

1 Single-cell characterization of menstrual fluid at 2 homeostasis and in endometriosis

3 Authors

4 Petra C. Schwalie^{1,2}, Cemsal Bafligil³, Julie Russeil¹, Magda Zachara¹, Marjan Biocanin¹,
5 Daniel Alpern¹, Evelin Aasna⁴, Bart Deplancke¹, Geraldine Canny⁵, Angela Goncalves^{3,4}

6 Affiliations

7 ¹Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life
8 Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015, Lausanne,
9 Switzerland.

10 ²Current: Roche Innovation Center Basel, Pharma Research and Early Development, Basel,
11 CH-4070, Switzerland

12 ³Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10
13 1SA, UK.

14 ⁴German Cancer Research Center (DKFZ), Somatic Evolution and Early Detection;
15 Heidelberg, 69120, Germany.

16 ⁵University College Dublin, Research Office, Dublin, Ireland.

17 Abstract

18 Progress in detecting and understanding endometrial conditions in women of fertile age, such
19 as endometriosis, has been hampered by the invasiveness of the sample collection procedure.
20 Menstrual fluid (MF) can be sampled non-invasively and could provide a unique opportunity
21 to study the physiological state of tissues in the reproductive system. Despite this potential, the
22 use of MF for diagnostics and research has been limited. Here we establish protocols and assess
23 the feasibility of collecting and processing MF in an outpatient setting. We characterize the
24 cellular contents of MF from 15 healthy women using flow cytometry and single-cell RNA-
25 sequencing, and demonstrate the ability to recover millions of live cells from the different
26 cellular fractions of interest (epithelial, stromal, endothelial, perivascular and blood). Through
27 computational integration of MF with endometrial samples we show that MF sampling is a
28 good surrogate for endometrial biopsy. In a proof-of-principle case-control study, we collect
29 MF from a further 7 women with a diagnosis of endometriosis and 11 healthy controls. Through
30 RNA sequencing of 93 MF samples from these women we highlight important differences
31 between *ex vivo* and cultured cells, identify impaired decidualisation, low apoptosis, high
32 proliferation, and both higher and lower inflammatory activity in different subsets of immune
33 cells as distinguishing features of endometriosis patients. Finally, we identify potential novel
34 pan-cell-type biomarkers for this neglected condition.

35 Introduction

36 Menstrual fluid (MF) is an accessible source of large numbers of live cells of high biomedical
37 interest. During menstruation, the female body sheds 86±48ml of MF (1) containing cervico-
38 vaginal secretions, blood and cells from the lining of uterus, cervix, vagina, and possibly the
39 fallopian tube and ovaries. This rich source of live cells has been previously proposed as a

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

40 scalable source of live uterine NK cells (2), mesenchymal stem cells (3–7), T-cells (8) and
41 epithelial cells (9,10) for research and therapy.

42 Menstrual fluid has also been proposed as a sampling strategy for potential diagnosis of
43 endometriosis (11–14), a chronic, currently incurable disease that affects an estimated 10% of
44 women of reproductive age. Endometriosis occurs when cells similar to those that line the
45 inside of the uterus – the eutopic endometrium – grow outside the uterine cavity in so called
46 ectopic lesions. These ectopic, endometrial-like cells can be found on the ovaries, fallopian
47 tubes, and the tissue lining the pelvis. In rare cases, they may spread beyond the pelvic region.
48 Like the endometrial cells in the uterus, these cells outside the uterus respond to the menstrual
49 cycle. This can cause severe menstrual pain, chronic lower abdominal and pelvic pain, painful
50 intercourse, and fertility issues. The exact cause of endometriosis is not well understood and
51 its diagnosis is often delayed, taking on average 8-10 years from the onset of symptoms. This
52 delay is partly due to the necessity of laparoscopic surgery for definitive diagnosis. Progress in
53 understanding the molecular basis of endometriosis has also been hampered by the
54 invasiveness of the sample collection procedure, as well as the lack of appropriate animal
55 models. Therefore, there is a real need for non-invasive, early diagnosis methods, as well as a
56 better understanding of the disease.

57 Despite its potential for research and diagnostics, quantitative parameters of MF cellular
58 composition, including cell-count, cell-viability, cell-type representation and transcriptional
59 similarity to endometrium remain only partially described (13,15–17). Moreover, it is unclear
60 whether endometriosis biomarkers identified in other tissue compartments are likewise
61 detectable in MF. To date, the identification of biomarkers for endometriosis has proven
62 challenging due to contradictory results that depend on the exact tissues being contrasted (e.g.
63 peritoneal fluid versus endometriotic lesion versus endometrial biopsy), on the phase of the
64 menstrual cycle at which the tissues were collected, or on the cell-type analysed (see (18) for a
65 systematic review of these issues).

66 MF can be acceptably and conveniently self-collected by study participants using menstrual
67 cups (19). Here, we profiled self-collected MF in a research setting at single-cell resolution to
68 ask whether MF contains a faithful representation of the cell-type composition and cellular
69 transcriptional states of the endometrium. We then compared the proportion of cell-types and
70 transcriptomes of the different cellular fractions of MF between endometriosis patients and
71 healthy controls.

72 **Results**

73 **Menstrual fluid collection is a viable sampling alternative to endometrial biopsy**

74 To quantify the cellular and molecular composition of MF, we analyzed 25 self-collected fresh
75 MF samples from 15 healthy women on day 2 of the menstrual cycle (Sup. Table 1). MF
76 samples were serially passed through 100µm and 70µm cell strainers and split into a flow-
77 through and a clump fraction. The retained cellular clumps were gently dissociated prior to
78 further processing. Flow cytometry was performed on each sample fraction staining for CD45+
79 (immune cells), CD45- EPCAM+ (epithelial) and CD45- EPCAM- (stromal) cell populations
80 (Fig. 1A).
81

82 *Volume, cell-viability and cell-type composition by fluorescence-activated cell sorting (FACS)*

83 The median volume of MF per sample was 5mL (Fig. 1B), with a potential donor effect (Fig.
84 1C-D, one-way ANOVA p-value = 0.09), but no age association (generalized linear mixed
85 model with donor random effect p-value = 0.43). Three out of 25 MF samples did not yield any
86 live cells, these samples came from two individuals, both of which subsequently donated again,
87 yielding samples with enough viable cells for processing. The remaining MF samples yielded
88 a median of 2.6 million live cells per mL of flow-through (Fig. 1E). There was no statistical
89 association (Pearson and Spearman correlation coefficient not significantly different from 0 at
90 a 5% significance level) between cell viability and processing time, with the sample with the
91 highest processing time delay after cup removal (12h) containing 63% live cells. After gating
92 out dead cells, flow-through MF was mainly composed of CD45+ cells (Fig. 1F, median 83%
93 of cells).

94
95 *Cell-type composition and comparison to endometrial biopsy by single-cell RNA-seq*

96 To characterize and quantify the cell-types found in MF, we then generated single-cell
97 transcriptomes from 7 MF samples of 4 individuals. For each sample we mixed the 3 sorted
98 populations at a 14/22/64% ratio (Fig. 2A) to obtain a balanced representation of cell-types.
99 Using established marker genes, we identified ciliated, glandular and exhausted epithelial cells,
100 mesenchymal and decidualized stromal cells, smooth muscle/endothelial cells, and immune
101 cells in all 7 samples (Fig. 2B-D, Methods). The ratio of epithelial to stromal and immune cells
102 was similar among MF samples, with epithelial cells comprising around 50% of cells (Fig. 2D).

103
104 We then asked whether the transcriptional profiles of the MF samples correspond to the
105 transcriptional profiles of freshly isolated endometrial cells by computationally integrating our
106 single-cell data with single-cells from 3 endometrial biopsies (20) in the late-secretory and
107 proliferative phases of the menstrual cycle (Fig. 2E-F). Transcriptionally, MF epithelial,
108 stromal and immune cells were more similar to their corresponding cell-types in endometrial
109 biopsies than to one another, indicating that these cells are able to retain their cellular identity
110 despite being shed (Fig. 2G). Epithelial and stromal cells exhibited the highest degree of
111 transcriptomic concordance between MF and endometrial biopsy cells, whereas the immune
112 cells were more divergent, indicating that immune cells are more sensitive to shedding or
113 experimental manipulation.

114
115 Despite the overall transcriptomic concordance, we expected expression differences between
116 MF and the endometrial biopsies to reflect differences in cycle stage. The endometrial biopsy
117 samples used came from cycle stages with high progesterone levels, which represses, among
118 others, expression of many matrix metalloproteinases (MMPs). Progesterone withdrawal at
119 menstruation triggers expression of MMPs (21), an effect mediated by cytokines including
120 tumor necrosis factor- α (TNF- α). In addition, a large number of neutrophils loaded with MMPs
121 are recruited through induction of chemokines such as interleukin 8 (CXCL8) (21). Differential
122 gene expression comparison between MF and the endometrial biopsy stromal cells revealed
123 much higher expression in MF of MMPs, members of the TNF- α signaling pathway and
124 *CXCL8* (Fig. 2H, Sup. Table 2).

125
126 In sum, our data show that sampling self-collected MF is a viable strategy for obtaining large
127 numbers of live epithelial, stromal and immune cells from the human endometrium in an
128 outpatient setting without prior scheduling. MF samples capture the main cell types of the
129 endometrium and are transcriptionally similar to cells collected via invasive biopsy.

130

131 **Properties of MF from patients with endometriosis and healthy controls**

132 To assess the potential of MF to provide insight into female reproductive tract disorders, we
133 collected 36 fresh MF samples from 7 women with a confirmed diagnosis of endometriosis and
134 11 age-matched healthy volunteers (Fig. 3A, Table 1, Sup. Table 3). Samples were processed
135 by density gradient to remove red blood cells. High volume samples were split into CD45+ and
136 CD45- fractions by MACS. For a subset of samples, we used FACS to estimate proportions of
137 CD45+ vs. CD45- cells and the CD45- fraction was further sorted into CD105+/- cells for a
138 estimation of the fraction of putative mesenchymal stem cells present (22).

