Unraveling the Molecular Landscape of Uterine Fibroids, Insights into *HMGA2* and Stem Cell Involvement

Emmanuel N. Paul[1], Tyler J. Carpenter[1], Laura A. Pavliscak[1], Abigail Z. Bennett[1], Maria Ariadna Ochoa-Bernal[1], Asgerally T. Fazleabas[1], Jose M. Teixeira[1*]

[1]Department of Obstetrics, Gynecology and Reproductive Biology, College of Human Medicine, Michigan State University, Grand Rapids, MI 49503, USA

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

Acknowledgments:
These studies have been funded by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NIH R01HD096259) to JMT and the SRI and Bayer Discovery/Innovation Grant to EP.

The authors have no conflicts to declare.

Abstract

Uterine fibroids are prevalent benign tumors in women that exhibit considerable heterogeneity in clinical presentation and molecular characteristics, necessitating a deeper understanding of their etiology and pathogenesis. *HMGA2* overexpression has been associated with fibroid development, yet its precise role remains elusive. Mutations in fibroids are mutually exclusive and largely clonal, suggesting that tumors originate from a single mutant cell. We explored a possible role for *HMGA2* overexpression in differentiated myometrial cells, hypothesizing its potential to induce a stem cell-like or dedifferentiating phenotype and drive fibroid development. Myometrial cells were immortalized and transduced with an *HMGA2* lentivirus to produce HMGA2hi cells. *In vitro* stem cell assays were conducted and RNA from HMGA2hi and control cells and fibroid-free myometrial and HMGA2 fibroid (HMGA2F) tissues were submitted for RNA-sequencing. HMGA2hi cells have enhanced self-renewal capacity, decreased proliferation, and have a greater ability to differentiate into other mesenchymal cell types. HMGA2hi cells exhibit a stem cell-like signature and share transcriptomic similarities with HMGA2F. Moreover, dysregulated extracellular matrix pathways are observed in both HMGA2hi cells and HMGA2F. Our findings suggest that HMGA2 overexpression drives myometrial cells to dedifferentiate into a more plastic phenotype and underscore a pivotal role for HMGA2 in fibroid pathogenesis.

Keywords:

Myometrium, fibroids, stem cell, transcriptome, extracellular matrix.

43 **Introduction**

44 Uterine fibroids, or leiomyomas, represent the most prevalent benign tumors in
45 women of reproductive age, affecting approximately 70-80% of women during their
46 lifetime (1). These tumors are mainly characterized by extensive extracellular matrix
47 (ECM) deposition and remodeling (2). Despite their high incidence, the etiology and
48 pathogenesis of uterine fibroids remain poorly understood.

49 Uterine fibroids exhibit considerable heterogeneity in terms of their clinical
50 presentation, growth patterns, and response to therapeutic interventions (3), which
51 suggests different phenotypic subtypes. Recent reports showing molecular
52 characterization of fibroids have provided unprecedented insights into the complex
53 interplay of genetic, epigenetic, and cellular factors contributing to the development and
54 progression of uterine fibroids (4-6) and support the classification of uterine fibroids into
55 distinct subtypes based on their combined molecular and histological characteristics.
56 Mutations identified in various genes, including *MED12*, *FH*, and *COL4A5*, have been
57 associated with distinct subtypes of fibroids, potentially contributing to the diverse clinical
58 manifestations observed in affected individuals (7,8). Unraveling the molecular
59 mechanisms driving the pathogenesis of the various fibroid subtypes is critical for
60 developing targeted therapeutic strategies and personalized treatment approaches.

61 Another genetic factor associated with fibroid development is overexpression of
62 the high-mobility group AT-hook 2 (*HMGA2*) gene (9-12). *HMGA2* is located on
63 chromosome 12q15 and encodes a non-histone chromatin-binding protein that plays a
64 crucial role in regulating gene expression especially during embryonic development,
65 including stem cell pluripotency, lineage specification, and organogenesis (13). *HMGA2*

66 transcripts are detected in most embryonic tissues and progressively become absent in
67 adult tissues, except for the testis, the lung, and the kidney (14). Its dynamic expression
68 and regulatory functions are tightly orchestrated to ensure proper development, with
69 aberrations in *HMGA2* expression linked to developmental abnormalities and diseases
70 (13,15,16). Overexpression of *HMGA2* commonly resulting from the translocation
71 between chromosome bands 12q15 and 14q24 has been previously described as the
72 second major subtype of uterine fibroids (HMGA2F), constituting approximately 7% of
73 fibroid cases (17,18).

74 The discovery of a potential stem/progenitor cell population within the myometrium
75 has generated interest in uncovering the involvement of these cells in the initiation and
76 growth of fibroids (19,20). Indeed, uterine fibroids are thought to be a clonal disease
77 (21,22), and since most clonal diseases have a single cell origin, the prevailing hypothesis
78 has been that a dysregulated myometrium stem cell could be the cell of origin for uterine
79 fibroids (23). However, the molecular mechanisms governing the interplay between
80 myometrium stem cells and *HMGA2* in uterine fibroids remain largely unexplored. In this
81 study, we explored an alternative hypothesis regarding the cell origin for uterine fibroids
82 by postulating that a mutated differentiated myometrial cell could give rise to a fibroid. We
83 show that *HMGA2* overexpression in myometrium cells induces stem/progenitor cell
84 properties, suggesting a potential gain of function for the gene in fibroid development.
85 Such insights may pave the way for the development of novel therapeutic targets that can
86 effectively address uterine fibroids in a subtype-specific manner.
87

88 **Methods**

89 **Myometrium cell line generation**

90 HEK293T cells were transfected with hTERT plasmid, Addgene #85140 (24) and helper
91 plasmids (25) to prepare hTERT lentivirus. Myometrial cells were dissociated and isolated
92 from a fibroid-free myometrium patient (MyoN) as previously described (26) and
93 incubated at 37°C overnight in growth medium (DMEM/F12, 10% FBS) prior to
94 transduction with lentiviral preparations. The subsequent day, transduced MyoN cells
95 were selected with 50 µg/mL of hygromycin for 1 week. Following selection, MyoN-TERT
96 cells were maintained in medium with 5 µg/mL hygromycin.

97

98 ***HMGA2* Overexpressing and Control MyoN-TERT Cell Lines**

99 MyoN-TERT cells were suspended in 1% bovine serum albumin blocking buffer,
100 incubated with a rabbit anti-human CRIP1 primary antibody (ThermoFisher, #PA5-24643).
101 Cells were then incubated with an Alexa-647 anti-rabbit secondary antibody and washed
102 with flow buffer, and resuspended in flow buffer with of 4′,6-diamidino-2-phenylindole
103 (DAPI). Cells were sorted by the flow cytometry core at Van Andel Research Institute
104 (VARI) using a FACSymphony S6 cytometer (BD Biosciences) and analyzed with FlowJo
105 Software (BD Biosciences, version 10.8.1). Lentiviral particles isolated from in HEK293T
106 cells transfected with either a custom *HMGA2* (VectorBuilder) or empty control plasmid
107 (VectorBuilder) were used to transduce CRIP1^{-} MyoN-TERT cells as described above.
108 Selection of successfully transduced MyoN-TERT live cells expressing GFP were sorted
109 by FACS. *HMGA2* overexpressing cells are called HMGA2hi cells.

