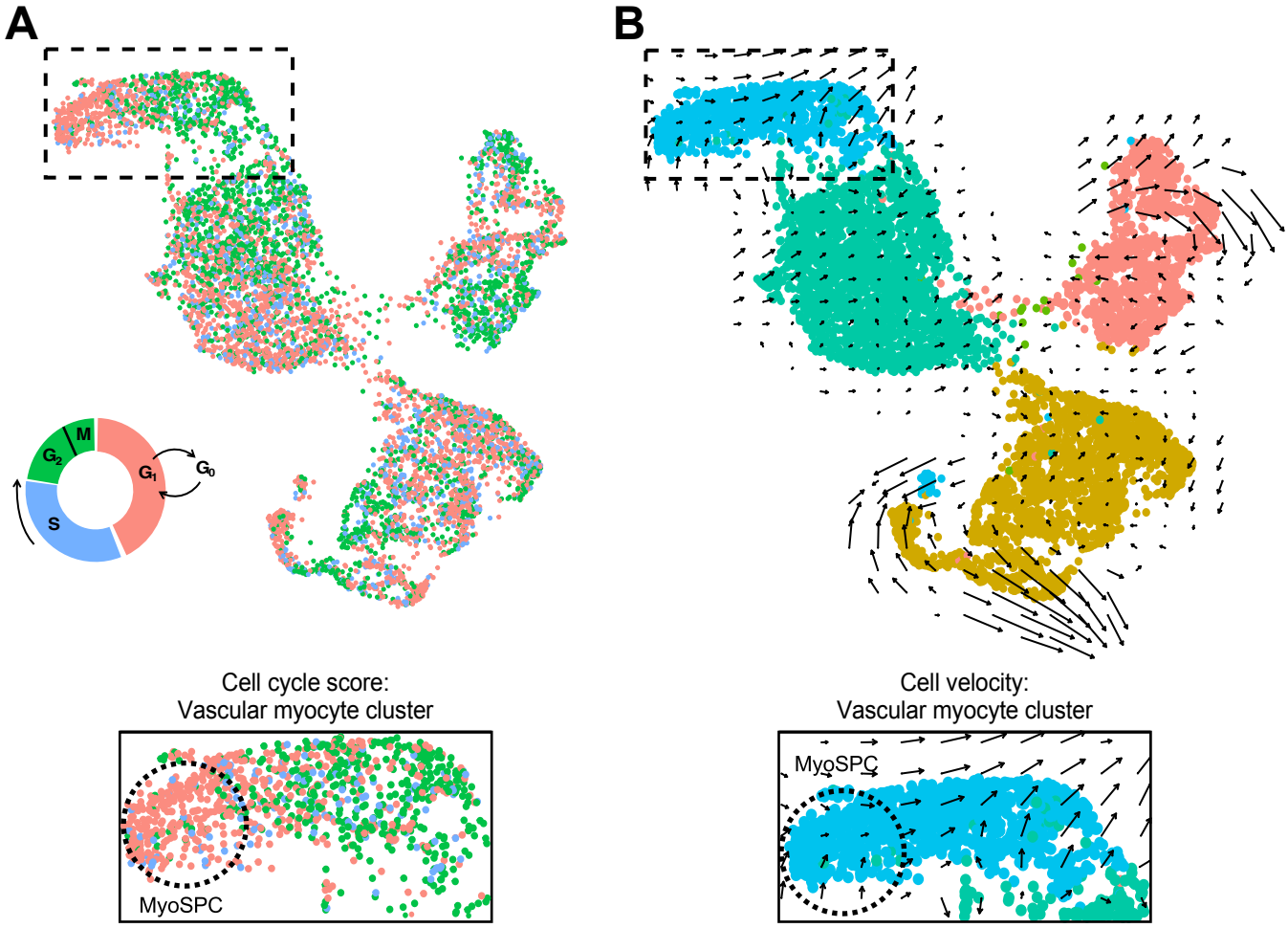
666 and the DEG in both analyses were represented in red dots. *CRIP1* is highly upregulated
667 in the MyoSPC cluster (log_2FC=3.1, adjusted p value = 4 x 10^{-251}) and in the SUSD2+
668 cells (log_2FC=1.9, FDR = 5 x 10^{-4}). UMAP (**C**) and Ridge plot (**D**) show *CRIP1* expression
669 in the different myometrium cell clusters by the color intensity or intensity level.
670
671 **Figure 6. CRIP1$^+$ cells have stem/progenitor cell characteristics.** (**A**)
672 Immunofluorescence imaging of human myometrium using PECAM1 as an endothelial
673 marker, SUSD2 as a mesenchymal stem cell marker, and CRIP1. Scale bar = 200 μm.
674 (**B**) Representative scatter plot of the gating strategy for CRIP1$^+$/PECAM$^-$ cell sort. (**C**)
675 Representative images of colonies formed by the CRIP1$^+$/PECAM$^-$ and depleted
676 myometrium cell population. (**D**) Plot with of number of colonies formed from
677 CRIP1$^+$/PECAM$^-$ and depleted myometrium cells after 2-3 weeks in culture (n = 6). (**E**)
678 Plot of colony forming efficiency represented as %CFUs (#CFU/cells seeded × 100) of
679 CRIP1$^+$/PECAM$^-$ and depleted myometrium cells (n = 6). (**F**) Total area of colony formed
680 in pixels from CRIP1$^+$/PECAM$^-$ and depleted myometrium cell population (n = 6). (**G**)
681 ACTA2 immunofluorescence in CRIP1$^+$/PECAM$^-$ myometrial cells after differentiation.
682 Scale bar = 100 μm. Representative images of CRIP1$^+$/PECAM$^-$ and depleted
683 myometrium cell population grown in control growth media and adipogenic (H) or
684 osteogenic (I) differentiation media. Adipogenic and control cultures were stained with Oil
685 Red O (red color, black arrows), and osteogenic and control cultures were stained for
686 alkaline phosphatase activity (purple color). Scale bar for the adipogenic and osteogenic
687 assays are 500 μm and 5000 μm, respectively. *$p<0.05$; **$p< 0.01$ by student t test.
688