Table 1: Demographic and disease characteristics of the endometriosis and healthy controls.

	Donors (n = 18)
Median age healthy controls (range)	31 (24 – 43)
Median age endometriosis (range)	40 (22 – 49)
Diagnosis	
No known endometriosis (controls)	11
Endometriosis stage 2	2
Endometriosis stage 3	2
Endometriosis stage 4	3
Known infertility healthy controls	0
Known infertility endometriosis	3

139

140

141 Once more, we observed a donor effect on volume (Fig. 3B, ANOVA p-value 0.054). However,
142 we did not observe any statistically significant differences between endometriosis and healthy
143 volunteer-derived MF samples in volume, cell number, cell viability, fraction of CD45+/- cells,
144 or fraction of mesenchymal stem cells (CD45-CD105+) (Fig. 3B,C). Moreover, MF volume
145 was not statistically significantly different between endometriosis stages at this sample size
146 (Fig. 3B). The vast majority of samples could be successfully cultured over multiple passages,
147 under standard mesenchymal cell culture conditions (Methods), irrespective of whether they
148 came from cases or controls.

149 *Cell-type composition of ex vivo and in vitro MF cells by RNA-seq*

150 To molecularly characterize both freshly isolated (*ex vivo*) and cultured (*in vitro*) samples, as
151 well as identify putative differences between endometriosis patients, we performed 3' bulk
152 RNA-sequencing on 93 samples (Fig. 4A). A principal component analysis of gene expression
153 patterns revealed two main sample groups: samples that were processed directly after isolation
154 (*ex vivo*) and those that were cultured (*in vitro*) (Fig. 4B). An over-representation analysis of
155 genes negatively correlated with PC1 (i.e. genes with higher expression in *ex vivo* samples)
156 revealed these to be highly enriched for blood cell-type markers, i.e. monocytes, dendritic and
157 macrophage-specific genes (Fig. 4C, Sup. Table 4). Cultured samples were transcriptionally
158 more homogeneous than freshly isolated samples, likely due culture-induced selection for
159 CD45- cells. Indeed, we observed that the CD45+ cell fraction typically failed to attach, in line
160 with previous reports that have shown preferential attachment of menstrual stromal cells (so-
161 called menstrual-derived stem cells, or colony-forming cells) upon culture (23). Nevertheless,
162 when comparing *in vitro* cells to CD45- *ex vivo* cells, we observed that the transcriptional
163 profiles remain distinct, suggesting major transcriptional shifts induced by cell culture (Fig.
164 4D, Sup. Table 5). In contrast, *ex vivo* CD45+ samples showed higher similarity to *ex vivo* non-

165 sorted (*unsorted*) samples (Fig. 4E), in line with the observation that most MF cells are CD45+
166 as determined by FACS.

167 To confirm the differences in cell-type composition, we used an *in silico* transcriptome
168 deconvolution method (24) and a single-cell reference dataset (13) to estimate cell-type
169 fractions from our bulk RNA-seq samples. We found that indeed *in vitro* samples are
170 preponderantly composed of stromal, smooth muscle and/or endothelial-like cells, whereas *ex*
171 *vivo* samples are preponderantly composed of myeloid cells (Fig. 4F).

172 In sum, MF collected *ex vivo* and *in vitro* display transcriptional changes consistent with
173 changes in the cell-type enriched, as well as a culture induced phenotype. *In vitro* samples only
174 capture a subset of the cell populations initially present in MF, preponderantly stromal-like
175 cells.

176 *Donor effect on transcriptomic profiles*

177 Cycle-to-cycle variability in MF transcriptomic profiles has not been previously characterised.
178 There is a concern that samples might be too variable due to unpredictability in tissue
179 breakdown and exact timing of sample collection. Leveraging on having multiple longitudinal
180 samples from the same donor (Fig. 3B), we asked if MF samples are transcriptionally stable
181 across cycles of the same donor. To do so, we looked at transcriptomic concordance among
182 CD45+ *ex vivo* samples, as these were the most homogeneous in terms of their composition
183 (Fig. 4F). Samples from the same donor clustered together more often than expected by chance,
184 suggesting that donor effects can be observed over multiple cycles (Fig. 4G).

185 **Identification of potential transcriptional biomarkers of endometriosis in MF**

186 To identify MF gene expression markers that might distinguish endometriosis from healthy
187 controls, we used three different statistical models on various subsets of the 93 RNA-seq assays
188 performed. First, we identified endometriosis biomarkers that are detectable in stromal-like
189 CD45- cells. Second, we identified biomarkers that are detectable in immune CD45+ cells.
190 Finally, we implemented a model to identify biomarkers that are detectable irrespective of cell-
191 type enriched or whether the sample is freshly isolated or cultured.

192 *Differentially expressed genes in CD45- stromal-like cells*

193 In this model, we sought to identify differentially expressed genes that can be found in *in vitro*
194 CD45- samples. We found 256 genes differentially expressed at a 5% false discovery rate (Fig.
195 5A, Sup. Table 6). These included the down-regulation of stromal decidualization genes
196 (20,25–27) in endometriosis patients when compared to healthy controls, supporting findings
197 that decidualization is compromised in cultured (28,29) and freshly isolated stromal cells from
198 endometriosis patients (11,13,14). Likewise, among genes downregulated in endometriosis
199 patients we found an enrichment for targets of estrogen receptor alpha (*ESR1*), whose
200 expression has been previously found to be suppressed in endometrial stromal cells (30).
201 Additionally, we found that inflammation related pathways (TNF α signaling and IL-2/STAT5),
202 were down-regulated in endometriosis patients, whereas epithelial to mesenchymal transition
203 (EMT) and angiogenesis were upregulated (Sup. Table 7).

204 Among down-regulated genes we found multiple endometriosis previously implicated in the
205 disease including *PENK*, *MLLT11*, *HAND2-AS1* and *PDE4B*.

206 The endogenous opioid peptide precursor proenkephalin (*PENK*) has been previously reported
207 to be up-regulated in whole-endometrium (31,32) of endometriosis patients when compared to
208 healthy controls. In contrast, *PENK* expression was found to be down-regulated in ectopic

209 stroma when compared to eutopic stroma of patients (33). The discrepancy in the direction of
210 regulation may have to do with the specific cell-types assayed, as well as the cycle phase at
211 which endometrial samples were obtained, as *PENK* expression is modulated by the menstrual
212 cycle (31). Interestingly, a decrease in opioid peptides has been previously suggested to be
213 involved in the maintenance of chronic inflammation in endometriosis (34).

214 Consistent with our findings, the expression of mixed-lineage leukemia translocated to 11
215 (*MLLT11*) was previously found to be reduced in the ectopic stromal cells of women with
216 advanced endometriosis compared to healthy controls, and its down-regulation found to be
217 associated with an increased stromal cell adhesion phenotype (35).

218 *HAND2* Antisense RNA 1 (*HAND2-ASI*), whose expression is coordinated with *HAND2* in
219 human endometrial stromal cells, was also previously found to be reduced in ectopic lesions
220 from patients compared controls, primarily expressed by stromal cells as opposed to epithelial,
221 and that its silencing impairs stromal cell decidualization (20,26,27).

222 Finally, genetic polymorphisms in cAMP-specific 3',5'-cyclic phosphodiesterase 4B (*PDE4B*)
223 were recently proposed to be implicated in the etiology of endometriosis (36).

224 Relatively fewer genes were found to be up-regulated in endometriosis patients when compared
225 to controls. Notably, the most significant up-regulated gene was fibronectin 1 (*FNI*). A
226 genome-wide meta-analysis and a subsequent targeted study have implicated single nucleotide
227 polymorphisms in *FNI* with moderate to severe endometriosis (37,38), and relaxed fibronectin
228 is being investigated as a potential target for imaging endometriotic lesions (39). At the protein
229 level, studies comparing eutopic and ectopic endometrium from patients have yielded
230 contradictory results (39). However, fibronectin levels in plasma and peritoneal fluid have
231 consistently reported elevated levels in endometriosis patients when compared to healthy
232 women (40,41).

233 *Differentially expressed genes in CD45+ immune cells*

234 We found 792 differentially expressed genes among the CD45+ *ex vivo* samples (Fig. 5B, Sup.
235 Table 8, the one sample with an abnormally high percentage of stromal cells in Fig. 4D was
236 excluded). We found a trend for lower expression of immune-related genes in CD45+ MF
237 samples originating from endometriosis patients, in particular for genes related to inflammatory
238 pathways (TNF α signaling, Inflammatory response, IFN-gamma response, Sup. Table 9).

239 This was intriguing, as the production of potentially pro-inflammatory cytokines, such as IL-
240 6, has been previously found to be up-regulated in cultured and fresh stromal cells from ectopic
241 and eutopic endometrium from patients compared to controls (13,42). To explore the activation
242 of inflammatory pathways in specific cell-types from endometriosis patients and controls, we
243 used AUCell (43) to score the activity of the MsigDB “inflammatory response” pathway in
244 each individual cell in the dataset from Shih et al. (13) (Methods).

245 In agreement with our findings, we confirmed that the activity of the inflammatory signaling
246 pathway was significantly downregulated in myeloid (the most abundance cell-type in CD45+
247 *ex vivo* samples) and CD4+ T cells, and in a subtype of stromal cells of endometriosis patients.
248 In contrast, the pathway was upregulated in uNK cells and two subtypes of CD8+ T cells of
249 endometriosis patients (Fig. 5C). A similar pattern emerged when looking at the expression of
250 specific inflammatory ligand receptors (Fig. 5D).

