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.

40 **Introduction**

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

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

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

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

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

Methods

Myometrium cell line generation

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

HMGA2 Overexpressing and Control MyoN-TERT Cell Lines

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

108 **Proliferation Assay**

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

115

116 **Colony Formation Assay**

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

126

127 **Differentiation Assays**

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

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

154 xylenes, and a coverslip was applied using Permount mounting medium (Fisher Scientific
155 #SP15). Collagen deposition was determined using the color deconvolution ImageJ
156 plugin (version 1.53k). Relative collagen area was calculated by the ratio of blue stain
157 (collagen) over the red stain (cytoplasm).
158
159 **Tissue Collection and *HMGA2* fibroid mutation detection**
160 The use of human tissue specimens was approved by the Spectrum Health Systems and
161 Michigan State University Institutional Review Boards (MSU IRB Study ID:
162 STUDY00003101, SR IRB #2017-198) as secondary use of biobank materials. Fibroids
163 samples (HMGA2F, n = 24) and fibroid-free myometrial (MyoN, n = 32) from pre-
164 menopausal women (aged 28-55) were obtained following total hysterectomy. All patients
165 who participated in the study gave consent to donate tissue through the Spectrum Health
166 Biorepository. Human samples were processed as previously described (4). Briefly,
167 samples were washed with phosphate-buffered saline, minced, and immersed in
168 RNAlater (Sigma, Saint Louis, MO, USA) for RNA-seq analyses. *HMGA2* fibroid
169 (HMGA2F) expression was determined by RT-qPCR using the iScript Advanced cDNA
170 Synthesis Kit (BioRad, #1725037) for the reverse transcription and *HMGA2* PrimePCR
171 primer (BioRad, #qHsaCED0021425) for the qPCR. RPL17 was use as a reference gene
172 with forward ACGAAAAGCCACGAAGTATCTG and reverse
173 GACCTTGTGTCCAGCCCCAT primers.
174
175 **RNA Isolation, Library Preparation and Sequencing**

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

values were generated using the R package ggplot2 (version 3.4.4). Venn diagrams of the overlapping DEG of the cells (HMGA2hi vs. control) and DEG of tissues (HMGA2F vs. MyoN) were constructed using the eulerr package (version 7.0.0). Gene set enrichment and overrepresentation analyses were completed with clusterProfiler package (version 4.8.2) using MSigDB (29) and DOSE (30) reference databases.

Statistics and Reproducibility

Bioinformatic statistics were performed using the listed packages in R (version 4.3.2). DEGs in the RNA-seq of HMGA2hi vs. control and the HMGA2F vs. MyoN were identified as those having a Benjamini–Hochberg FDR corrected $p < 0.05$ (31). Data with unequal variances were log-transformed, and the homogeneity of variances was verified before the completion of analyses. Hypergeometric testing was performed using the function phyper from R. For the colony formation assays, comparison of two means was performed with a two-sided student t test, and significance was determined at $p < 0.05$ after confirming normal distribution using GraphPad Prism (version 10.2.2).

216 **Results**

217 ***HMGA2* overexpression increases myometrium cell plasticity.**

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

236

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

239 Principal component analysis (PCA) plot of HMGA2hi and control cells revealed sample
240 separation primarily along principal component (PC) 2, which accounted for 29% of the
241 variance, indicating that *HMGA2* overexpression induces significant transcriptomic
242 changes in the immortalized myometrial cells (Figure 2a). Notably, the most prominent
243 separation observed in PC1, representing 49% of the variance, was associated with cell
244 passage and was thus incorporated as a co-variable for subsequent analyses. A volcano
245 plot depicted a total of 709 DEGs with an FDR < 0.05 in HMGA2hi compared to control
246 cells (Figure 2b). Among these, 304 genes were downregulated, while 405, including
247 *HMGA2* (Log_2FC = 2.2, FDR = 3.4 x 10^{-55}), were upregulated (Table S1). Additionally,
248 *CSPG4*, a gene associated with mesenchymal stem cells (26), was significantly
249 upregulated (Log_2FC = 0.4, FDR = 3.1 x 10^{-2}) in HMGA2hi cells (Figure 2c). However,
250 other putative myometrial stem/progenitor-related genes such as *PDGFRβ* (32), *SUSD2*
251 (32), *MCAM* (32), *CD44* (19), and *CD34* (33) did not exhibit significantly (FDR > 0.05)
252 higher expression in the HMGA2hi cells compared to control cells, except for *ITGA6* (also
253 known as *CD49f*) (33), which was downregulated (Figure S1). Gene set enrichment
254 analysis using the MSigDB (29) for cell type signature gene sets (C8) identified 62
255 significant pathways (FDR <0.05) (Table S2), including two upregulated stem cell gene
256 sets (Figure 2d).
257
258 **HMGA2hi cells share a similar transcriptomic profile with HMGA2F.**
259 Disease ontology analysis using DisGeNET database (34) of HMGA2hi compared to
260 control cells showed enrichment of multiple diseases associated with fibrosis, including
261 renal fibrosis (Unified Medical Language System: C0151650), chronic kidney disease

