


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

Petra C Schwalie, Cemsal Bafligil, Julie Russeil, Magda Zachara, Marjan Biocanin, Daniel Alpern, Evelin Aasna, Bart Deplancke, Geraldine Canny, Angela Goncalves 

Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland • Roche Innovation Center Basel, Pharma Research and Early Development, Basel, Switzerland • Wellcome Trust Sanger Institute, Wellcome Genome Campus, Cambridge, UK • German Cancer Research Center (DKFZ), Somatic Evolution and Early Detection Heidelberg, Germany • University College Dublin, Research Office, Dublin, Ireland

 [https://en.wikipedia.org/wiki/Open\\_access](https://en.wikipedia.org/wiki/Open_access)

 Copyright information

## eLife Assessment

This basic research study presents **useful** data concerning the menstrual fluid composition and its potential for endometriosis biomarker research. However, despite **solid** bioinformatics analyses, the choice of markers used to separate or identify the different cell types needs to be justified and the results better discussed in relation to current knowledge of the pathophysiology of endometriosis.

<https://doi.org/10.7554/eLife.99558.1.sa2>

## Abstract

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

patients. Finally, we identify potential novel pan-cell-type biomarkers for this neglected condition.

## Introduction

Menstrual fluid (MF) is an accessible source of large numbers of live cells of high biomedical interest. During menstruation, the female body sheds  $86\pm 48$  ml of MF (1 [↗](#)) containing cervico-vaginal secretions, blood and cells from the lining of uterus, cervix, vagina, and possibly the fallopian tube and ovaries. This rich source of live cells has been previously proposed as a scalable source of live uterine NK cells (2 [↗](#)), mesenchymal stem cells (3 [↗](#)–7 [↗](#)), T-cells (8 [↗](#)) and epithelial cells (9 [↗](#), 10 [↗](#)) for research and therapy.

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

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

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

## Results

## Menstrual fluid collection is a viable sampling alternative to endometrial biopsy

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

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

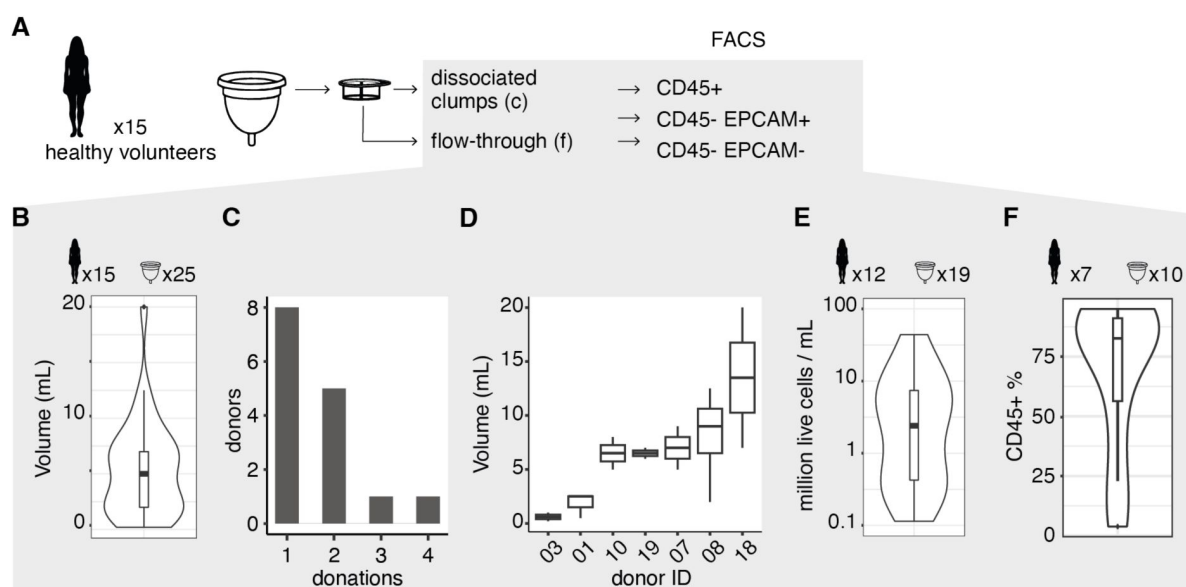
The median volume of MF per sample was 5mL (Fig. 1B), with a potential donor effect (Fig. 1C-D, one-way ANOVA p-value = 0.09), but no age association (generalized linear mixed model with donor random effect p-value = 0.43). Three out of 25 MF samples did not yield any live cells, these samples came from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