251 At the opposite end of the spectrum, we identified several genes significantly more highly
252 expressed in endometriosis patients versus healthy controls, many of which had also previously
253 shown to have increased expression in the eutopic endometrium of endometriosis patients

254 compared to healthy controls (e.g. *CALDI* (44) and *SEMA3A* (45), which has been proposed
255 to be involved in macrophage recruitment to endometriotic lesions (46)), or increased
256 expression in the eutopic endometrium of severe compared to mild endometriosis patients
257 (*TEAD2* (47)). Unlike the downregulated genes, these genes were typically not immune-cell
258 specific, but rather connected to tissue invasiveness, epithelial to mesenchymal transition or
259 myc targets (Fig. 5B). C-myc expression has been previously reported to be altered in the
260 eutopic endometrium of endometriosis patients when compared to controls which, together
261 with a reduction in cell death by apoptosis, is thought to play a role in facilitating the invasive
262 features of endometriotic endometrium (48).

263 Interestingly, among immune-related genes, endometriosis patients also had significantly
264 higher expression of *CD8A* (Fig. 5B). The abundance and activation levels of CD8 T cells has
265 been a topic of considerable interest in endometriosis research. However, results vary
266 depending on the tissue analysed (49). Our study supports the notion that CD8⁺ T cells (or
267 subsets thereof (13)) are enriched in the menstrual fluid of endometriosis patients.

268 *Differentially expressed genes in all cell and sample types*

269 In the model designed to identify biomarkers that are detectable irrespective of cell type
270 enriched or whether the sample is freshly isolated or cultured, we identified a large number of
271 differentially expressed genes (Fig. 5F, Sup. Table 10). Many of these genes, such as
272 *MTRNR2L1* and *HBG2*, were not, to our knowledge, previously described as biomarkers of
273 endometriosis.

274 *MTRNR2L1* was detected by all 3 models as being substantially upregulated in healthy controls
275 compared to endometriosis patients (Fig. 5G). As high *MTRNR2L1* expression in human
276 endometrium is highly specific to endometrial epithelial glandular cells (50), the lower
277 expression of this gene in endometriosis patients might be driven either by lower expression
278 levels in their epithelial cells, or by a proportion of epithelial cells present in their MF.
279 Interestingly, despite not having been previously linked to endometriosis, *MTRNR2L1* was
280 found to be downregulated in the eutopic endometrium of adenomyosis patients compared to
281 healthy controls (51).

282 *HBG2* is more highly expressed in endometriosis patients than in controls (Fig. 5H). As high
283 expression of *HBG2* is very highly specific to macrophages and erythroid cells and lowly
284 expressed in all cell-types of the endometrium (50), we hypothesize that the origin of this signal
285 may be in a higher proportion of macrophage cells in endometriosis patients, or in a higher
286 amount of ambient RNA from higher levels of cell death in the samples of endometriosis
287 patients. Intriguingly, *HBG2* was previously found to be over-expressed in the menstrual
288 endometrium of women with heavy menstrual bleeding when compared to controls (52).

289 **Discussion**

290 This study demonstrates the feasibility of using self-collected MF to obtain a comprehensive
291 cellular and molecular snapshot of the human endometrium. Our data show that MF collection
292 is not only viable but also yields a substantial number of live cells, including epithelial, stromal,
293 and immune cells. The similarity in transcriptional profiles between MF and endometrial
294 biopsies and recapitulation of menstrual cycle dynamics reinforces the reliability of MF as a
295 representative sample of the endometrial environment. This representativeness opens avenues
296 for exploring menstrual fluid in research settings and potential clinical diagnostics.

297 Our pilot study reports on useful parameters for future large-scale studies. Our data show that
298 MF is primarily composed by CD45+ myeloid cells, which dominate the transcriptomic signal.
299 Unsorted *ex vivo* samples vary widely in cell-type composition with no association with disease
300 state, highlighting the importance of enriching for the target cell-type in both bulk or single-
301 cell transcriptomics assays. We report differences between *ex vivo* and *in vitro* (cultured cells)
302 cells, with the former containing a large number of immune-related cells, and the latter being
303 composed mainly of stromal-like cells. This stresses the importance of probing fresh cells
304 directly or having alternative specialized culture conditions if the immune compartment is to
305 be assessed. We also report on genes whose transcription in stromal cells has been previously
306 reported to be dysregulated in endometriosis patients, but which we found to be affected by
307 cell culture (e.g. *MARCH4*, *INMT* or *GABRP* (32,53)) or menstrual cycle stage (18) (e.g.
308 *PENK*).

309 In patients with endometriosis, our MF analysis revealed distinctive gene expression patterns.
310 Notably, the differential expression of genes such as *FNI*, *PDE4B*, *PENK*, *MLLT11* and
311 *HAND2-AS1* in stromal-like cells of endometriosis patients aligns with previous findings and
312 underscores their potential roles in the disease's pathogenesis. Furthermore, the dysregulation
313 of immune-related genes in endometriosis patients suggests an altered immune landscape in
314 endometriosis, possibly contributing to the disease's inflammatory component. The
315 identification of *MTRNR2L1* and *HBG2* as potential biomarkers that are robust to cell culture
316 and cell-type enrichment is intriguing. These findings necessitate further research to validate
317 and understand the roles of these biomarkers in endometriosis.

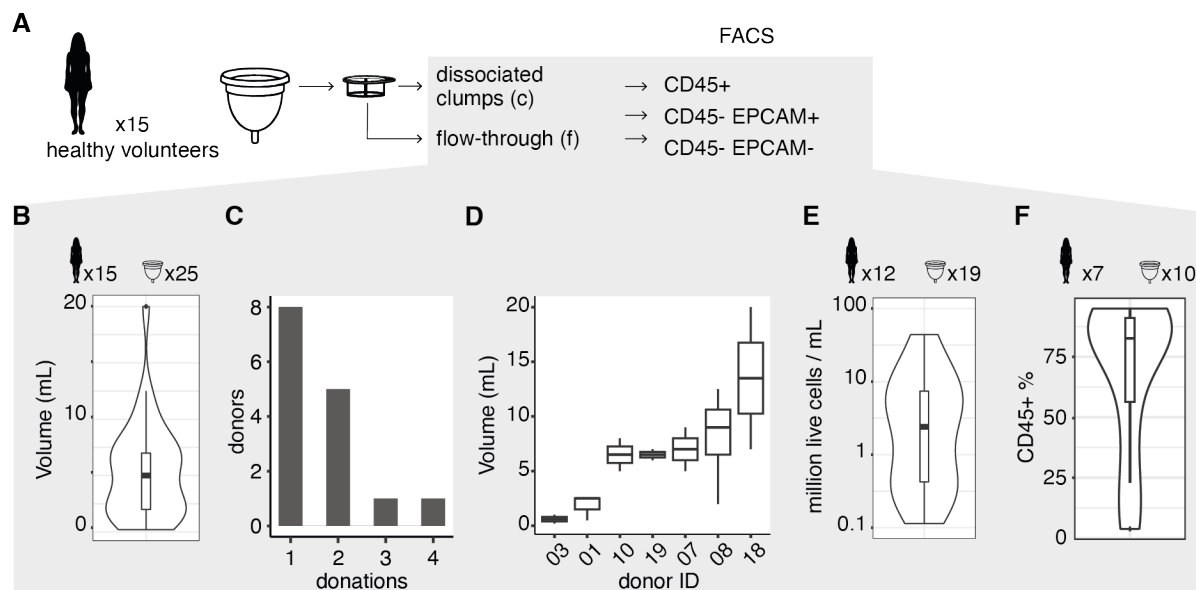
318 The use of MF for diagnosing endometriosis offers a promising non-invasive alternative to
319 current diagnostic methods, which are often invasive and delayed. Our pilot study was limited
320 by a relatively small sample size. Further research with larger cohorts with standardized
321 collection procedures is essential to validate our findings and establish MF-based diagnostics
322 in clinical practice.

323 In conclusion, our study highlights the potential of menstrual fluid as a rich, non-invasive
324 source for studying the endometrium and diagnosing endometriosis. This work paves the way
325 for further studies to explore the full diagnostic potential of menstrual fluid and its application
326 in women's health. Combined with banking stem cells isolated from MF, MF sampling can
327 enable powerful functional studies of gynecological disease.