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

284 defined structures and exhibited significantly elevated collagen deposition compared to
285 their respective controls.
286
287 **Discussion**
288 The clonal origin of fibroids from a mutated myometrial stem cell is the prevailing
289 hypothesis for the etiology of the disease (21,22). In this study, we explored an alternative
290 hypothesis, i.e., that fibroids could arise from a mutated differentiated myometrial cell.
291 Since elevated *HMGA2* expression in adult tissues is frequently encountered in neoplastic
292 processes, it is thought to be a key driver of tumorigenesis, but the mechanisms involved
293 are largely unknown. HMGA2 plays a crucial role in embryonic development and stem
294 cell maintenance (13). While its expression is generally low in most normal adult tissues
295 (14), it is notably elevated in some tumors (13,15,16), including HMGA2 fibroids, one of
296 the major fibroid subtypes (17,18). We exploited these HMGA2 characteristics to
297 investigate whether its overexpression was sufficient to induce a fibroid-like phenotype in
298 myometrial cell culture models. We found that overexpression of *HMGA2* induces a more
299 stem cell-like phenotype in HMGA2hi cells characterized by decreased cell proliferation,
300 enhanced colony formation, and enhanced differentiation into osteocytes and adipocytes
301 compared to control cells, all of which support a role for HMGA2 in inducing stem cell
302 properties in myometrial cells. Whether similar stem cell like properties can be induced in
303 myometrial cells with other fibroid subtype genetic alterations, for example with MED12
304 mutation or HMGA1 overexpression, remains to be determined.
305 Increased colony formation in myometrial cells overexpressing *HMGA2* compared
306 to control cells has been previously reported (12), but whether the parental cells used in

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

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

326 In addition to the acquisition of the stem-like properties induced by *HMGA2*
327 overexpression in the myometrial cell line, we have also uncovered significant
328 transcriptomic alterations associated with *HMGA2* overexpression, indicating a pivotal
329 role for HMGA2 in modulating key cellular processes in fibroid cells. HMGA2hi cells

330 showed a similar transcriptomic profile to HMGA2F based on DEGs. Disease ontology
331 analysis utilizing the sequencing data unveiled that HMGA2hi cells were enriched for the
332 uterine fibroids gene set and over 60% of the DEGs in the HMGA2hi cells overlapped
333 with those identified in HMGA2F tissues. This convergence in transcriptomic signatures
334 between HMGA2hi cells and HMGA2F tissues underscores the potential role of *HMGA2*
335 dysregulation in driving fibroid pathogenesis and further supports its involvement in
336 modulating gene expression patterns associated with fibroid development. Mas *et al*.
337 reported that overexpression of the truncated form of *HMGA2* in primary myometrial side
338 population and primary myometrium cells resulted in the formation of leiomyoma-like
339 tissue with similarities, including ECM content, to human fibroids in a xenotransplantation
340 *in vivo* model (11). Dysregulated ECM pathways were also found in the gene ontology
341 analysis in both HMGA2hi cells and HMGA2F tissues compared to their respective
342 controls. Additionally, collagen deposition was significantly increased with HMGA2hi cells
343 in a spheroid model, mimicking observations made using primary fibroid tissue cells
344 where collagen deposition is a key histological characteristic of uterine fibroids (37).