### Cell-type composition and comparison to endometrial biopsy by single-cell RNA-seq

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

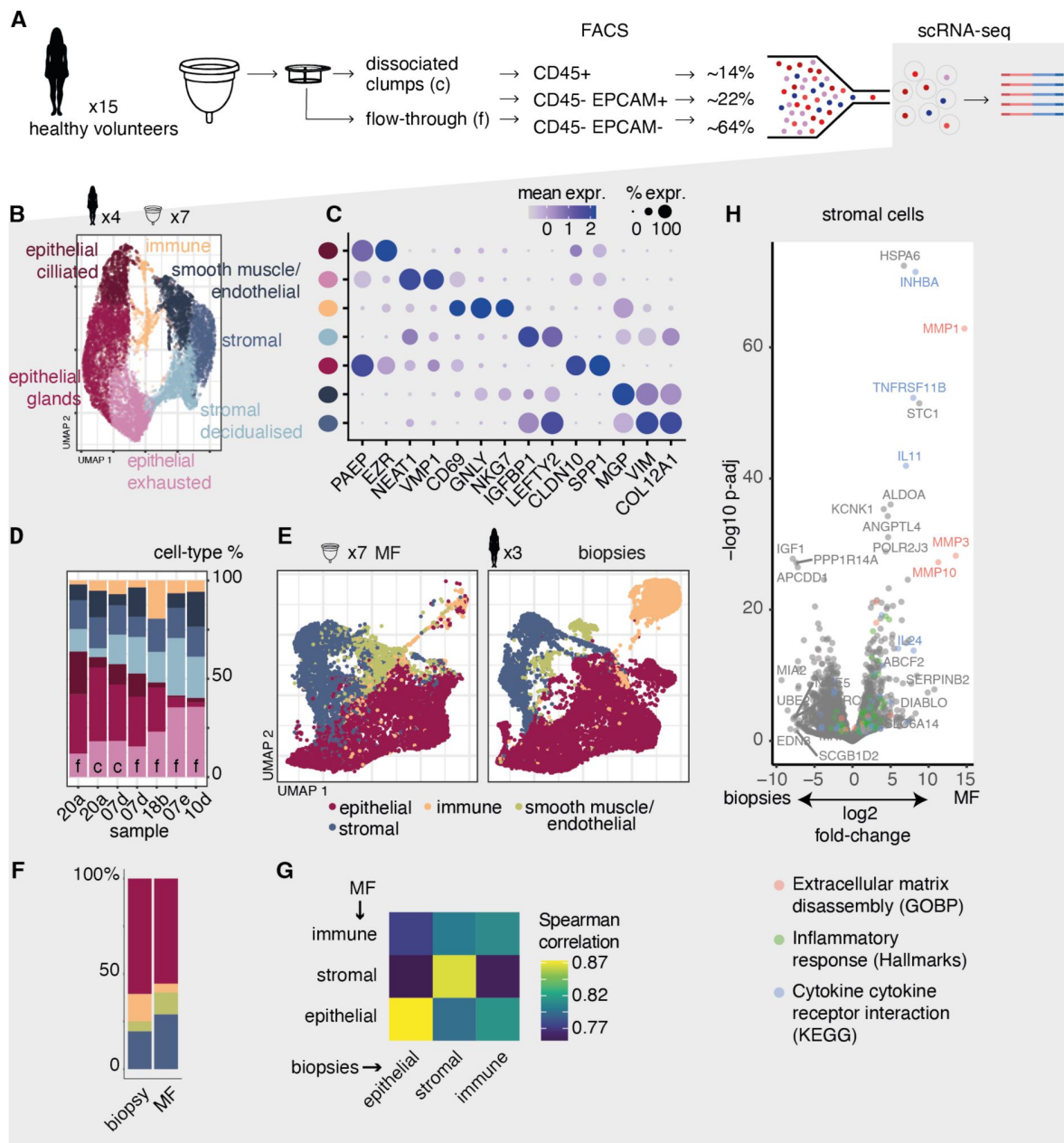
We then asked whether the transcriptional profiles of the MF samples correspond to the transcriptional profiles of freshly isolated endometrial cells by computationally integrating our single-cell data with single-cells from 3 endometrial biopsies (20) in the late-secretory and proliferative phases of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