328

329 **Figures**

330 **Figure 1**



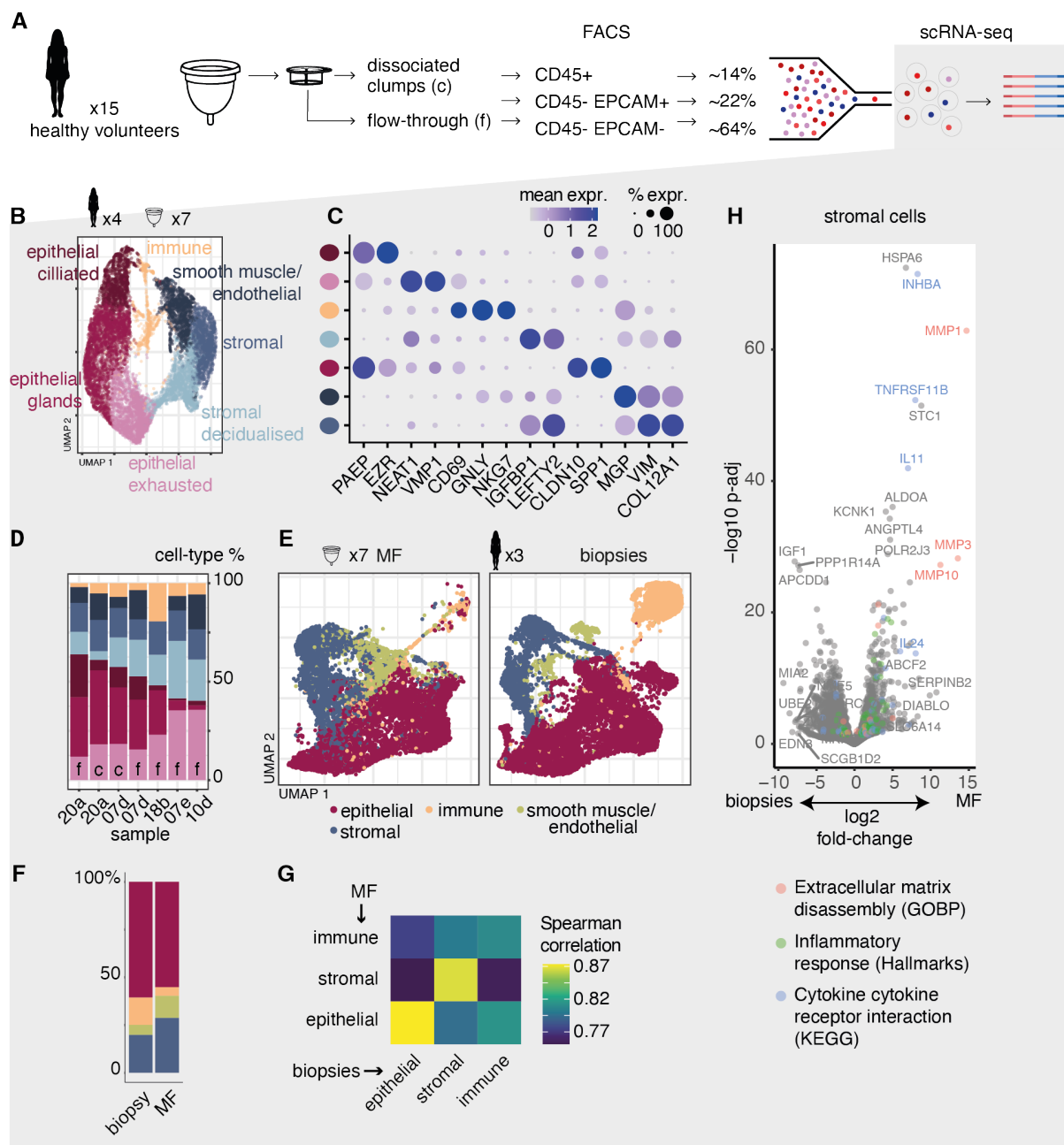
331

332 **Figure 1:** Characterization of menstrual fluid from healthy donors. (A) Sample preparation for
333 the FACS assays. (B) Total sample volumes in mL. (C) Frequency of the number of donations
334 per donor. (D) Volume in mL of MF across 2-4 donations from the same donor. (E) Million of
335 live cells per mL of flow-through. (F) Percentage of CD45+ cells in flow-through.

336

337

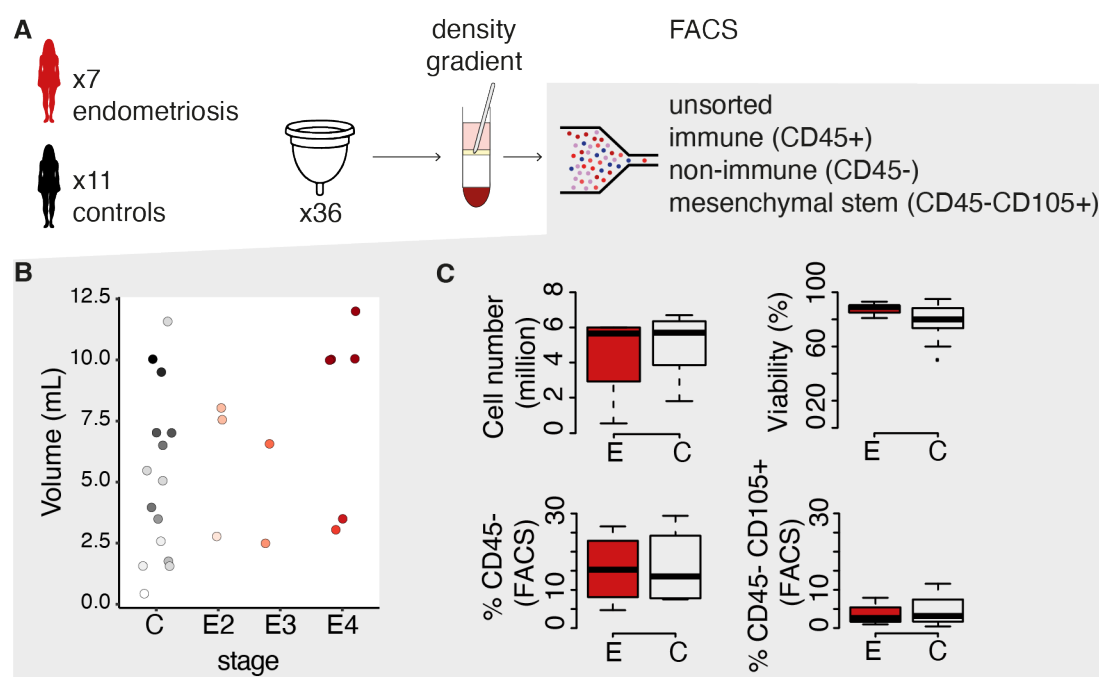
338 **Figure 2**



339

340 **Figure 2:** Characterisation of menstrual fluid from healthy donors using single-cell RNA-seq
 341 (scRNA-seq). (A) Sample preparation for the scRNA-seq assays. (B) Single-cell UMAP of
 342 ciliated, glandular and exhausted epithelial cells, mesenchymal and decidualized stroma,
 343 smooth muscle/endothelial and immune cells. (C) Marker genes used for the cell-type
 344 annotation. (D) Cell-type abundances in each MF sample (c - clumps, f - flow-through). (E)
 345 UMAP of the MF samples integrated with endometrial biopsies from Wang et al. (20). (F) Cell-
 346 type abundances in MF and biopsies. (G) Gene expression correlation between the MF and
 347 biopsies. (H) Differential gene expression between stromal cells in MF and biopsies. Gene
 348 names are shown for the top 15 genes with largest absolute log₂ fold changes and for genes
 349 with -log₁₀ adjusted p-values greater than 20. Colors indicate if genes are members of the
 350 pathways indicated. The pathways chosen were significantly over-represented (FDR q-value
 351 ≤ 0,00027) in gene set enrichment analysis.

352 **Figure 3**



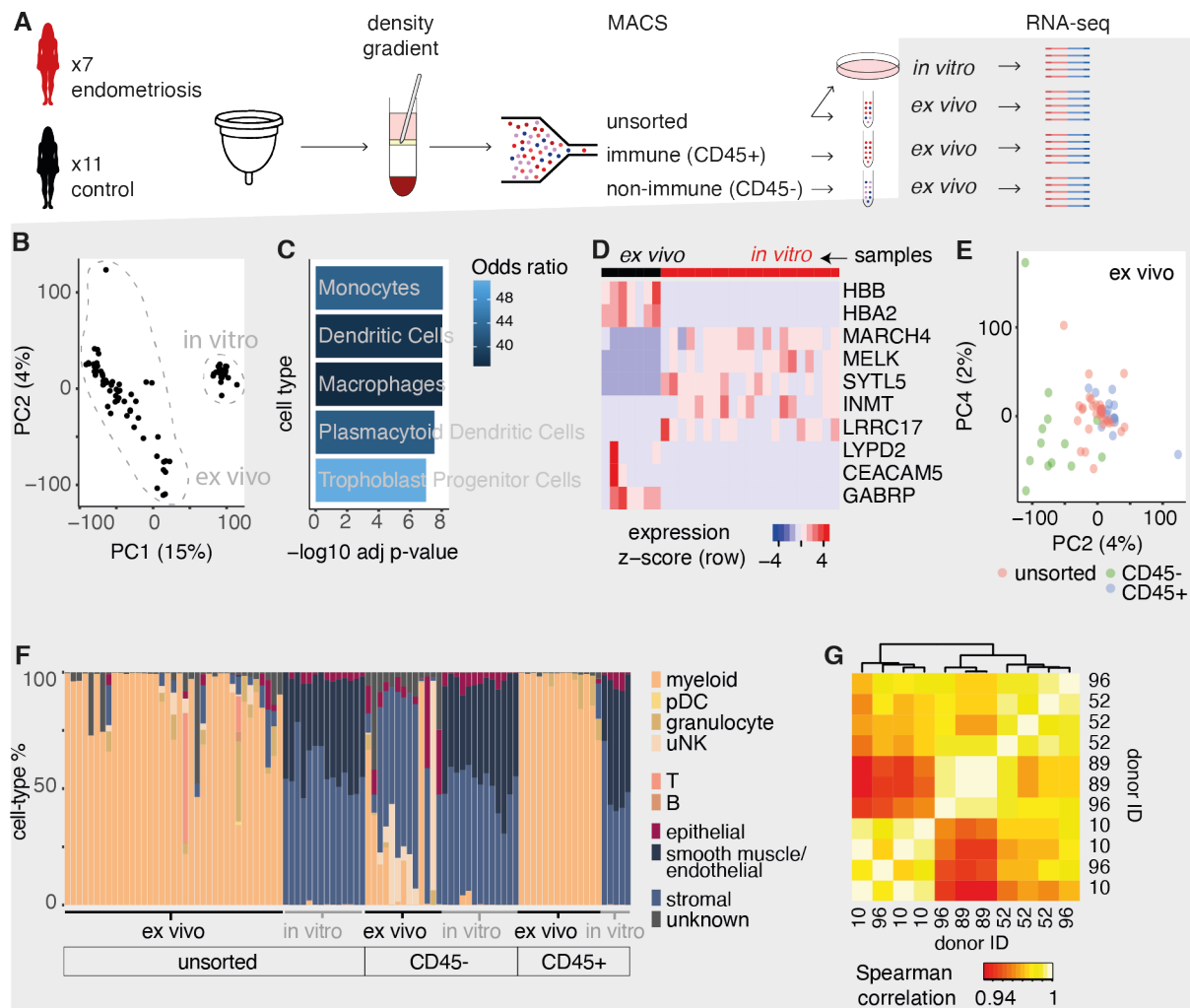
353

354 **Figure 3:** Comparison of menstrual fluid volume, cell number and cell-type proportions
355 between endometriosis patients and healthy controls. (A) Sample preparation procedure. (B)
356 MF volume in controls - C, and endometriosis stages 2 to 4 (E2-E4). Points are coloured by
357 donor. (C) Cell number, viability, fraction of CD45+ cells and fraction of CD45+CD105+ cells
358 in endometriosis - E, and controls - C.