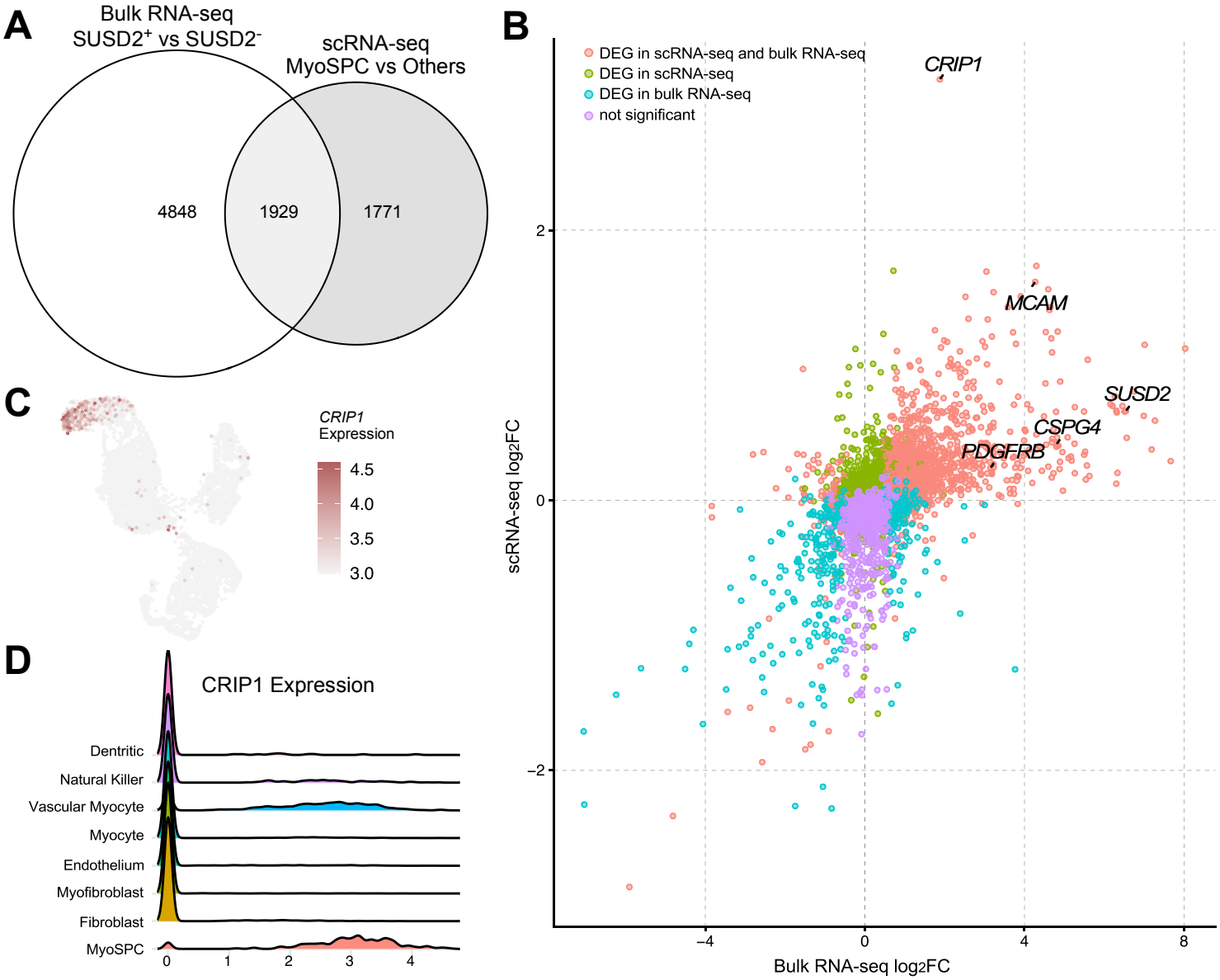
689 **Figure S1. Cell distribution across cell clusters in the single cell RNA-seq.** Uniform
690 manifold approximation and projection (UMAP) visualization of 9775 isolated cells from 5
691 human myometrial samples (n = 5). Each color dots represents cells from a myometrium
692 sample. UMAP plot shows that patient's cells are equally distributed across clusters.
693
694 **Figure S2. *CRIP1* expression in the side population.** *CRIP1* expression in $\log_2$CPM of
695 the RNA-seq results from the side population (SP^+) and the negative population (SP-)
696 showing no significant differences (FDR>0.05) between the two population.'