345 In conclusion, *HMGA2* emerges as a central player in the pathogenesis of uterine
346 fibroids, influencing cellular behavior and gene expression profiles. Our study sheds light
347 on the intricate interplay between HMGA2 and fibroid development, revealing its capacity
348 to induce a stem cell-like phenotype, alter transcriptomic profiles leading to significant
349 changes in ECM pathways, and mirror observations in fibroid tissues, all of which
350 underscores its role in driving fibroid pathogenesis. Further research is needed to
351 investigate if MED12mt cells exhibit similar characteristics, confirming the theory that the
352 mutation leads to a stem cell-like phenotype. Overall, our study highlights the significance

353 of HMGA2 in uterine fibroid biology and lays the groundwork for future studies aimed at
354 improving clinical management and treatment outcomes for individuals affected by this
355 condition.
356
357 **Study approval**
358 The use of human tissue specimens was approved by the Spectrum Health Systems and
359 Michigan State University Institutional Review Boards (MSU IRB Study ID:
360 STUDY00003101, SR IRB #2017-198) as secondary use of biobank materials.
361
362 **Author contributions**
363 Experimental design (ENP, JMT), collected data and performed experiments (ENP, TJC,
364 LP, AB, MAOB), analyzed data (ENP, JMT), wrote/reviewed manuscript (ENP, TJC, LP,
365 AB, MAOB, ATF, JMT).
366
367 **Acknowledgments**
368 We would like to thank the patients who consented for the study, the Spectrum Health
369 Systems Universal Biorepository staff, and the Van Andel Research Institute Genomics
370 Core for their help with the study. This work was supported in part through computational
371 resources and services provided by the Institute for Cyber-Enabled Research at Michigan
372 State University.

373 **References**

374 1. Stewart EA. Uterine fibroids. *Lancet*. 2001;357(9252):293-298.
375 2. Borahay MA, Al-Hendy A, Kilic GS, Boehning D. Signaling Pathways in
376 Leiomyoma: Understanding Pathobiology and Implications for Therapy. *Mol Med*.
377 2015;21(1):242-256.
378 3. Stewart EA, Laughlin-Tommaso SK, Catherino WH, Lalitkumar S, Gupta D,
379 Vollenhoven B. Uterine fibroids. *Nat Rev Dis Primers*. 2016;2:16043.
380 4. Paul EN, Burns GW, Carpenter TJ, Grey JA, Fazleabas AT, Teixeira JM.
381 Transcriptome Analyses of Myometrium from Fibroid Patients Reveals Phenotypic
382 Differences Compared to Non-Diseased Myometrium. *Int J Mol Sci*. 2021;22(7).
383 5. George JW, Fan H, Johnson B, Carpenter TJ, Foy KK, Chatterjee A, Patterson AL,
384 Koeman J, Adams M, Madaj ZB, Chesla D, Marsh EE, Triche TJ, Shen H, Teixeira
385 JM. Integrated Epigenome, Exome, and Transcriptome Analyses Reveal Molecular
386 Subtypes and Homeotic Transformation in Uterine Fibroids. *Cell Rep*.
387 2019;29(12):4069-4085 e4066.
388 6. Paul EN, Grey JA, Carpenter TJ, Madaj ZB, Lau KH, Givan SA, Burns GW,
389 Chandler RL, Wegienka GR, Shen H, Teixeira JM. Transcriptome and DNA
390 methylome analyses reveal underlying mechanisms for the racial disparity in
391 uterine fibroids. *JCI Insight*. 2022;7(20).
392 7. Mehine M, Kaasinen E, Heinonen HR, Makinen N, Kampjarvi K, Sarvilinna N,
393 Aavikko M, Vaharautio A, Pasanen A, Butzow R, Heikinheimo O, Sjoberg J,
394 Pitkanen E, Vahteristo P, Aaltonen LA. Integrated data analysis reveals uterine
395 leiomyoma subtypes with distinct driver pathways and biomarkers. *Proceedings of*
396 *the National Academy of Sciences of the United States of America*.
397 2016;113(5):1315-1320.
398 8. Makinen N, Mehine M, Tolvanen J, Kaasinen E, Li Y, Lehtonen HJ, Gentile M, Yan
399 J, Enge M, Taipale M, Aavikko M, Katainen R, Virolainen E, Bohling T, Koski TA,
400 Launonen V, Sjoberg J, Taipale J, Vahteristo P, Aaltonen LA. MED12, the mediator
401 complex subunit 12 gene, is mutated at high frequency in uterine leiomyomas.
402 *Science*. 2011;334(6053):252-255.
403 9. Nilbert M, Heim S, Mandahl N, Floderus UM, Willen H, Mitelman F. Characteristic
404 chromosome abnormalities, including rearrangements of 6p, del(7q), +12, and
405 t(12;14), in 44 uterine leiomyomas. *Hum Genet*. 1990;85(6):605-611.
406 10. Klemke M, Meyer A, Nezhad MH, Bartnitzke S, Drieschner N, Frantzen C, Schmidt
407 EH, Belge G, Bullerdiek J. Overexpression of HMGA2 in uterine leiomyomas points
408 to its general role for the pathogenesis of the disease. *Genes Chromosomes*
409 *Cancer*. 2009;48(2):171-178.
410 11. Mas A, Cervello I, Fernandez-Alvarez A, Faus A, Diaz A, Burgues O, Casado M,
411 Simon C. Overexpression of the truncated form of High Mobility Group A proteins
412 (HMGA2) in human myometrial cells induces leiomyoma-like tissue formation. *Mol*
413 *Hum Reprod*. 2015;21(4):330-338.
414 12. Li Y, Qiang W, Griffin BB, Gao T, Chakravarti D, Bulun S, Kim JJ, Wei JJ. HMGA2-
415 mediated tumorigenesis through angiogenesis in leiomyoma. *Fertil Steril*.
416 2020;114(5):1085-1096.