110

111 **Proliferation Assay**

112 HMGA2hi and control cells were plated in triplicate a 96 well plate at 2,000 cells/well, and
113 cell proliferation was recorded every 4 h using an Incucyte SX5 instrument (Sartorius),
114 following the manufacturer's protocols, until cell confluence was reached. Cell
115 proliferation was quantified using GFP and phase contrast images at each time point with
116 the cell-by-cell software module of the instrument. The experiment was repeated three
117 times using three different cell passages.

118

119 **Colony Formation Assay**

120 Matched passages of HMGA2hi and control cells were plated at 50 cells/cm^2 in triplicate
121 overnight, then switched to MesenPro RS (Thermo Fisher, # 12746012) for 2 to 3 weeks.
122 Cells were cultured in different wells of the same plate, and both cell types were assayed
123 on the same day. Cultures were fixed in 4% paraformaldehyde (PFA) and stained with
124 crystal violet for colony visualization. Images were taken using a Nikon SMZ18
125 microscope and Ds-Ri1 camera (Nikon Instruments Inc.). Colonies were counted and total
126 surface area was estimated using ImageJ (version 1.53k), and the percent colony-forming
127 units (CFUs) was calculated as (number of colonies/number of cells plated) × 100 and
128 averaged for triplicates.

129

130 **Differentiation Assays**

131 For osteogenic and adipogenic differentiation, matched passage HMGA2hi and control
132 cells were plated in duplicate at 50% to 80% confluency in growth media overnight and
133 then cultured for 10 days in fresh StemPro Adipogenesis Differentiation (Thermo Fisher

134 Scientific, #A1007001) or StemPro Osteogenesis Differentiation (Thermo Fisher
135 Scientific, #A1007201) media according to the manufacturer's instructions. Cells were
136 cultured in regular growth media to serve as differentiation controls. To assay adipogenic
137 differentiation, cultures were fixed in 4% PFA and stained using Oil Red O (Sigma,
138 #01391) according to the manufacturer's instructions. To assay osteogenic differentiation,
139 cultures were stained for alkaline phosphatase activity using the Alkaline Phosphatase
140 (AP), Leukocyte kit (Sigma, #86R-1KT) according to the manufacturer's instructions.
141
142 **3D Culture and Trichrome Staining**
143 60,000 cells of HMGA2hi and control cells, as well as primary MyoN and fibroid cells,
144 were seeded in scaffold-free 96-well agarose micro-molds (MilliporeSigma, Z764043),
145 following the protocol outlined by Wiwatpanit et al. (27). Spheroids were cultured for 4
146 days in growth medium. Following culture, agarose micro-molds were sealed with
147 agarose and fixed in 4% PFA at 4°C overnight. Subsequently, they were briefly rinsed in
148 PBS and stored in 70% ethanol at 4°C until processed for standard paraffin embedding
149 and sectioning at 6 μm. Spheroid sections were stained with trichrome stain according to
150 the manufacturer's protocol (StatLab #KTTRBPT). Briefly, sections were deparaffinized
151 with two xylene baths and rehydrated with three absolute ethanol washes. After hydration,
152 spheroid sections were incubated in Bouin's fluid at room temperature (RT) overnight.
153 The next day, after rinsing in running tap water, slides were stained with Mayer's
154 hematoxylin and rinsed in running tap water. Spheroid sections were then immersed in
155 the one-step trichrome stain and quickly rinsed in running tap water. Slides were
156 dehydrated through three changes of absolute ethanol, cleared in three changes of

157 xylenes, and a coverslip was applied using Permount mounting medium (Fisher Scientific
158 #SP15). Collagen deposition was determined using the color deconvolution ImageJ
159 plugin (version 1.53k). Relative collagen area was calculated by the ratio of blue stain
160 (collagen) over the red stain (cytoplasm).
161
162 **Tissue Collection and *HMGA2* fibroid mutation detection**
163 The use of human tissue specimens was approved by the Spectrum Health Systems and
164 Michigan State University Institutional Review Boards (MSU IRB Study ID:
165 STUDY00003101, SR IRB #2017-198) as secondary use of biobank materials. Fibroids
166 samples (HMGA2F, n = 23) and fibroid-free myometrial (MyoN, n = 32) from pre-
167 menopausal women (aged 28-55) were obtained following total hysterectomy. All patients
168 who participated in the study gave consent to donate tissue through the Spectrum Health
169 Biorepository. Human samples were processed as previously described (4). Briefly,
170 samples were washed with phosphate-buffered saline, minced, and immersed in
171 RNAlater (Sigma, Saint Louis, MO, USA) for RNA-seq analyses. *HMGA2* fibroid
172 (HMGA2F) expression was determined by RT-qPCR using the iScript Advanced cDNA
173 Synthesis Kit (BioRad, #1725037) for the reverse transcription and *HMGA2* PrimePCR
174 primer (BioRad, #qHsaCED0021425) for the qPCR. RPL17 was use as a reference gene
175 with forward ACGAAAAGCCACGAAGTATCTG and reverse
176 GACCTTGTGTCCAGCCCCAT primers.
177
178 **RNA Isolation, Library Preparation and Sequencing**

179 Total RNA from HMGA2hi and control cells was isolated from 3 different passages using
180 an RNeasy mini kit (Qiagen) and stored at −80 °C in nuclease-free water. RNA integrity
181 values from cells and tissues were determined with an Agilent 2100 Bioanalyzer
182 (ThermoFisher), and values ≥7.5 were used for library preparation and paired-end (2 ×
183 100 bp) RNA-sequencing on an Illumina NextSeq 6000 instrument (Illumina). Libraries
184 were prepared using a Kapa RNA HyperPrep kit with ribosomal reduction, pooled, and
185 sequenced on flowcells to yield approximately 50–60 million reads/sample. Raw fastq
186 files were deposited in the NCBI Gene Expression Omnibus (GSEXXXX).
187
188 **RNA Seq Analysis**
189 RNA-seq reads were trimmed for quality and adapters using TrimGalore (version 0.6.4),
190 and quality trimmed reads were assessed with FastQC (version 0.11.7). Trimmed reads
191 were mapped to Homo sapiens genome assembly GRCh38 (hg38) using STAR (version
192 2.6.0c). Reads overlapping Ensembl annotations (version 99) were quantified with STAR
193 prior to model-based differential expression analysis using the edgeR-robust method.
194 Genes with low counts per million (CPM) were removed using the filterByExpr function
195 from edgeR (28). Scatterplots of two selected principal components were constructed
196 with the PCAtools R package (version 2.12.0) to verify sample separation prior to
197 statistical testing. Generalized linear models were used to determine if principal
198 components were significantly associated with cell type. Differential expression was
199 calculated by comparing HMGA2hi to control cells and HMGA2F to MyoN tissues and
200 differentially expressed genes (DEGs) were visualized with a volcano plot generated
201 using the EnhancedVolcano (version 1.18.0) package in R. Boxplots of the $\log_2$(CPM)