359

360

361 **Figure 4**

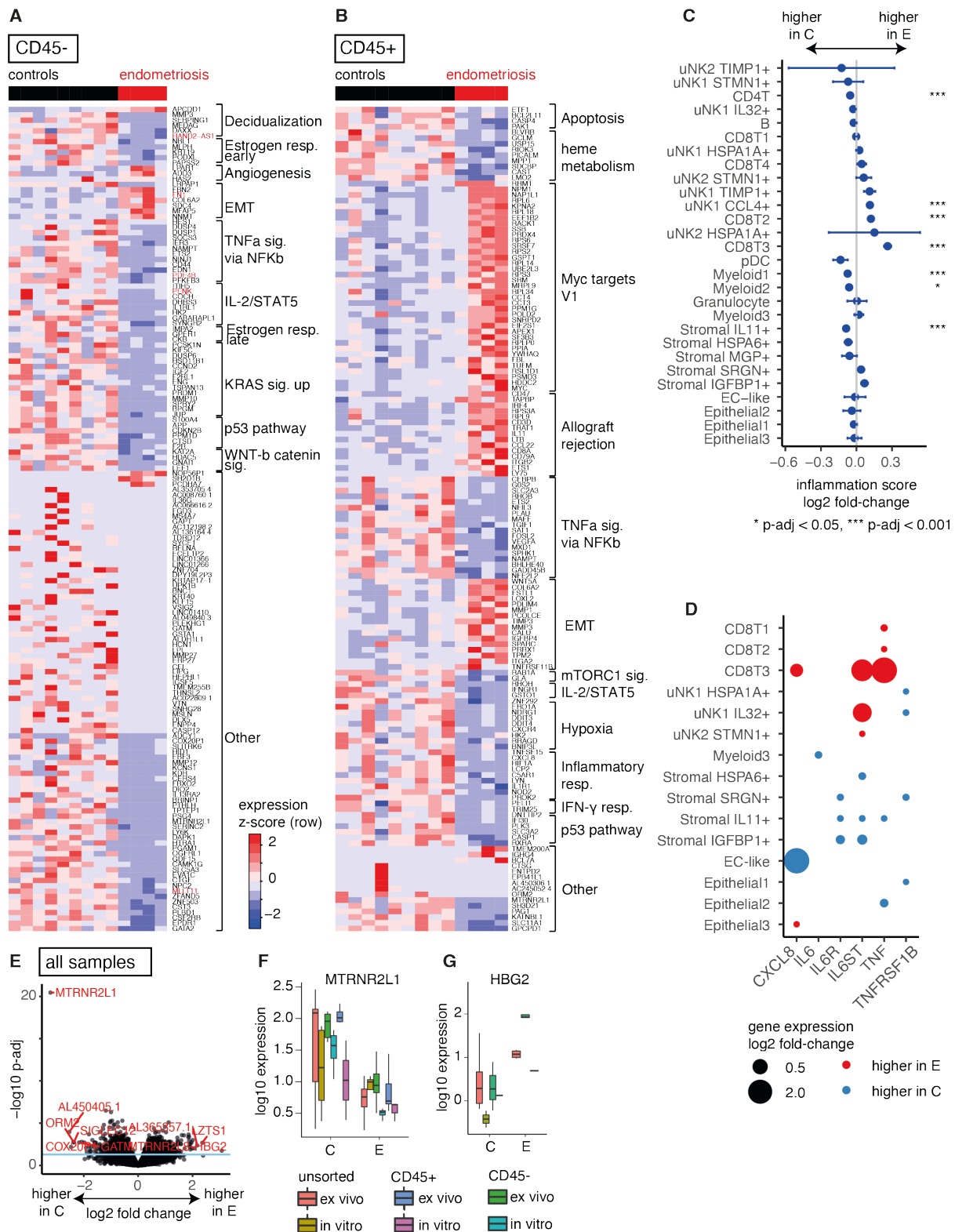


362

363 **Figure 4:** Cell-type composition of MF samples by cell-type enrichment method. (A) Sample
 364 preparation procedure. (B) Principal component analysis of all samples using the full
 365 transcriptome. Shown in parentheses on the axes is the percentage of variance explained by
 366 each of the principal components. (C) Over-representation analysis of the top 30 genes
 367 positively associated with PC1 in the cell-type marker gene-sets from PanglaoDB(54). P-values
 368 were calculated using Fisher's exact test and corrected for multiple testing using the Benjamini-
 369 Hochberg procedure. (D) Gene expression heatmap of the 10 most significantly differentially
 370 expressed genes between CD45- *ex vivo* and all *in vitro* samples. (E) Principal component
 371 analysis of *ex vivo* samples. (F) *In silico* cell-type deconvolution of the bulk RNA-seq samples
 372 using MuSiC. pDC - plasmacytoid dendritic cells, uNK – uterine NK cells. (G) Spearman
 373 correlation of the whole CD45+ *ex vivo* transcriptomes for donors that donated more than once.
 374 The row-side dendrogram (identical to the column-side one) was omitted.

375

376 **Figure 5**



377

378 **Figure 5:** Comparison of the transcriptomes between endometriosis patients and healthy
 379 controls. (A) Gene expression heatmap of 150 genes differentially expressed genes in CD45-
 380 *in vitro* samples (columns). The genes shown are either 1) members of pathways significantly
 381 enriched among differentially expressed genes, or 2) top most differentially expressed genes
 382 according to p-value and absolute log2 fold-change (“Other”). (B) Gene expression heatmap

383 of 150 genes differentially expressed genes in CD45+ *ex vivo* samples (columns). The genes
384 shown are either 1) members of pathways significantly enriched among differentially expressed
385 genes, or 2) top most differentially expressed genes according to p-value and absolute log₂ fold
386 change (“Other”). (C) Log₂ fold-change of the inflammatory signaling pathway activity
387 between endometriosis patients – E, and controls – C in the different cell-types defined in Shih.
388 et al. Horizontal bars show standard errors and asterisks indicate significance level of a t-test
389 performed for each cell-type independently. P-values were adjusted for multiple testing with
390 the method from Benjamini-Hochberg. (D) Log₂ fold-change of the expression of selected
391 ligand-receptors involved in inflammatory pathways between endometriosis – E and controls
392 – C. (E) Volcano plot of the differential expression of endometriosis - E - versus healthy
393 controls – C using all available samples. Points above the horizontal blue line indicate genes
394 with multiple-testing adjusted p-values below 0.05. (F) Normalised expression levels of
395 MTRNR2L1 in endometriosis - E and control - C samples. (G) Normalised expression levels
396 of HBG2 in endometriosis - E and control - C samples.

397 **Methods**

398 **Sample collection for the healthy donor cohort**

399 The study was approved by the South West - Frenchay Research Ethics Committee (IRAS
400 project ID 221351). Fifteen healthy participants (menstruating, self-defined healthy females
401 with no other selection criteria) were recruited from employees, visitors and students at the
402 Wellcome Trust Sanger Institute. All samples and data collected were solely identified by a
403 randomly assigned identifier and the research team had no access to the participant names. No
404 links were kept between A) the consent forms containing the participant name and B) the
405 samples and sample associated data collected. Participants were between 24 and 47 years old
406 with self-reported regular menstrual cycles. Two of the 15 participants were users of combined
407 oral contraceptive pills, and a third one of a contraceptive implant. All women were requested
408 to collect their complete menstrual fluid during day 2 of their period using a silicone menstrual
409 cup (Mooncup®). These cups were purchased by the investigation team for the purpose of this
410 investigation. We do not hold any commercial link with this provider. Size A was
411 recommended if the person was aged 30 and over and/or has given birth vaginally at any age.
412 Size B was recommended if the person was under 30 and had not given birth vaginally.
413 Participants were asked to pour and/or pipette the sample from the Mooncup into a sample
414 collection tube and to deliver the sample tube to the study collection boxes within 6 hours of
415 menstrual cup removal. Donors were able to donate samples up to 6 times each and at each
416 donation filled in a questionnaire with demographic and biologic data (age, time of cup
417 removal, current use of contraception, medication and ongoing illness). All sample collection
418 tubes and questionnaires within a kit were labeled with the same randomly assigned identifier,
419 thereby completely anonymising the samples but allowing multiple donations from the same
420 participant to be linked. Sample delivery was unscheduled and samples were processed upon
421 receipt from the collection boxes (within 30m to 12h from the time of menstrual cup removal,
422 of this time up to 3h at RT and the remainder in the fridge at 5°C).

423 **Sample processing for the healthy donor cohort**

424 The total volume, appearance (color, haemolysis and viscosity scores) and presence of
425 clumps/mucus were assessed for each sample. Five of the samples had high mucous content
426 and were centrifuged at 300g for 5 minutes. The supernatant was transferred to a new tube,
427 spun at 2000g for 3 minutes and resuspended with 50mL PBS. All samples were then serially
428 filtered with 100µm and 70µm nylon mesh sieves (Falcon). Clumps were backwashed from the
429 sieves with PBS and digested with collagenase for 20–60 minutes at 37°C. Flowthrough and
430 clumps were briefly centrifuged and subject to red blood cell lysis (Biolegend Inc.). Cells were
431 counted with Trypan blue. Samples with more than 1 million live cells were processed by flow
432 cytometry.