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

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.

Figures

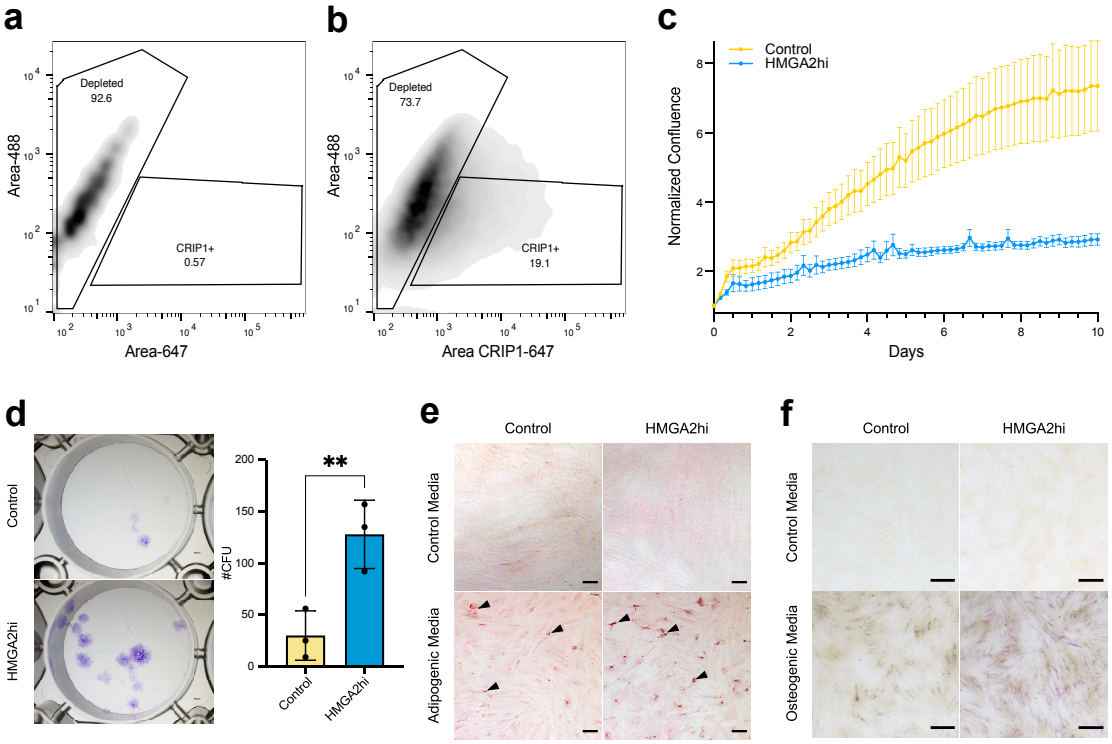


Figure 1. *HMGA2* overexpression increases myometrial cell plasticity. Scatter plot of the gating strategy used for cell sorting of depleted stem/progenitor MyoN-TERT, using the unstained sample as control **(a)** and the full stain sample to isolate the depleted cell population **(b)**. Quantification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