202 values were generated using the R package ggplot2 (version 3.4.4). Venn diagrams of
203 the overlapping DEG of the cells (HMGA2hi vs. control) and DEG of tissues (HMGA2F
204 vs. MyoN) were constructed using the eulerr package (version 7.0.0). Gene set
205 enrichment and overrepresentation analyses were completed with clusterProfiler
206 package (version 4.8.2) using MSigDB (29) and DOSE (30) reference databases.
207
208 **Statistics and Reproducibility**
209 Bioinformatic statistics were performed using the listed packages in R (version 4.3.2).
210 DEGs in the RNA-seq of HMGA2hi vs. control and the HMGA2F vs. MyoN were identified
211 as those having a Benjamini–Hochberg FDR corrected $p < 0.05$ (31). Data with unequal
212 variances were log-transformed, and the homogeneity of variances was verified before
213 the completion of analyses. Hypergeometric testing was performed using the function
214 phyper from R. For the colony formation assays, comparison of two means was
215 performed with a two-sided student t test, and significance was determined at $p < 0.05$
216 after confirming normal distribution using GraphPad Prism (version 10.2.2).
217
218

Results

HMGA2 overexpression increases myometrium cell plasticity.

Our initial objective was to examine whether *HMGA2* overexpression enhances stem cell/progenitor-like properties in a differentiated myometrium cell line. Primary myometrial cells were immortalized (MyoN-TERT) and subsequently depleted for stem/progenitor cells via fluorescence-activated cell sorting (FACS) (Figure 1a and 1b), utilizing the cell surface marker CRIP1, a putative myometrium stem/progenitor cell marker recently identified in our laboratory (26). Among the MyoN-TERT cells, 73.7% of the MyoN-TERT were CRIP1^{-} (Figure 1b) and were sorted for downstream experiments. CRIP1-depleted MyoN-TERT cells were then transduced with either *HMGA2* (HMGA2hi) or a control virus. We performed proliferation, colony formation and differentiation assays to determine stem cell-like properties induced by *HMGA2* overexpression. The proliferation assay revealed reduced cell proliferation in HMGA2hi cells compared to the control cells (Figure 1c). Quantitative analysis of the colony formation assay showed a significant 4.3-fold increase in the number of cells with self-renewal capacity in the HMGA2hi cells (mean = 128) compared to control cells (mean = 30) (Figure 1d). While both control and HMGA2hi cells exhibited lipid droplets when cultured in adipogenic media (Figure 1e), indicating their ability to differentiate into adipocytes, staining in the osteogenic differentiation assay revealed alkaline phosphatase activity only in HMGA2hi cells and not in control cells when cultured in osteogenic media (Figure 1f).

Stem cell-like transcriptomic signature is induced by *HMGA2* overexpression in myometrial cells.

242 Principal component analysis (PCA) plot of HMGA2hi and control cells revealed sample
243 separation primarily along principal component (PC) 2, which accounted for 29% of the
244 variance, indicating that *HMGA2* overexpression induces significant transcriptomic
245 changes in the immortalized myometrial cells (Figure 2a). Notably, the most prominent
246 separation observed in PC1, representing 49% of the variance, was associated with cell
247 passage and was thus incorporated as a co-variable for subsequent analyses. A volcano
248 plot depicted a total of 709 DEGs with an FDR < 0.05 in HMGA2hi compared to control
249 cells (Figure 2b). Among these, 304 genes were downregulated, while 405, including
250 *HMGA2* (Log_2FC = 2.2, FDR = 3.4 x 10^{-55}), were upregulated (Table S1). Additionally,
251 *CSPG4*, a gene associated with mesenchymal stem cells (26), was significantly
252 upregulated (Log_2FC = 0.4, FDR = 3.1 x 10^{-2}) in HMGA2hi cells (Figure 2c). However,
253 other putative myometrial stem/progenitor-related genes such as *PDGFRβ* (32), *SUSD2*
254 (32), *MCAM* (32), *CD44* (19), and *CD34* (33) did not exhibit significantly (FDR > 0.05)
255 higher expression in the HMGA2hi cells compared to control cells, except for *ITGA6* (also
256 known as *CD49f*) (33), which was downregulated (Figure S1). Gene set enrichment
257 analysis using the MSigDB (29) for cell type signature gene sets (C8) identified 62
258 significant pathways (FDR <0.05) (Table S2), including two upregulated stem cell gene
259 sets (Figure 2d).
260

261 **HMGA2hi cells share a similar transcriptomic profile with HMGA2F.**

262 Disease ontology analysis using DisGeNET database (34) of HMGA2hi compared to
263 control cells showed enrichment of multiple diseases associated with fibrosis, including
264 renal fibrosis (Unified Medical Language System: C0151650), chronic kidney disease

265 (C1561643), lung diseases (C0024115), and fatty liver disease (C4529962) (Figure 3a).
266 In addition, the uterine fibroids gene set (C0042133) was also enriched by the
267 overexpression of *HMGA2*. To determine if the transcriptomic profile of HMGA2hi cells is
268 similar to uterine fibroids, we sequenced RNA from *HMGA2* fibroids (HMGA2F) and
269 myometrium from non-fibroid patients (MyoN). The transcriptomes of the tissues
270 separated in PC1 with 41% of variance (Figure 3b) in the PCA plot. 12,929 DEG, 6,788
271 downregulated and 6,141 upregulated including *HMGA2* (Log_2FC = 11.4, FDR = 3.61 x
272 10^{-34}), were found in HMGA2F compared to MyoN samples (Table S3). A Venn diagram
273 of the DEGs from the HMGA2hi vs control cells and from the HMGA2F vs MyoN tissues
274 showed a significant overlap with 61% of the DEGs (n = 434) from the cells also found in
275 the tissue comparison HMGA2F vs. MyoN (Figure 3c).
276
277 ***HMGA2*hi cells and HMGA2F tissue had dysregulated extracellular matrix.**
278 Overrepresentation analysis of gene ontology molecular functions (35) revealed co-
279 regulated significant pathways enriched in both cells (HMGA2hi compared to control) and
280 tissue (HMGA2F compared to MyoN). These co-regulated pathways included functions
281 related to ECM, cell adhesion, stem cell differentiation, and muscle development (Figure
282 4a). To assess ECM dysregulation in both cell lines and primary cells, 3D culture followed
283 by trichrome staining analyses were conducted. Spheroids derived from primary fibroid
284 cells exhibited well-defined round structures, and collagen quantification revealed
285 significantly higher deposition in fibroid spheroids compared to spheroids from MyoN cells
286 (Figure 4b and 4c). Similarly, spheroids derived from HMGA2hi cells displayed well-