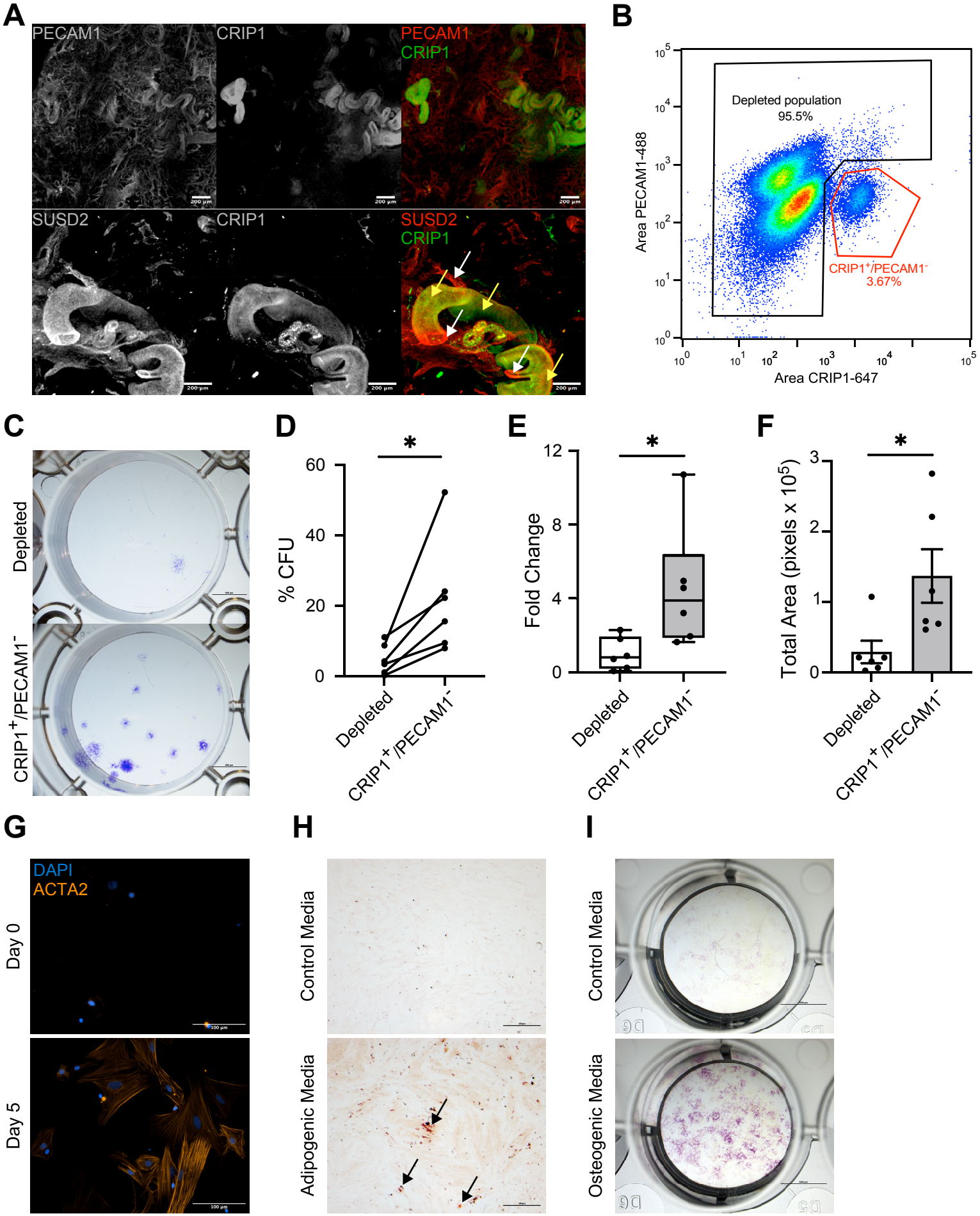

A
SUSD2+
32.8%
SUSD2-
45.0%
Area SUSD2-PE
Area PE-Cy7
B
SUSD2⁻
SUSD2⁺
S3
S1
S2
S4
S5
Principal Component 2 (13%)
Principal Component 1 (39%)
C
MCAM
SUSD2
CSPG4
PDGFRβ
-Log10FDR
Log2 Fold Change
D
Log2CPM
SUSD2⁻
SUSD2⁺
SUSD2
MCAM
PDGFRβ
CSPG4
E
CSPG4
PDGFRβ
SUSD2
MCAM
S5
S1
S4
S2
S3
S2
S1
S5
S3
S4


A
Area Hoechst Blue
Area Hoechst Red
SP+
1.99%
SP-
75.2%
B
Area Hoechst Blue
Area Hoechst Red
SP+
0.10%
SP-
74.3%
+ Verapamil
C
Principal Component 2 (25%)
Principal Component 1 (32%)
S1
S2
S3
SP−
SP+
D
-Log10FDR
Log2 Fold Change
SELE
XCL2
GATA3
CD69
IL7R
KLRD1
IL18R1
VCAM1
TNFRSF10A
SPOCK2
E
Log2CPM
ns
SP−
SP+
ABCB1
ABCC1
ABCG2
ESR1
PGR
F
Log2CPM
ns
SP−
SP+
SUSD2
MCAM
PDGFRβ
CSPG4
CD44
CD34
ITGA6


A
UMAP 2
UMAP 1
10
5
0
−5
−10
0
10
Fibroblast
Myofibroblast
Endothelium
Myocyte
Vascular Myocyte
Natural Killer
Dentritic
B
Dentritic
Natural Killer
Vascular Myocyte
Myocyte
Myofibroblast
Fibroblast
Endothelium
RNASE1
PECAM1
PCDH17
VWF
ADGRL4
LUM
COL6A3
DCN
DES
CNN1
ACTA2
MCAM
BGN
GNLY
NKG7
FCER1A
Average Expression
2
1
0
Percent Expressed
0
25
50
75
C
Dentritic