511 media and adipogenic **(e)** or osteogenic **(f)** differentiation media (n = 3). Adipogenic and
512 control cultures were stained with Oil Red O (red color, black arrows), and osteogenic
513 and control cultures were stained for alkaline phosphatase activity (purple color). Scale
514 bar for the adipogenic and osteogenic assays are 500 μm. **$p<0.01$; by student t-test.
515

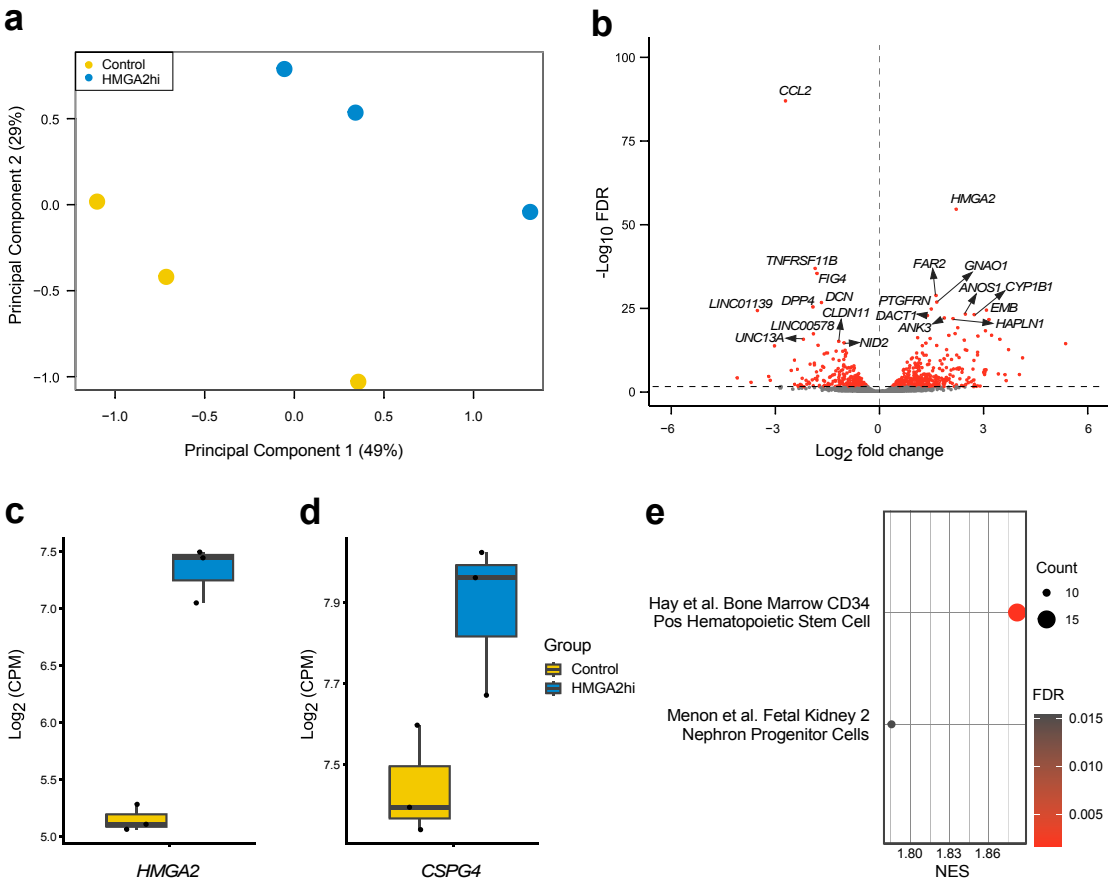