287 defined structures and exhibited significantly elevated collagen deposition compared to
288 their respective controls.
289
290 **Discussion**
291 The clonal origin of fibroids from a mutated myometrial stem cell is the prevailing
292 hypothesis for the etiology of the disease (21,22). In this study, we explored an alternative
293 hypothesis, i.e., that fibroids could arise from a mutated differentiated myometrial cell
294 (Figure 5). Since elevated *HMGA2* expression in adult tissues is frequently encountered
295 in neoplastic processes, it is thought to be a key driver of tumorigenesis, but the
296 mechanisms involved are largely unknown. HMGA2 plays a crucial role in embryonic
297 development and stem cell maintenance (13). While its expression is generally low in
298 most normal adult tissues (14), it is notably elevated in some tumors (13,15,16), including
299 HMGA2 fibroids, one of the major fibroid subtypes (17,18). We exploited these HMGA2
300 characteristics to investigate whether its overexpression was sufficient to induce a fibroid-
301 like phenotype in myometrial cell culture models. We found that overexpression of
302 *HMGA2* induces a more stem cell-like phenotype in HMGA2hi cells characterized by
303 decreased cell proliferation, enhanced colony formation, and enhanced differentiation into
304 osteocytes and adipocytes compared to control cells, all of which support a role for
305 HMGA2 in inducing stem cell properties in myometrial cells. Whether similar stem cell like
306 properties can be induced in myometrial cells with other fibroid subtype genetic
307 alterations, for example with *MED12* mutation or *HMGA1* overexpression, remains to be
308 determined.

309 Increased colony formation in myometrial cells overexpressing *HMGA2* compared
310 to control cells has been previously reported (12), but whether the parental cells used in
311 the study were first depleted of putative myometrial stem cells was not described.
312 Contradictory findings regarding cell proliferation were also noted with this previous report
313 and the current study. While our study indicated decreased proliferation, Li, *et al.* reported
314 a higher Ki67 proliferation index in *HMGA2*-overexpressing cells compared to the control
315 cells. This discrepancy may be attributed to differences in the methods used to assess
316 proliferation. In our study, we employed a continuous monitoring approach over a 10-day
317 period in a 2D model, whereas the study by Li et al. assessed Ki67 staining at a specific
318 time point in a 3D culture. The significance of this discrepancy is not clear and requires
319 further investigation.

320 Transcriptomic analysis of HMGA2hi cells further revealed increased stemness, as
321 evidenced by a significant increase in the expression of another mesenchymal marker,
322 *CSPG4* (26), along with enrichment for two stem-related pathways. Several studies have
323 demonstrated the capability of myometrial and fibroid stem/progenitor cells to differentiate
324 into various cell types (26,32,36). However, to our knowledge, no previous reports have
325 investigated whether myometrium and fibroid cells depleted for stem/progenitor cells
326 retain the ability to differentiate into other cell types. Our findings suggest that the
327 myometrial cell line can differentiate into adipocytes, and notably, can differentiate into
328 osteocytes only when *HMGA2* is overexpressed, suggesting that *HMGA2*, a gene
329 expressed by stem cells, enhances cell plasticity in myometrium cells.

330 In addition to the acquisition of the stem-like properties induced by *HMGA2*
331 overexpression in the myometrial cell line, we have also uncovered significant

332 transcriptomic alterations associated with *HMGA2* overexpression, indicating a pivotal
333 role for HMGA2 in modulating key cellular processes in fibroid cells. HMGA2hi cells
334 showed a similar transcriptomic profile to HMGA2F based on DEGs. Disease ontology
335 analysis utilizing the sequencing data unveiled that HMGA2hi cells were enriched for the
336 uterine fibroid gene set and over 60% of the DEGs in the HMGA2hi cells overlapped with
337 those identified in HMGA2F tissues. This convergence in transcriptomic signatures
338 between HMGA2hi cells and HMGA2F tissues underscores the potential role of *HMGA2*
339 dysregulation in driving fibroid pathogenesis and further supports its involvement in
340 modulating gene expression patterns associated with fibroid development. Mas *et al*.
341 reported that overexpression of the truncated form of *HMGA2* in primary myometrial side
342 population and primary myometrium cells resulted in the formation of leiomyoma-like
343 tissue with similarities, including ECM content, to human fibroids in a xenotransplantation
344 *in vivo* model (11). Dysregulated ECM pathways were also found in the gene ontology
345 analysis in both HMGA2hi cells and HMGA2F tissues compared to their respective
346 controls. Additionally, collagen deposition was significantly increased with HMGA2hi cells
347 in a spheroid model, mimicking observations made using primary fibroid tissue cells
348 where collagen deposition is a key histological characteristic of uterine fibroids (37).
349 In conclusion, *HMGA2* emerges as a central player in the pathogenesis of uterine
350 fibroids, influencing cellular behavior and gene expression profiles. Our study sheds light
351 on the intricate interplay between HMGA2 and fibroid development, revealing its capacity
352 to induce a stem cell-like phenotype, alter transcriptomic profiles leading to significant
353 changes in ECM pathways, and mirror observations in fibroid tissues, all of which
354 underscores its role in driving fibroid pathogenesis. Further research is needed to

355 investigate if MED12mt cells exhibit similar characteristics, confirming the theory that the
356 mutation leads to a stem cell-like phenotype. Overall, our study highlights the significance
357 of HMGA2 in uterine fibroid biology and lays the groundwork for future studies aimed at
358 improving clinical management and treatment outcomes for individuals affected by this
359 condition.
360

361 **Study approval**

362 The use of human tissue specimens was approved by the Spectrum Health Systems and
363 Michigan State University Institutional Review Boards (MSU IRB Study ID:
364 STUDY00003101, SR IRB #2017-198) as secondary use of biobank materials.
365

366 **Author contributions**

367 Experimental design (ENP, JMT), collected data and performed experiments (ENP, TJC,
368 LP, AB, MAOB), analyzed data (ENP, JMT), wrote/reviewed manuscript (ENP, TJC, LP,
369 AB, MAOB, ATF, JMT).
370

371 **Acknowledgments**

372 We would like to thank the patients who consented for the study, the Spectrum Health
373 Systems Universal Biorepository staff, and the Van Andel Research Institute Genomics
374 Core for their help with the study. This work was supported in part through computational
375 resources and services provided by the Institute for Cyber-Enabled Research at Michigan
376 State University.