433 **Flow cytometry staining and acquisition for the healthy donor cohort**

434 After digestion, cells were washed and centrifuged at 2000rpm for 3 minutes. Cells were
435 resuspended in PBS plus Human TruStain FcX Blocking solution (BioLegend) and incubated
436 for 5 minutes at room temperature in the dark. 5µl/million cells of CD45 (clone 2D1,
437 BioLegend) and EPCAM (clone 3C4, BioLegend) antibodies were added and incubated for 1h
438 at 4°C in the dark. Post staining, cells were washed and centrifuged at 2000rpm for 3 minutes.
439 1µl/mL of DAPI was added and incubated for 5 minutes at RT. Cells were washed twice,
440 filtered with a 50µm FACS filter and analyzed using SH800S, MoFlo XDP or BD Influx cell
441 sorters, depending on instrument availability, according to manufacturer's instructions and
442 gating strategies. After sorting, the CD45+, CD45- EPCAM- and EPCAM+ fractions were
443 collected for 10x single-cell RNA-sequencing.

444 **Generation of single cell transcriptomes**

445 Single-cell suspensions were loaded per channel of Chromium™ Single Cell Chips (10X
446 Genomics® Chromium Single Cell 3' Reagent Kits v2.0), aiming for a recovery of 4,600 cells.
447 Reverse transcription and library construction were carried out according to the manufacturer's
448 recommendations. Libraries were sequenced on an Illumina HiSeq 4000 using paired-end runs
449 of 150bp.

450 **Computational analysis of the single-cell RNA-seq data**

451 Raw sequencing reads were processed using the Cell Ranger analysis pipeline (55). The
452 "cellranger count" command was used to generate filtered and raw matrices. Reads were
453 aligned against the human genome version GRCh38. Raw gene barcode count matrices were
454 processed using CellBender for unsupervised denoising (56) and further analyzed using the R
455 package Seurat (57). To remove low quality cells, an adaptive filtering threshold approach was
456 used based on extreme numbers of counts (count depth) and extreme numbers of genes per
457 barcode. Cells were filtered when counts or genes per barcode were less than 99% of all cells,
458 or when the mitochondrial content was higher than 10%.

459 To annotate cell types, we used two strategies: annotation of individually processed samples,
460 and annotation of all samples integrated together. For the individually processed samples,
461 counts were normalized using the SCT normalization approach of Seurat. Sample-specific

462 UMAPs were constructed using a subset of genes exhibiting high cell-to-cell variation which
463 were identified by modeling the mean-variance relationship. The top 3000 features were used
464 to perform PCA analysis. To cluster the cells, a K-nearest neighbor (kNN) graph based on the
465 euclidean distance in PCA space was first constructed using the first 30 PC components as
466 input. Next, the Louvain algorithm was applied to iteratively group cells. We identified the cell
467 types in each cluster using a combination of manual and automated approaches using known
468 marker genes (Fig. 1E). First, clusters were assigned to known cell populations using cell type-
469 specific markers obtained through the FindAllMarkers function. Multiple testing correction
470 was performed using Benjamini-Hochberg procedure. Second, the R package Garnett (58) in
471 cluster extension mode was used to annotate cells in a semi-automated manner. Integration of
472 samples was performed using reference-based canonical correlation analysis together with SCT
473 normalization (Seurat), and marker based annotation was performed as above. A consensus of
474 the three annotation strategies was used to annotate each sample.

475 Integration of the MF samples with the endometrial biopsies was performed using reference-
476 based canonical correlation analysis together with SCT normalization (Seurat) using 3000
477 integration features.

478 Differential expression analysis of menstrual fluid versus the endometrial biopsies was
479 performed by generating pseudo-bulks for each of the main cell-types: stromal, epithelial and
480 immune. Pseudo-bulking was performed by summing up the counts of all cells of each cell-
481 type by sample. Only genes that had non-zero counts in at least one MF and one biopsy sample
482 were kept. Samples were then used as replicates in a DESeq2 analysis using default parameters.
483 Log fold change shrinkage using the lfcShrink function with apeglm was performed on the
484 results.

485 Gene set enrichment analyses were performed using the GSEA tool (v4.3.3, (59)). Sorted
486 differentially expressed genes by log fold change were used as input to a weighted analysis
487 using the following gene-sets: Hallmarks, GO Biological Process, Reactome and KEGG.

488 **Sample collection and processing for the endometriosis cohort**

489 We recruited age-matched (20-45) healthy female volunteers and endometriosis patients
490 through advertising at endometriosis support groups as well as the EPFL campus. The study
491 received ethical approval by the Swiss Cantonal Authorities Vaud, CER-VD, project nr. 2016-
492 00770. All samples and data collected were solely identified by a randomly assigned identifier.
493 Samples were processed within 6h by a standard density gradient-based cell isolation protocol
494 (Histopaque). The viability and the number of nucleated cells in the cell suspension was
495 determined using a Nucleostainer and when possible, further processed by MACS into a
496 CD45+ and CD45- fraction. For a subset of samples, surface marker gene expression was
497 determined by FACS. The isolated single cell suspension was diluted to 1×10^7 cells/ml with
498 FACS buffer (DPBS -/-) with 1% human platelet serum and the following fluorophore-
499 conjugated antibodies were added: anti-human CD45 for identifying hematopoietic cells, and
500 anti-human CD105 for identifying mesenchymal stromal cells. 7-Aminoactinomycin D (7-
501 AAD) was used for assessing viability and Syto40 was used for discerning nucleated cells. Cell
502 culture was performed using high glucose MEMalpha medium supplemented with 5% platelet
503 serum and 50 ug/ml Primocin. TrypLE Select reagent was used to collect the cells from the cell
504 culture plates.

505 **Generation of bulk RNA-seq for the endometriosis cohort**

506 RNA was collected from fresh or cultured samples into Tri-Reagent and bulk RNA-seq was
507 performed as previously described, using BRB-seq, a highly multiplexed 3' end protocol (60).

508 **Computational analysis of the RNA-seq for the endometriosis cohort**

509 Sequencing results were processed in the Deplancke laboratory by a standard in-house pipeline,
510 consistent with published methodology to obtain digital gene expression values for each cell
511 and estimated expression for each sample. In brief, sequenced tags were demultiplexed (at
512 sample and library level) and fastq files containing 62-bp-long single-end sequenced tags
513 (reads) were trimmed and filtered using prinseq 0.20.3 and cutadapt 1.5 and subsequently
514 aligned to the Ensembl 84 gene annotation of the hg19 human genome using STAR 2.4.0g.
515 The number of tags per gene was calculated using htseq-count 0.6.0 with the parameters 'htseq-
516 count -m intersection-nonempty -s no -a 10 -t exon -i gene_id'.

517 The over-representation analysis in Fig. 4C was performed by taking the top 30 genes most
518 anticorrelated with PC1 from Fig. 4B and testing their enrichment in the cell-type marker gene
519 sets from PanglaoDB (54), using the hypergeometric test implemented in Enrichr (61).

520 The decomposition of the bulk-RNA-seq data into cell-type compositions was performed using
521 raw counts and cell-type annotations from the dataset in Shih et al.(13) as reference, and the
522 sample deconvolution method implemented in the MuSiC R package (24), with default
523 parameters.

524 The differential gene expression analyses of endometriosis versus healthy controls were
525 performed using the standard DESeq2 workflow (62). For the CD45- *in vitro* and CD45+ *ex*
526 *vivo* cell analyses we used endometriosis (yes or no) as a fixed effect. For the "all cell and
527 sample types" analyses we used type (*ex vivo* or *in vitro*), cell-type (unsorted, CD45+ or
528 CD45-) and endometriosis (yes or no) as fixed effect terms.

529 Over representation analysis for Fig. 5A-B was performed by taking all genes significantly up
530 or down regulated in each analysis and testing for their enrichment in the MSigDB Hallmarks
531 gene sets (63) using the hypergeometric test implemented in Enrichr (61) and correcting for
532 multiple testing. Additionally, in Fig. 5A we included differentially expressed decidualization-
533 associated genes. We defined decidualization-associated genes as genes whose expression
534 peaks in stromal cells during decidualization in Wang et al. (20) or that are known from the
535 literature.

536 For the scoring of the "Inflammatory signaling" pathway in the single-cells from Shih et
537 al.(13), we downloaded the gene names for this pathway from MSigDB (63) and used AUCell
538 (43) with maxrank set to 800. To test for significance of differential scoring between cell-types
539 we transformed the scores with $\log_2(\text{score}+0.0000001)$ and then used a linear model with
540 endometriosis status (yes or no) as predictor.

541 **Supplementary Materials**

542 Sup. Table 1: Donor and sample characteristics for the healthy single-cell cohort.

543 Sup. Table 2: Differentially expressed genes between MF and endometrial biopsies.

544 Sup. Table 3: Donor and sample characteristics for the endometriosis and controls bulk-
545 RNA-seq cohort.
546 Sup. Table 4: Over-representation analysis of in vitro associated genes in the PanglaoDB cell-
547 type marker gene sets.
548 Sup. Table 5: Differentially expressed genes between in vitro and ex vivo CD45- samples.
549 Sup. Table 6: Differentially expressed genes between CD45- in vitro samples of
550 endometriosis patients versus controls.
551 Sup. Table 7: Over-representation analysis of the significant ($p_{adj} < 0.05$) genes from Sup.
552 Table 6.
553 Sup. Table 8: Differentially expressed genes between CD45+ ex vivo samples of
554 endometriosis patients versus controls.
555 Sup. Table 9: Over-representation analysis of the significant ($p_{adj} < 0.05$) genes from Sup.
556 Table 8.
557 Sup. Table 10: Differentially expressed genes between all samples of endometriosis patients
558 versus controls.