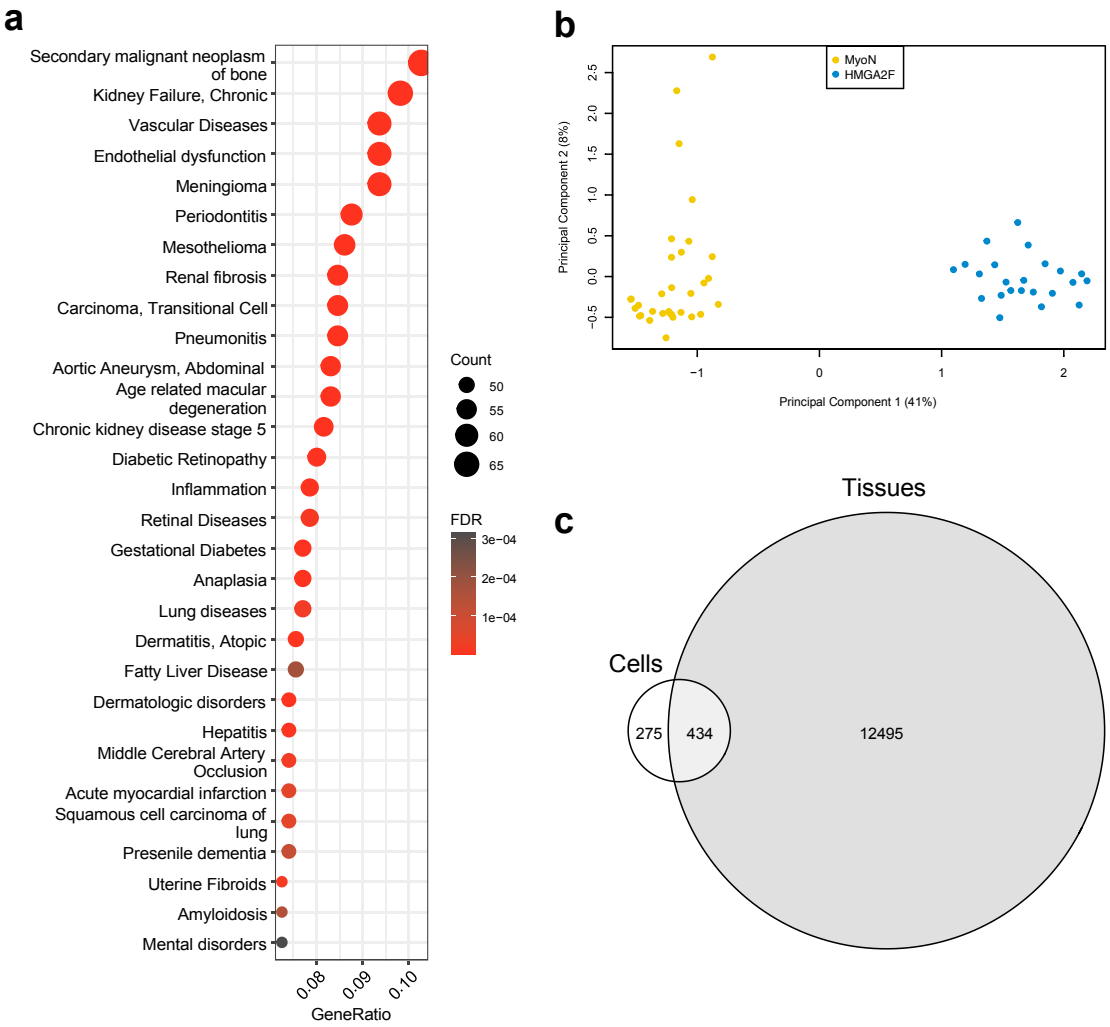


Figure 2. HMGA2hi transcriptome is enriched for stem cell markers and pathways. Principal component (PC) plot analyses for PC1 vs PC2 shows samples for control and HMGA2hi cells (n = 3/group) in two-dimensional space. Each dot represents one sample **(a)**. Volcano plots showing the DEGs from the comparison 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