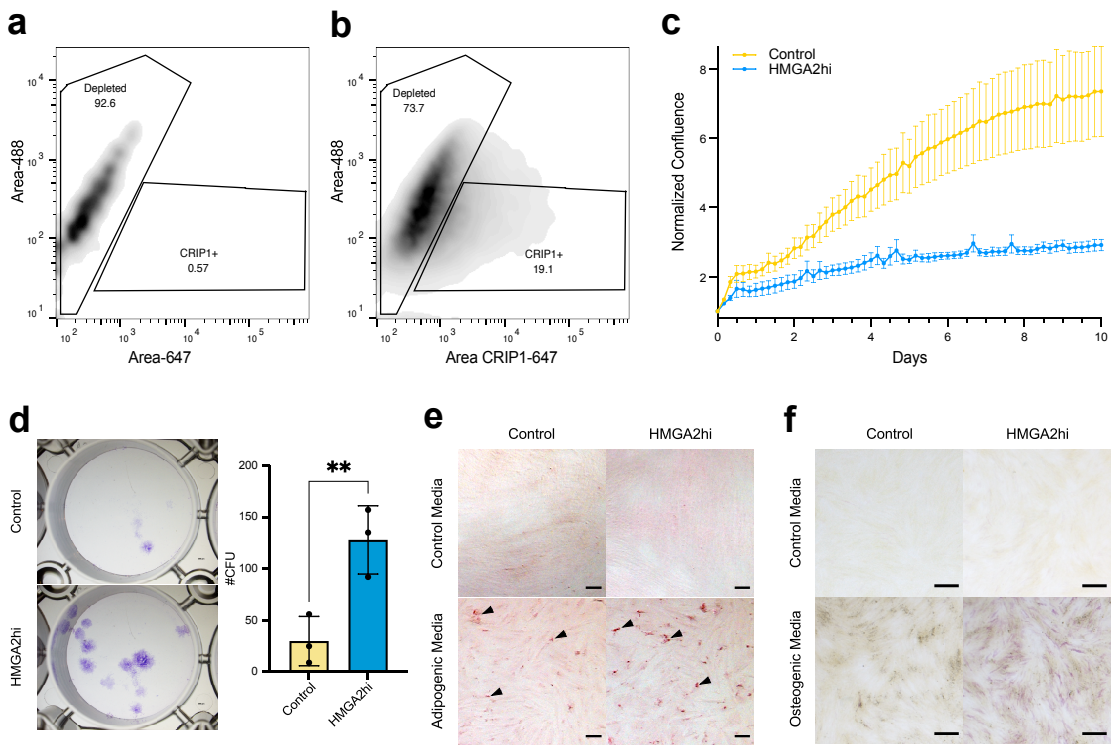
377 **References**

378 1. Stewart EA. Uterine fibroids. *Lancet*. 2001;357(9252):293-298.
379 2. Borahay MA, Al-Hendy A, Kilic GS, Boehning D. Signaling Pathways in
380 Leiomyoma: Understanding Pathobiology and Implications for Therapy. *Mol Med*.
381 2015;21(1):242-256.
382 3. Stewart EA, Laughlin-Tommaso SK, Catherino WH, Lalitkumar S, Gupta D,
383 Vollenhoven B. Uterine fibroids. *Nat Rev Dis Primers*. 2016;2:16043.
384 4. Paul EN, Burns GW, Carpenter TJ, Grey JA, Fazleabas AT, Teixeira JM.
385 Transcriptome Analyses of Myometrium from Fibroid Patients Reveals Phenotypic
386 Differences Compared to Non-Diseased Myometrium. *Int J Mol Sci*. 2021;22(7).
387 5. George JW, Fan H, Johnson B, Carpenter TJ, Foy KK, Chatterjee A, Patterson AL,
388 Koeman J, Adams M, Madaj ZB, Chesla D, Marsh EE, Triche TJ, Shen H, Teixeira
389 JM. Integrated Epigenome, Exome, and Transcriptome Analyses Reveal Molecular
390 Subtypes and Homeotic Transformation in Uterine Fibroids. *Cell Rep*.
391 2019;29(12):4069-4085 e4066.
392 6. Paul EN, Grey JA, Carpenter TJ, Madaj ZB, Lau KH, Givan SA, Burns GW,
393 Chandler RL, Wegienka GR, Shen H, Teixeira JM. Transcriptome and DNA
394 methylome analyses reveal underlying mechanisms for the racial disparity in
395 uterine fibroids. *JCI Insight*. 2022;7(20).
396 7. Mehine M, Kaasinen E, Heinonen HR, Makinen N, Kampjarvi K, Sarvilinna N,
397 Aavikko M, Vaharautio A, Pasanen A, Butzow R, Heikinheimo O, Sjoberg J,
398 Pitkanen E, Vahteristo P, Aaltonen LA. Integrated data analysis reveals uterine
399 leiomyoma subtypes with distinct driver pathways and biomarkers. *Proceedings of*
400 *the National Academy of Sciences of the United States of America*.
401 2016;113(5):1315-1320.
402 8. Makinen N, Mehine M, Tolvanen J, Kaasinen E, Li Y, Lehtonen HJ, Gentile M, Yan
403 J, Enge M, Taipale M, Aavikko M, Katainen R, Virolainen E, Bohling T, Koski TA,
404 Launonen V, Sjoberg J, Taipale J, Vahteristo P, Aaltonen LA. MED12, the mediator
405 complex subunit 12 gene, is mutated at high frequency in uterine leiomyomas.
406 *Science*. 2011;334(6053):252-255.
407 9. Nilbert M, Heim S, Mandahl N, Floderus UM, Willen H, Mitelman F. Characteristic
408 chromosome abnormalities, including rearrangements of 6p, del(7q), +12, and
409 t(12;14), in 44 uterine leiomyomas. *Hum Genet*. 1990;85(6):605-611.
410 10. Klemke M, Meyer A, Nezhad MH, Bartnitzke S, Drieschner N, Frantzen C, Schmidt
411 EH, Belge G, Bullerdiek J. Overexpression of HMGA2 in uterine leiomyomas points
412 to its general role for the pathogenesis of the disease. *Genes Chromosomes*
413 *Cancer*. 2009;48(2):171-178.
414 11. Mas A, Cervello I, Fernandez-Alvarez A, Faus A, Diaz A, Burgues O, Casado M,
415 Simon C. Overexpression of the truncated form of High Mobility Group A proteins
416 (HMGA2) in human myometrial cells induces leiomyoma-like tissue formation. *Mol*
417 *Hum Reprod*. 2015;21(4):330-338.
418 12. Li Y, Qiang W, Griffin BB, Gao T, Chakravarti D, Bulun S, Kim JJ, Wei JJ. HMGA2-
419 mediated tumorigenesis through angiogenesis in leiomyoma. *Fertil Steril*.
420 2020;114(5):1085-1096.