559 **Author contributions**

560 AG and PCS: designed the study.
561 GC: provided critical input on study design and volunteer recruitment.
562 CB, JR, MB, MZ and DA: collected, processed samples and performed experiments.
563 AG, PCS and EA: analysed data.
564 AG: wrote the manuscript.
565 BD: provided resources and input on data interpretation.
566 PCS: edited the manuscript.
567 All authors: read, commented and approved the manuscript.

568 **Acknowledgements**

569 This work was supported by an Early Detection Gynaecological Cancers Pump Priming award
570 by CRUK to AG. PCS was supported by an HFSP Postdoctoral Fellowship and a Diversa
571 Foundation Grant. We thank Suzanne Ratte for help with experimental technical assistance,
572 Johanne Doleman for support with the ethical application and Daniel Gaffney for support with
573 resources. We thank all volunteers for their participation in this study.

574 **Data access**

575 Count matrices for all sequencing data DOI: [10.5281/zenodo.11105267](https://doi.org/10.5281/zenodo.11105267)
576 ArrayExpress: accession numbers for sc and bulk RNA-seq requested, assignment pending.

577 **Declaration/Conflict of interest**

578 • The cups (LadyCup) for the endometriosis cohort were donated by Ladyplanet GmbH.
579 • DA, PCS and MB conducted this work while employed at EPFL in Switzerland. DA is
580 currently employed by and holds Alithea Genomics company stock. MB is currentl
581 employed by Lonza AG. PCS is currently employed by and holds F. Hoffmann-La
582 Roche Ltd. company stock.

584 **References**

- 585 1. Donoso MB, Serra R, Rice GE, Gana MT, Rojas C, Khoury M, et al. Normality Ranges
586 of Menstrual Fluid Volume During Reproductive Life Using Direct Quantification of
587 Menses with Vaginal Cups. *Gynecol Obstet Invest.* 2019;84(4):390–5.
- 588 2. Ivarsson MA, Stiglund N, Marquardt N, Westgren M, Gidlöf S, Björkström NK.
589 Composition and dynamics of the uterine NK cell KIR repertoire in menstrual blood.
590 *Mucosal Immunology.* 2017 Mar;10(2):322–31.
- 591 3. Alcayaga-Miranda F, Cuenca J, Luz-Crawford P, Aguila-Díaz C, Fernandez A, Figueroa
592 FE, et al. Characterization of menstrual stem cells: angiogenic effect, migration and
593 hematopoietic stem cell support in comparison with bone marrow mesenchymal stem
594 cells. *Stem Cell Res Ther.* 2015 Dec;6(1):32.
- 595 4. Rossignoli F, Caselli A, Grisendi G, Piccinno S, Burns JS, Murgia A, et al. Isolation,
596 Characterization, and Transduction of Endometrial Decidual Tissue Multipotent
597 Mesenchymal Stromal/Stem Cells from Menstrual Blood. *BioMed Research*
598 *International.* 2013;2013:1–14.
- 599 5. Faramarzi H, Mehrabani D, Fard M, Akhavan M, Zare S, Bakhshalizadeh S, et al. The
600 Potential of Menstrual Blood-Derived Stem Cells in Differentiation to Epidermal
601 Lineage: A Preliminary Report. 2016;
- 602 6. Liu Y, Niu R, Yang F, Yan Y, Liang S, Sun Y, et al. Biological characteristics of human
603 menstrual blood-derived endometrial stem cells. *J Cell Mol Med.* 2018 Mar;22(3):1627–
604 39.
- 605 7. Patel AN, Park E, Kuzman M, Benetti F, Silva FJ, Allickson JG. Multipotent Menstrual
606 Blood Stromal Stem Cells: Isolation, Characterization, and Differentiation. *Cell*
607 *Transplant.* 2008 Mar;17(3):303–11.
- 608 8. Sabbaj S, Hel Z, Richter HE, Mestecky J, Goepfert PA. Menstrual Blood as a Potential
609 Source of Endometrial Derived CD3+ T Cells. Landay A, editor. *PLoS ONE.* 2011 Dec
610 9;6(12):e28894.
- 611 9. Cindrova-Davies T, Zhao X, Elder K, Jones CJP, Moffett A, Burton GJ, et al. Menstrual
612 flow as a non-invasive source of endometrial organoids. *Commun Biol.* 2021 Jun
613 17;4(1):651.
- 614 10. Filby CE, Wyatt KA, Mortlock S, Cousins FL, McKinnon B, Tyson KE, et al.
615 Comparison of Organoids from Menstrual Fluid and Hormone-Treated Endometrium:
616 Novel Tools for Gynecological Research. *JPM.* 2021 Dec 6;11(12):1314.
- 617 11. Warren LA, Shih A, Renteira SM, Seekin T, Blau B, Simpfendorfer K, et al. Analysis of
618 menstrual effluent: diagnostic potential for endometriosis. *Mol Med.* 2018 Dec;24(1):1.
- 619 12. Schmitz T, Hoffmann V, Olliges E, Bobinger A, Popovici R, Nöbner E, et al. Reduced
620 frequency of perforin-positive CD8+ T cells in menstrual effluent of endometriosis
621 patients. *Journal of Reproductive Immunology.* 2021 Nov;148:103424.

- 622 13. Shih AJ, Adelson RP, Vashistha H, Khalili H, Nayyar A, Puran R, et al. Single-cell
623 analysis of menstrual endometrial tissues defines phenotypes associated with
624 endometriosis. *BMC Med.* 2022 Sep 15;20(1):315.
- 625 14. Nayyar A, Saleem MI, Yilmaz M, DeFranco M, Klein G, Elmaliki KM, et al. Menstrual
626 Effluent Provides a Novel Diagnostic Window on the Pathogenesis of Endometriosis.
627 *Front Reprod Health.* 2020 Jul 22;2:3.
- 628 15. Van Der Molen RG, Schutten JHF, Van Cranenbroek B, Ter Meer M, Donckers J,
629 Scholten RR, et al. Menstrual blood closely resembles the uterine immune micro-
630 environment and is clearly distinct from peripheral blood. *Human Reproduction.* 2014
631 Feb 1;29(2):303–14.
- 632 16. Fraser IS, McCarron G, Markham R, Resta T. Blood and total fluid content of menstrual
633 discharge. *Obstetrics and gynecology.* 1985;65(2):194–8.
- 634 17. Hosseini S, Shokri F, Tokhmechy R, Savadi-Shiraz E, Jeddi-Tehrani M, Rahbari M, et al.
635 Menstrual blood contains immune cells with inflammatory and anti-inflammatory
636 properties. *J of Obstet and Gynaecol.* 2015 Nov;41(11):1803–12.
- 637 18. Brulport A, Bourdon M, Vaiman D, Drouet C, Pocate-Cheriet K, Bouzid K, et al. An
638 integrated multi-tissue approach for endometriosis candidate biomarkers: a systematic
639 review. *Reprod Biol Endocrinol.* 2024 Feb 10;22(1):21.
- 640 19. Koks CAM, Dunselman GAJ, De Goeij AFPM, Arends JW, Evers JLH. Evaluation of a
641 menstrual cup to collect shed endometrium for in vitro studies. *Fertility and Sterility.*
642 1997 Sep;68(3):560–4.
- 643 20. Wang W, Vilella F, Alama P, Moreno I, Mignardi M, Isakova A, et al. Single-cell
644 transcriptomic atlas of the human endometrium during the menstrual cycle. *Nat Med.*
645 2020 Oct;26(10):1644–53.
- 646 21. Gaide Chevronnay HP, Selvais C, Emonard H, Galant C, Marbaix E, Henriët P.
647 Regulation of matrix metalloproteinases activity studied in human endometrium as a
648 paradigm of cyclic tissue breakdown and regeneration. *Biochimica et Biophysica Acta*
649 *(BBA) - Proteins and Proteomics.* 2012 Jan;1824(1):146–56.
- 650 22. Mutlu L, Hufnagel D, Taylor HS. The Endometrium as a Source of Mesenchymal Stem
651 Cells for Regenerative Medicine1. *Biology of Reproduction* [Internet]. 2015 Jun 1 [cited
652 2024 Mar 6];92(6). Available from: [https://academic.oup.com/biolreprod/article-](https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.114.126771)
653 [lookup/doi/10.1095/biolreprod.114.126771](https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.114.126771)
- 654 23. Bozorgmehr M, Gurung S, Darzi S, Nikoo S, Kazemnejad S, Zarnani AH, et al.
655 Endometrial and Menstrual Blood Mesenchymal Stem/Stromal Cells: Biological
656 Properties and Clinical Application. *Front Cell Dev Biol.* 2020 Jul 9;8:497.
- 657 24. Wang X, Park J, Susztak K, Zhang NR, Li M. Bulk tissue cell type deconvolution with
658 multi-subject single-cell expression reference. *Nat Commun.* 2019 Jan 22;10(1):380.
- 659 25. Chen ST, Shi WW, Lin YQ, Yang ZS, Wang Y, Li MY, et al. Embryo-derive TNF
660 promotes decidualization via fibroblast activation. *eLife.* 2023 Jul 17;12:e82970.