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

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

528



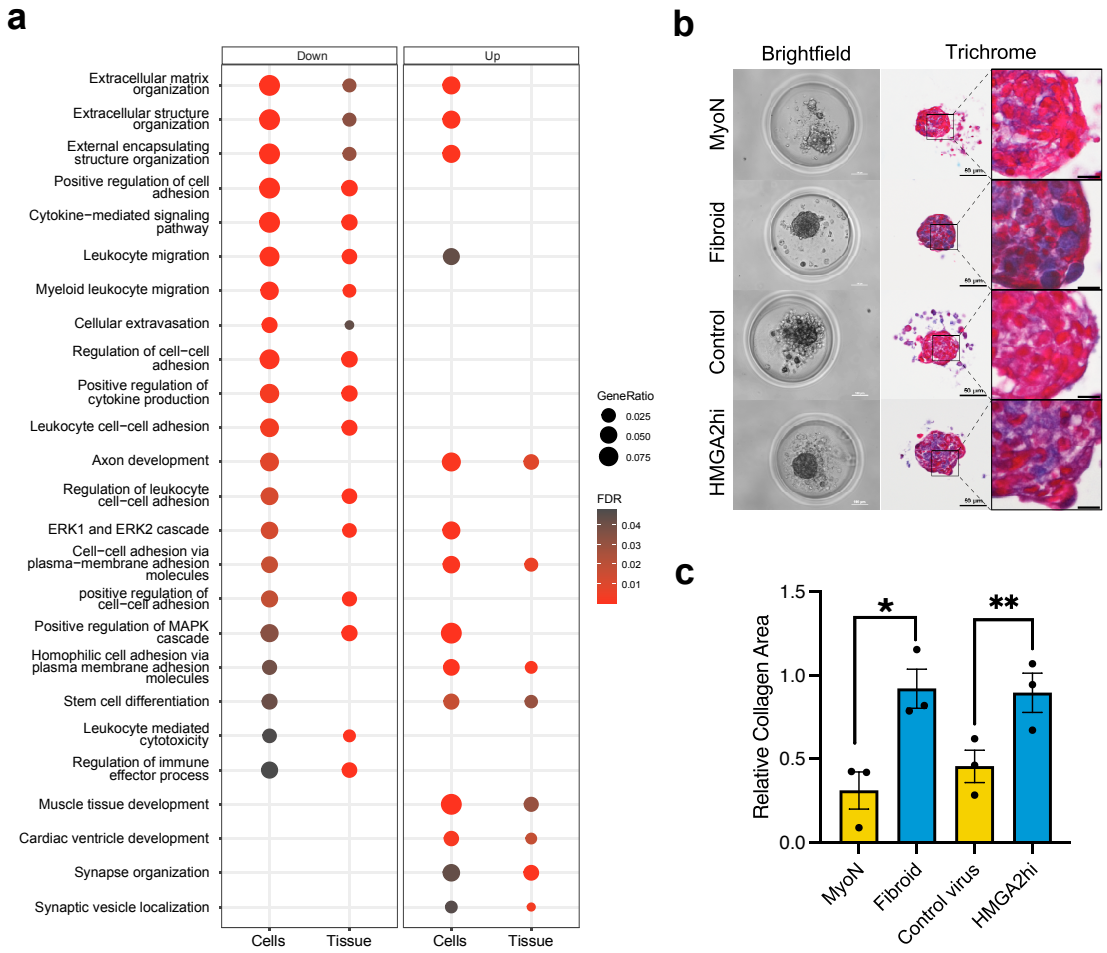
529

530 **Figure 3. *HMGA2*hi cells have a similar transcriptomic profile with HMGA2 fibroid**
531 **tissues.** Disease-associated gene analysis using the DisGeNET database showing the
532 top 30 significantly enriched pathways (FDR < 0.05) (**a**). Gene count and significance
533 level are shown by the size and color of each circle. Principal component (PC) plot
534 analyses for PC1 vs PC2 show samples from MyoN (n = 32) and HMGA2F tissues (n
535 = 24) in two-dimensional space. Each dot represents one tissue sample (**b**). Venn

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

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

538



539

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

548 blue/red stain, was calculated using the trichrome-stained cross sections from primary

549 cells and cell lines (**c**). *$p<0.05$, **$p<0.01$; by student t test.