Natural Killer
Vascular Myocyte
Myocyte
Myofibroblast
Fibroblast
Endothelium
MCAM
PDGFRβ
CSPG4
SUSD2
Average Expression
2
1
0
Percent Expressed
0
20
40
60
D
MCAM
PDGFRB
CSPG4
SUSD2
Average Expression
3
2
1


A
B
M
G2
G1
G0
S
Cell cycle score:
Vascular myocyte cluster
MyoSPC
Cell velocity:
Vascular myocyte cluster
MyoSPC


A
Bulk RNA-seq
SUSD2+ vs SUSD2-
scRNA-seq
MyoSPC vs Others
4848
1929
1771
B
DEG in scRNA-seq and bulk RNA-seq
DEG in scRNA-seq
DEG in bulk RNA-seq
not significant
CRIP1
MCAM
SUSD2
CSPG4
PDGFRB
scRNA-seq log2FC
Bulk RNA-seq log2FC
2
0
−2
−4
0
4
8
C
CRIP1
Expression
4.5
4.0
3.5
3.0
D
CRIP1 Expression
Dentritic
Natural Killer
Vascular Myocyte
Myocyte
Endothelium
Myofibroblast
Fibroblast
MyoSPC
0
1
2
3
4


A
PECAM1
CRIP1
PECAM1
CRIP1
SUSD2
CRIP1
SUSD2
CRIP1
200 µm
B
Depleted population
95.5%
CRIP1+/PECAM1-
3.67%
Area PECAM1-488
Area CRIP1-647
C
Depleted
CRIP1+/PECAM1-
D
% CFU
Depleted
CRIP1+/PECAM1-
E
Fold Change
Depleted
CRIP1+/PECAM1-
F
Total Area (pixels x 10^5)
Depleted
CRIP1+/PECAM1-
G
DAPI
ACTA2
Day 0
Day 5
100 µm
H
Control Media
Adipogenic Media
I
Control Media
Osteogenic Media