- 661 26. Liu L, Dong H, Guan Y, Fan T, Sun W, Bagchi IC, et al. Regulation of HAND2
662 Expression by LncRNA *HAND2-AS1* in Ovarian Endometriosis Involving DNA
663 Methylation. *Journal of the Endocrine Society*. 2023 May 5;7(6):bvad049.
- 664 27. Huyen DV, Bany BM. Evidence for a conserved function of heart and neural crest
665 derivatives expressed transcript 2 in mouse and human decidualization.
666 REPRODUCTION. 2011 Aug;142(2):353–68.
- 667 28. Klemmt PAB, Carver JG, Kennedy SH, Koninckx PR, Mardon HJ. Stromal cells from
668 endometriotic lesions and endometrium from women with endometriosis have reduced
669 decidualization capacity. *Fertility and Sterility*. 2006 Mar;85(3):564–72.
- 670 29. Barragan F, Irwin JC, Balayan S, Erikson DW, Chen JC, Houshdaran S, et al. Human
671 Endometrial Fibroblasts Derived from Mesenchymal Progenitors Inherit Progesterone
672 Resistance and Acquire an Inflammatory Phenotype in the Endometrial Niche in
673 Endometriosis1. *Biology of Reproduction* [Internet]. 2016 May 1 [cited 2024 Apr
674 3];94(5). Available from: [https://academic.oup.com/biolreprod/article-](https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.115.136010)
675 [lookup/doi/10.1095/biolreprod.115.136010](https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.115.136010)
- 676 30. Bulun S, Monsavais D, Pavone M, Dyson M, Xue Q, Attar E, et al. Role of Estrogen
677 Receptor- β in Endometriosis. *Semin Reprod Med*. 2012 Jan;30(01):39–45.
- 678 31. Burney RO, Talbi S, Hamilton AE, Vo KC, Nyegaard M, Nezhat CR, et al. Gene
679 Expression Analysis of Endometrium Reveals Progesterone Resistance and Candidate
680 Susceptibility Genes in Women with Endometriosis. *Endocrinology*. 2007 Aug
681 1;148(8):3814–26.
- 682 32. Chen S, Chai X, Wu X. Bioinformatical analysis of the key differentially expressed genes
683 and associations with immune cell infiltration in development of endometriosis. *BMC*
684 *Genom Data*. 2022 Mar 18;23(1):20.
- 685 33. Lavogina D, Samuel K, Lavrits A, Meltsov A, Söritsa D, Kadastik Ü, et al.
686 Chemosensitivity and chemoresistance in endometriosis – differences for ectopic versus
687 eutopic cells. *Reproductive BioMedicine Online*. 2019 Oct;39(4):556–68.
- 688 34. Velho RV, Taube E, Sehouli J, Mechsner S. Neurogenic Inflammation in the Context of
689 Endometriosis—What Do We Know? *IJMS*. 2021 Dec 3;22(23):13102.
- 690 35. Proestling K, Husslein H, Hudson QJ, Witzmann-Stern M, Widmar B, Bagó-Horváth Z,
691 et al. MLLT11 Regulates Endometrial Stroma Cell Adhesion, Proliferation and Survival
692 in Ectopic Lesions of Women with Advanced Endometriosis. *IJMS*. 2023 Dec
693 28;25(1):439.
- 694 36. Yang F, Wu Y, Hockey R, Doust J, Mishra GD, Montgomery GW, et al. Evidence of
695 shared genetic factors in the etiology of gastrointestinal disorders and endometriosis and
696 clinical implications for disease management. *Cell Reports Medicine*. 2023
697 Nov;4(11):101250.
- 698 37. Sapkota Y, Steinhorsdottir V, Morris AP, Fassbender A, Rahmioglu N, De Vivo I, et al.
699 Meta-analysis identifies five novel loci associated with endometriosis highlighting key
700 genes involved in hormone metabolism. *Nat Commun*. 2017 May 24;8(1):15539.

- 701 38. Matalliotaki C, Matalliotakis M, Rahmioglu N, Mavromatidis G, Matalliotakis I,
702 Koumantakis G, et al. Role of FN1 and GREB1 gene polymorphisms in endometriosis.
703 Mol Med Report [Internet]. 2019 May 15 [cited 2024 Apr 3]; Available from:
704 <http://www.spandidos-publications.com/10.3892/mmr.2019.10247>
- 705 39. Trachsel B, Imobersteg S, Valpreda G, Singer G, Grabherr R, Ormos M, et al. Relaxed
706 fibronectin: a potential novel target for imaging endometriotic lesions. EJMIMI Res.
707 2024 Feb 10;14(1):17.
- 708 40. Lis-Kuberka J, Kubik P, Chrobak A, Pająk J, Chelmońska-Soyta A, Orczyk-Pawilowicz
709 M. Fibronectin Molecular Status in Plasma of Women with Endometriosis and Fertility
710 Disorders. IJMS. 2021 Oct 22;22(21):11410.
- 711 41. Warzecha D, Załęcka J, Mańka G, Kiecka M, Lipa M, Spaczyński R, et al. Plasma and
712 Peritoneal Fluid Fibronectin and Collagen IV Levels as Potential Biomarkers of
713 Endometriosis. IJMS. 2022 Dec 10;23(24):15669.
- 714 42. Tseng JF, Ryan IP, Milam TD, Murai JT, Schriock ED, Landers DV, et al. Interleukin-6
715 secretion in vitro is up-regulated in ectopic and eutopic endometrial stromal cells from
716 women with endometriosis. The Journal of Clinical Endocrinology & Metabolism. 1996
717 Mar;81(3):1118–22.
- 718 43. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G,
719 et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods.
720 2017 Nov;14(11):1083–6.
- 721 44. Meola J, Hidalgo GDS, Silva JCRE, Silva LECM, Paz CCP, Ferriani RA. Caldesmon:
722 New Insights for Diagnosing Endometriosis1. Biology of Reproduction [Internet]. 2013
723 May 1 [cited 2024 Mar 5];88(5). Available from:
724 <https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.112.103598>
- 725 45. Liang Y, Wang W, Huang J, Tan H, Liu T, Shang C, et al. Potential Role of Semaphorin
726 3A and Its Receptors in Regulating Aberrant Sympathetic Innervation in Peritoneal and
727 Deep Infiltrating Endometriosis. Zhou R, editor. PLoS ONE. 2015 Dec
728 31;10(12):e0146027.
- 729 46. Wu J, Xie H, Yao S, Liang Y. Macrophage and nerve interaction in endometriosis. J
730 Neuroinflammation. 2017 Dec;14(1):53.
- 731 47. Aghajanova L, Giudice LC. Molecular Evidence for Differences in Endometrium in
732 Severe Versus Mild Endometriosis. Reprod Sci. 2011 Mar;18(3):229–51.
- 733 48. Johnson MC, Torres M, Alves A, Bacallao K, Fuentes A, Vega M, et al. Augmented cell
734 survival in eutopic endometrium from women with endometriosis: Expression of c-myc,
735 TGF-beta1 and bax genes. Reprod Biol Endocrinol. 2005 Dec;3(1):45.
- 736 49. Kisovar A, Becker CM, Granne I, Southcombe JH. The role of CD8+ T cells in
737 endometriosis: a systematic review. Front Immunol. 2023 Jul 11;14:1225639.
- 738 50. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al.
739 Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419.

- 740 51. Juárez-Barber E, Corachán A, Carbajo-García MC, Faus A, Vidal C, Giles J, et al.
741 Transcriptome analysis of adenomyosis eutopic endometrium reveals molecular
742 mechanisms involved in adenomyosis-related implantation failure and pregnancy
743 disorders. *Reprod Biol Endocrinol*. 2024 Jan 9;22(1):10.
- 744 52. Girling JE, Lockhart MG, Olshansky M, Paiva P, Woodrow N, Marino JL, et al.
745 Differential Gene Expression in Menstrual Endometrium From Women With Self-
746 Reported Heavy Menstrual Bleeding. *Reprod Sci*. 2017 Jan;24(1):28–46.
- 747 53. Kobayashi H, Yamashita Y, Iwase A, Yoshikawa Y, Yasui H, Kawai Y, et al. The
748 ferroimmunomodulatory role of ectopic endometriotic stromal cells in ovarian
749 endometriosis. *Fertility and Sterility*. 2012 Aug;98(2):415-422.e12.
- 750 54. Franzén O, Gan LM, Björkegren JLM. PanglaoDB: a web server for exploration of
751 mouse and human single-cell RNA sequencing data. *Database [Internet]*. 2019 Jan 1
752 [cited 2024 Mar 6];2019. Available from:
753 <https://academic.oup.com/database/article/doi/10.1093/database/baz046/5427041>
- 754 55. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively
755 parallel digital transcriptional profiling of single cells. *Nat Commun*. 2017 Jan
756 16;8(1):14049.
- 757 56. Fleming SJ, Chaffin MD, Arduini A, Akkad AD, Banks E, Marioni JC, et al.
758 Unsupervised removal of systematic background noise from droplet-based single-cell
759 experiments using CellBender. *Nat Methods [Internet]*. 2023 Aug 7 [cited 2023 Aug 16];
760 Available from: <https://www.nature.com/articles/s41592-023-01943-7>
- 761 57. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated
762 analysis of multimodal single-cell data. *Cell*. 2021 Jun;184(13):3573-3587.e29.
- 763 58. Pliner HA, Shendure J, Trapnell C. Supervised classification enables rapid annotation of
764 cell atlases. *Nat Methods*. 2019 Oct;16(10):983–6.
- 765 59. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al.
766 Gene set enrichment analysis: A knowledge-based approach for interpreting genome-
767 wide expression profiles. *Proc Natl Acad Sci USA*. 2005 Oct 25;102(43):15545–50.
- 768 60. Alpern D, Gardeux V, Russeil J, Mangeat B, Meireles-Filho ACA, Breysse R, et al.
769 BRB-seq: ultra-affordable high-throughput transcriptomics enabled by bulk RNA
770 barcoding and sequencing. *Genome Biol*. 2019 Dec;20(1):71.
- 771 61. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive
772 and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013
773 Dec;14(1):128.
- 774 62. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
775 RNA-seq data with DESeq2. *Genome Biol*. 2014 Dec;15(12):550.
- 776 63. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The
777 Molecular Signatures Database Hallmark Gene Set Collection. *Cell Systems*. 2015
778 Dec;1(6):417–25.