421 13. Parisi S, Piscitelli S, Passaro F, Russo T. HMGA Proteins in Stemness and
422 Differentiation of Embryonic and Adult Stem Cells. *Int J Mol Sci*. 2020;21(1).
423 14. Vignali R, Marracci S. HMGA Genes and Proteins in Development and Evolution.
424 *Int J Mol Sci*. 2020;21(2).
425 15. Mansoori B, Mohammadi A, Ditzel HJ, Duijf PHG, Khaze V, Gjerstorff MF,
426 Baradaran B. HMGA2 as a Critical Regulator in Cancer Development. *Genes*
427 *(Basel)*. 2021;12(2).
428 16. Li Y, Zhao Z, Xu C, Zhou Z, Zhu Z, You T. HMGA2 induces transcription factor Slug
429 expression to promote epithelial-to-mesenchymal transition and contributes to
430 colon cancer progression. *Cancer Lett*. 2014;355(1):130-140.
431 17. Nibert M, Heim S. Uterine leiomyoma cytogenetics. *Genes Chromosomes Cancer*.
432 1990;2(1):3-13.
433 18. Berta DG, Kuisma H, Valimaki N, Raisanen M, Jantti M, Pasanen A, Karhu A,
434 Kaukomaa J, Taira A, Cajuso T, Nieminen S, Penttinen RM, Ahonen S, Lehtonen
435 R, Mehine M, Vahteristo P, Jalkanen J, Sahu B, Ravantti J, Makinen N, Rajamaki
436 K, Palin K, Taipale J, Heikinheimo O, Butzow R, Kaasinen E, Aaltonen LA.
437 Deficient H2A.Z deposition is associated with genesis of uterine leiomyoma.
438 *Nature*. 2021;596(7872):398-403.
439 19. Mas A, Nair S, Laknaur A, Simon C, Diamond MP, Al-Hendy A. Stro-1/CD44 as
440 putative human myometrial and fibroid stem cell markers. *Fertil Steril*.
441 2015;104(1):225-234 e223.
442 20. Yin P, Ono M, Moravek MB, Coon JSt, Navarro A, Monsivais D, Dyson MT,
443 Druschitz SA, Malpani SS, Serna VA, Qiang W, Chakravarti D, Kim JJ, Bulun SE.
444 Human uterine leiomyoma stem/progenitor cells expressing CD34 and CD49b
445 initiate tumors in vivo. *J Clin Endocrinol Metab*. 2015;100(4):E601-606.
446 21. Canevari RA, Pontes A, Rosa FE, Rainho CA, Rogatto SR. Independent clonal
447 origin of multiple uterine leiomyomas that was determined by X chromosome
448 inactivation and microsatellite analysis. *Am J Obstet Gynecol*. 2005;193(4):1395-
449 1403.
450 22. Zhang P, Zhang C, Hao J, Sung CJ, Quddus MR, Steinhoff MM, Lawrence WD.
451 Use of X-chromosome inactivation pattern to determine the clonal origins of uterine
452 leiomyoma and leiomyosarcoma. *Hum Pathol*. 2006;37(10):1350-1356.
453 23. Bulun SE. Uterine fibroids. *N Engl J Med*. 2013;369(14):1344-1355.
454 24. Hayer A, Shao L, Chung M, Joubert LM, Yang HW, Tsai FC, Bisaria A, Betzig E,
455 Meyer T. Engulfed cadherin fingers are polarized junctional structures between
456 collectively migrating endothelial cells. *Nat Cell Biol*. 2016;18(12):1311-1323.
457 25. Dull T, Zufferey R, Kelly M, Mandel RJ, Nguyen M, Trono D, Naldini L. A third-
458 generation lentivirus vector with a conditional packaging system. *J Virol*.
459 1998;72(11):8463-8471.
460 26. Paul EN, Carpenter TJ, Fitch S, Sheridan R, Lau KH, Arora R, Teixeira JM.
461 Cysteine-rich intestinal protein 1 is a novel surface marker for human myometrial
462 stem/progenitor cells. *Commun Biol*. 2023;6(1):686.
463 27. Wiwatpanit T, Murphy AR, Lu Z, Urbanek M, Burdette JE, Woodruff TK, Kim JJ.
464 Scaffold-Free Endometrial Organoids Respond to Excess Androgens Associated
465 With Polycystic Ovarian Syndrome. *J Clin Endocrinol Metab*. 2020;105(3):769-
466 780.

28. Zhou X, Lindsay H, Robinson MD. Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic Acids Res*. 2014;42(11):e91.
29. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;102(43):15545-15550.
30. Yu G, Wang LG, Yan GR, He QY. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*. 2015;31(4):608-609.
31. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Statistical Methodology*. 1995;57(1):289-300.
32. Patterson AL, George JW, Chatterjee A, Carpenter TJ, Wolfrum E, Chesla DW, Teixeira JM. Putative human myometrial and fibroid stem-like cells have mesenchymal stem cell and endometrial stromal cell properties. *Human reproduction*. 2020;35(1):44-57.
33. Ono M, Kajitani T, Uchida H, Arase T, Oda H, Uchida S, Ota K, Nagashima T, Masuda H, Miyazaki K, Asada H, Hida N, Mabuchi Y, Morikawa S, Ito M, Bulun SE, Okano H, Matsuzaki Y, Yoshimura Y, Maruyama T. CD34 and CD49f Double-Positive and Lineage Marker-Negative Cells Isolated from Human Myometrium Exhibit Stem Cell-Like Properties Involved in Pregnancy-Induced Uterine Remodeling. *Biology of reproduction*. 2015;93(2):37.
34. Pinero J, Sauch J, Sanz F, Furlong LI. The DisGeNET cytoscape app: Exploring and visualizing disease genomics data. *Comput Struct Biotechnol J*. 2021;19:2960-2967.
35. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou LP, Mi H. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci*. 2022;31(1):8-22.
36. Banerjee S, Xu W, Chowdhury I, Driss A, Ali M, Yang Q, Al-Hendy A, Thompson WE. Human Myometrial and Uterine Fibroid Stem Cell-Derived Organoids for Intervening the Pathophysiology of Uterine Fibroid. *Reprod Sci*. 2022;29(9):2607-2619.
37. Leppert PC, Jayes FL, Segars JH. The extracellular matrix contributes to mechanotransduction in uterine fibroids. *Obstet Gynecol Int*. 2014;2014:783289.

503 **Figures**



504

505 **Figure 1. *HMGA2* overexpression increases myometrial cell plasticity.** Scatter plot
506 of the gating strategy used for cell sorting of depleted stem/progenitor MyoN-TERT, using
507 the unstained sample as control **(a)** and the full stain sample to isolate the depleted cell
508 population **(b)**. Quantification of the cell proliferation from the cell-by-cell function of the
509 Incucyte, measuring the area occupied by cells (cm^2/image) for each image scan
510 normalized to time = 0 (n = 3, error bars represent standard of the mean, sem) (**c**). The
511 yellow curve represents the control MyoN-TERT cells, and the blue curve the HMGA2hi
512 cells, error bars represent standard error of the mean. Representative images of colonies
513 formed by the control and HMGA2hi cells and plot of colony forming efficiency
514 represented as %CFUs (#CFU/cells seeded × 100) (n = 3, error bars = sem) **(d)**.

515 Representative images of control and HMGA2hi cells grown in control growth media and
516 adipogenic **(e)** or osteogenic **(f)** differentiation media (n = 3). Adipogenic and control
517 cultures were stained with Oil Red O (red color, black arrows), and osteogenic and control
518 cultures were stained for alkaline phosphatase activity (purple color). Scale bar for the
519 adipogenic and osteogenic assays are 500 μm. **$p<0.01$; by two-tailed student t-test.
520

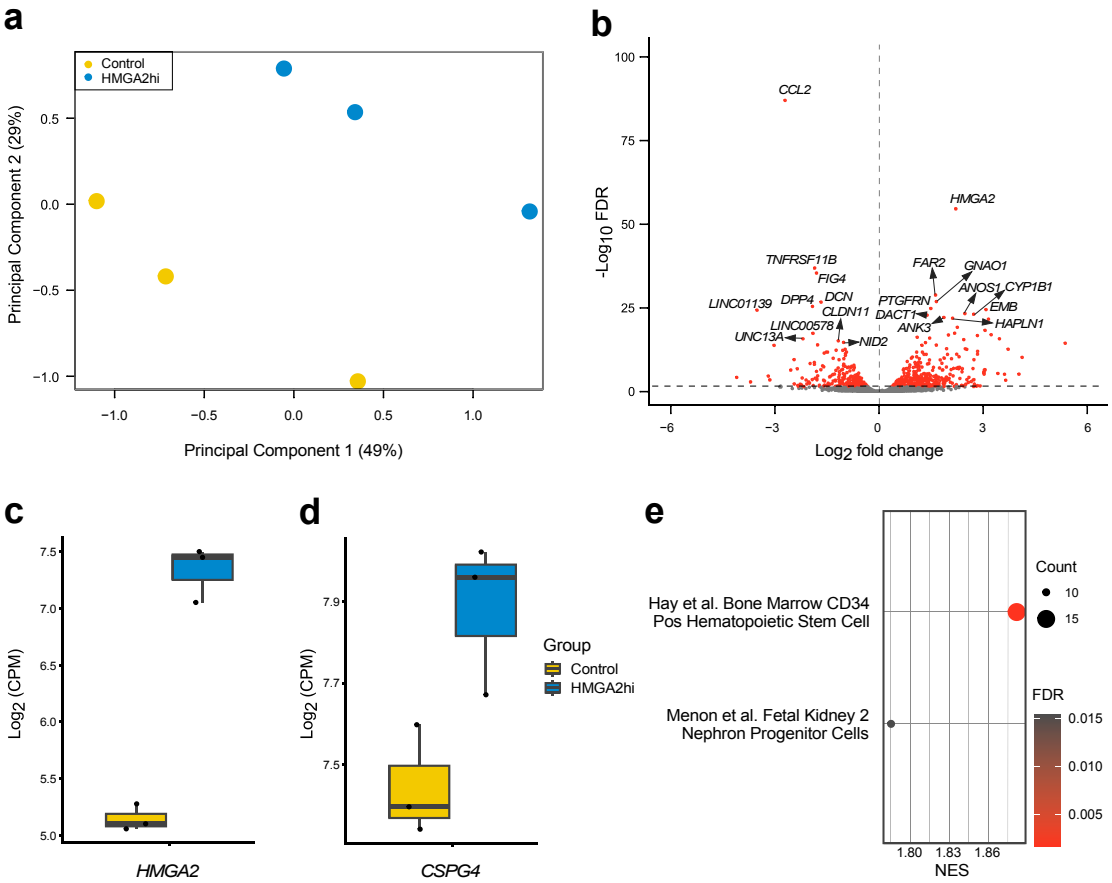
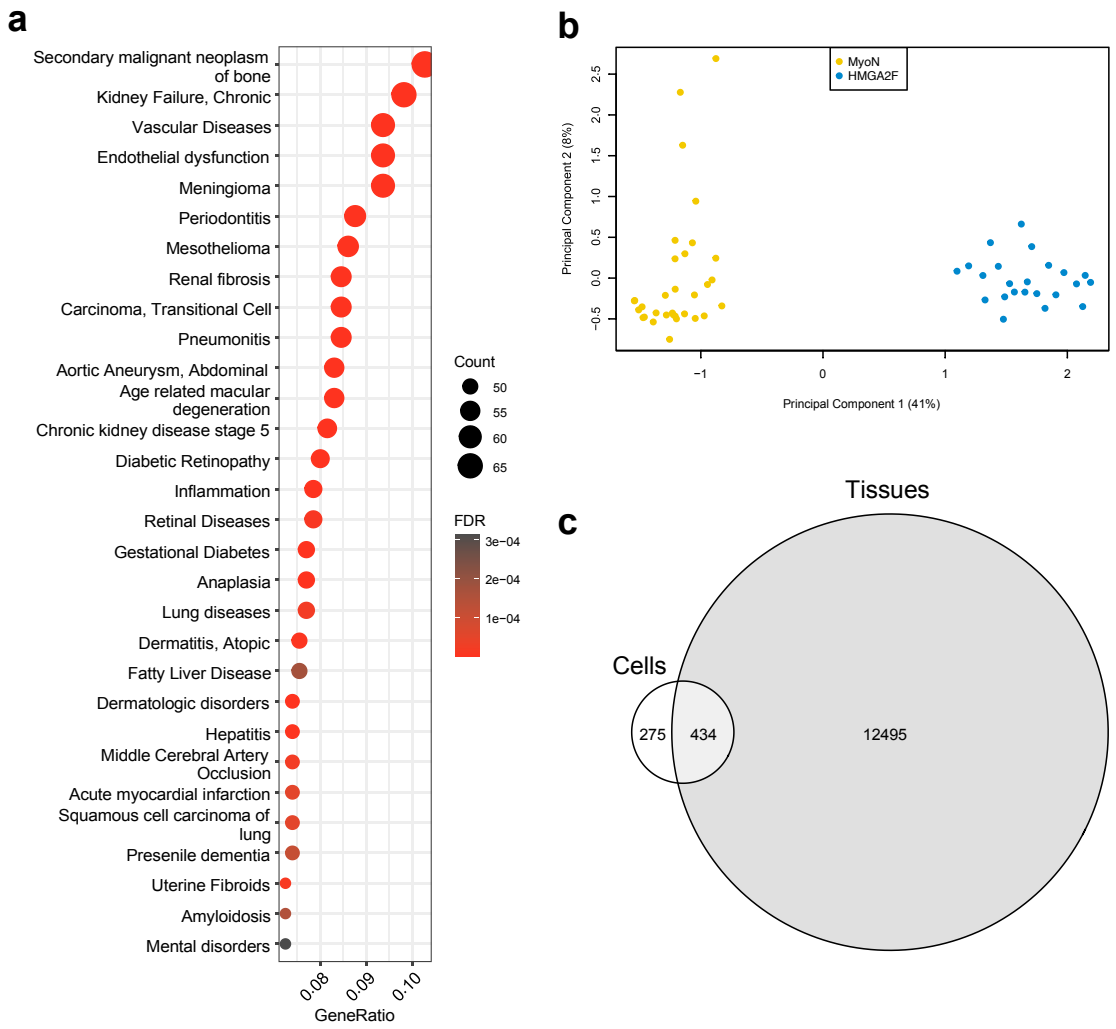


Figure 2. HMGA2hi transcriptome is enriched for stem cell markers and pathways. Principal component (PC) plot analyses for PC1 vs PC2 shows samples for control and HMGA2hi cells (n = 3/group) in two-dimensional space. Each dot represents one sample **(a)**. Volcano plots showing the DEGs from the comparison of HMGA2hi to control cells **(b)**. Those DEGs with an FDR < 0.05 are depicted as red dots, and the top 10 down- and up-regulated genes are annotated. Boxplot of two stem cells markers, *HMGA2* (**c**) and *CSPG4* (**d**), that are significantly induced (FDR<0.05) in the HMGA2hi cells compared to control cells (n = 3 each). Gene expression is shown as $\log_2$CPM. Gene set enrichment analysis (GSEA) using the cell type signature gene sets showing two upregulated stem

531 cell-related pathways in HMGA2hi cells compared to control cells (**e**). Gene count and

532 significance level are shown by the size and color of each circle.

533



534

535 **Figure 3. *HMGA2*hi cells have a similar transcriptomic profile with HMGA2 fibroid**

536 **tissues.** Disease-associated gene analysis using the DisGeNET database showing the

537 top 30 significantly enriched pathways (FDR < 0.05) (**a**). Gene count and significance

538 level are shown by the size and color of each circle. Principal component (PC) plot

539 analyses for PC1 vs PC2 show samples from MyoN (n = 32) and HMGA2F tissues (n

540 = 24) in two-dimensional space. Each dot represents one tissue sample (**b**). Venn

541 diagram illustrate the significant (hypergeometric test; $p = 2.2 \times 10^{-22}$) overlap of DEGs

542 between the HMGA2hi vs. control cells and the HMGA2F vs. MyoN tissue samples (**c**).

543

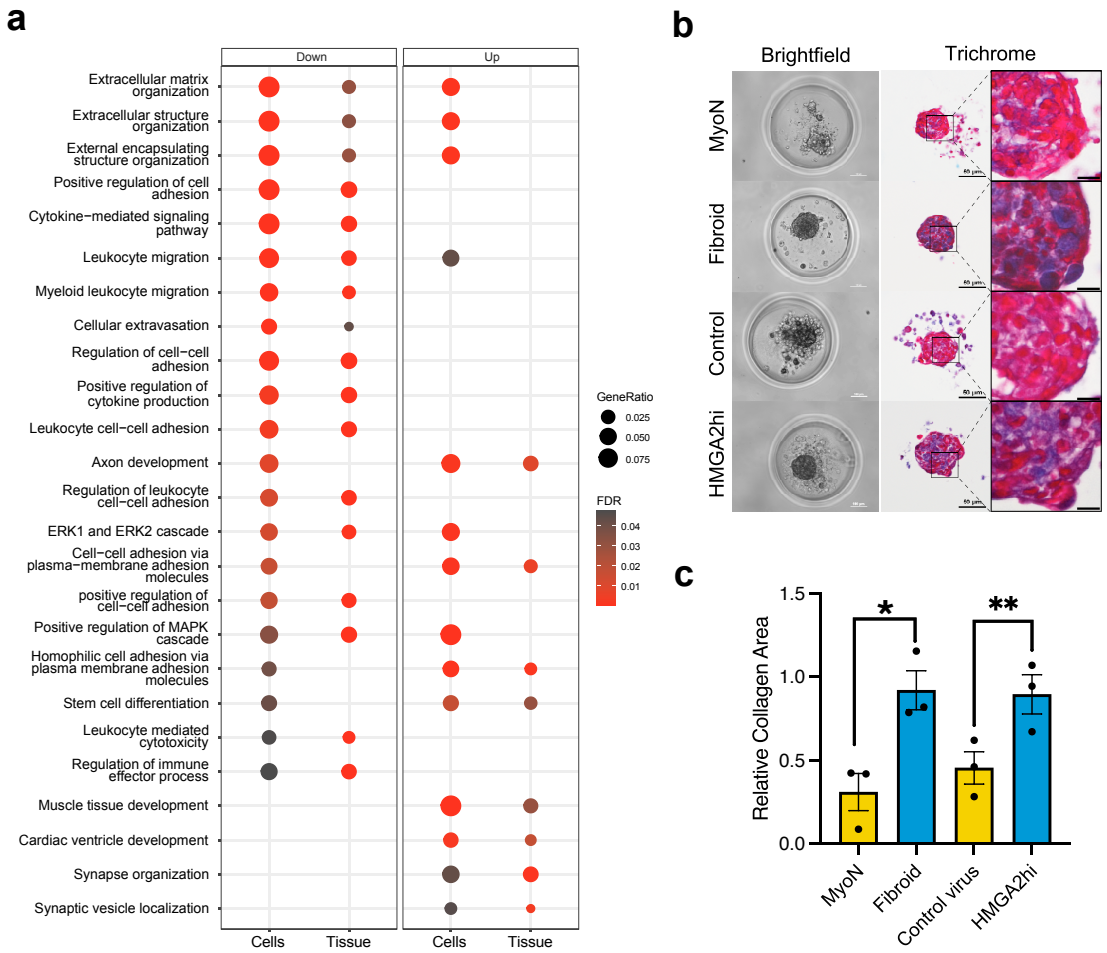
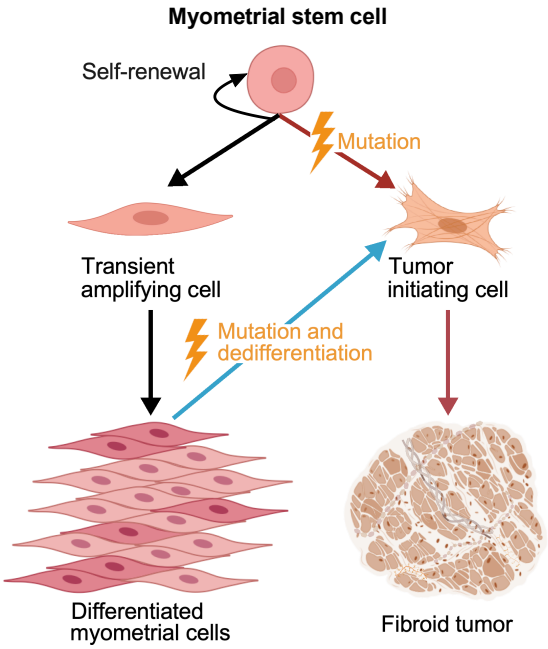


Figure 4. *HMGA2* overexpression in MyoN cells modulates extracellular matrix-associated pathways similarly to HMGA2F tissue. Gene ontology (GO) analysis using the biological processes (BP) gene sets showing the top 25 down- and up-regulated pathways in both cells (HMGA2hi vs. control) and tissues (HMGA2F vs. MyoN) (**a**). Gene ratio and significance level are shown by the size and color of each circle. Representative brightfield and trichrome-stained cross-section images of spheroids from primary cells and cell lines (**b**). The scale bar for the trichrome staining at higher magnitude is 10 μm. Collagen is stained in blue and cytoplasm in red. Relative collagen area, ratio of the

553 blue/red stain, was calculated using the trichrome-stained cross sections from primary
554 cells and cell lines (n=3 and error bars = sem) (**c**). *$p<0.05$, **$p<0.01$; by two-tailed student
555 t test.



556

557 **Figure 5. Proposed models for uterine fibroid development**. Black arrows represent
558 the traditional physiological origin of differentiated myometrial cells from a myometrial
559 stem cell. Red arrows depict the previously proposed origin for fibroid development from
560 a mutated myometrial stem cell. The blue arrow indicates our alternative hypothesis,
561 suggesting fibroid tumor initiation through the dedifferentiation of myometrial cells into a
562 more plastic tumor-initiating cell. Created with BioRender.com.