Despite the overall transcriptomic concordance, we expected expression differences between MF and the endometrial biopsies to reflect differences in cycle stage. The endometrial biopsy samples used came from cycle stages with high progesterone levels, which represses, among others, expression of many matrix metalloproteinases (MMPs). Progesterone withdrawal at menstruation triggers expression of MMPs (21), an effect mediated by cytokines including tumor necrosis factor-α (TNF-α). In addition, a large number of neutrophils loaded with MMPs are recruited through induction of chemokines such as interleukin 8 (CXCL8) (21). Differential gene



**Figure 1**

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



**Figure 2**

Characterisation of menstrual fluid from healthy donors using single-cell RNA-seq (scRNA-seq). (A) Sample preparation for the scRNA-seq assays. (B) Single-cell UMAP of ciliated, glandular and exhausted epithelial cells, mesenchymal and decidualized stroma, smooth muscle/endothelial and immune cells. (C) Marker genes used for the cell-type annotation. (D) Cell-type abundances in each MF sample (c - clumps, f - flow-through). (E) UMAP of the MF samples integrated with endometrial biopsies from Wang et al. (2024). (F) Cell-type abundances in MF and biopsies. (G) Gene expression correlation between the MF and biopsies. (H) Differential gene expression between stromal cells in MF and biopsies. Gene names are shown for the top 15 genes with largest absolute log2 fold changes and for genes with  $-\log_{10}$  adjusted p-values greater than 20. Colors indicate if genes are members of the pathways indicated. The pathways chosen were significantly over-represented (FDR q-value  $\leq 0.00027$ ) in gene set enrichment analysis.

expression comparison between MF and the endometrial biopsy stromal cells revealed much higher expression in MF of MMPs, members of the TNF- $\alpha$  signaling pathway and *CXCL8* (Fig. 2H [↗](#), Sup. Table 2).

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

## Properties of MF from patients with endometriosis and healthy controls

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

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

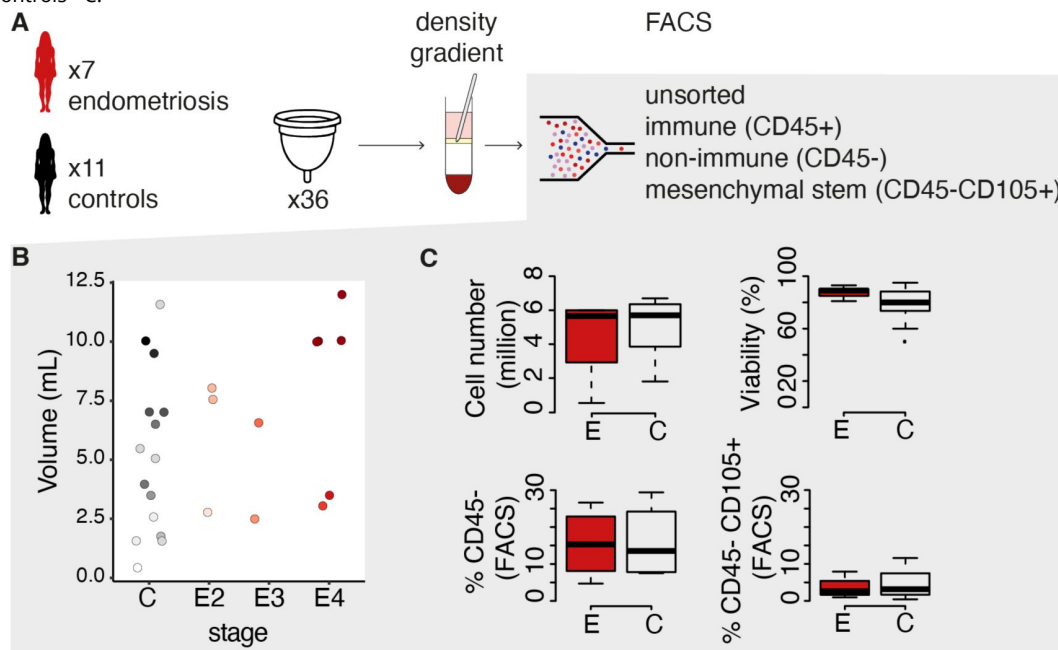
## Cell-type composition of ex vivo and in vitro MF cells by RNA-seq

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

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

**Figure 3**

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

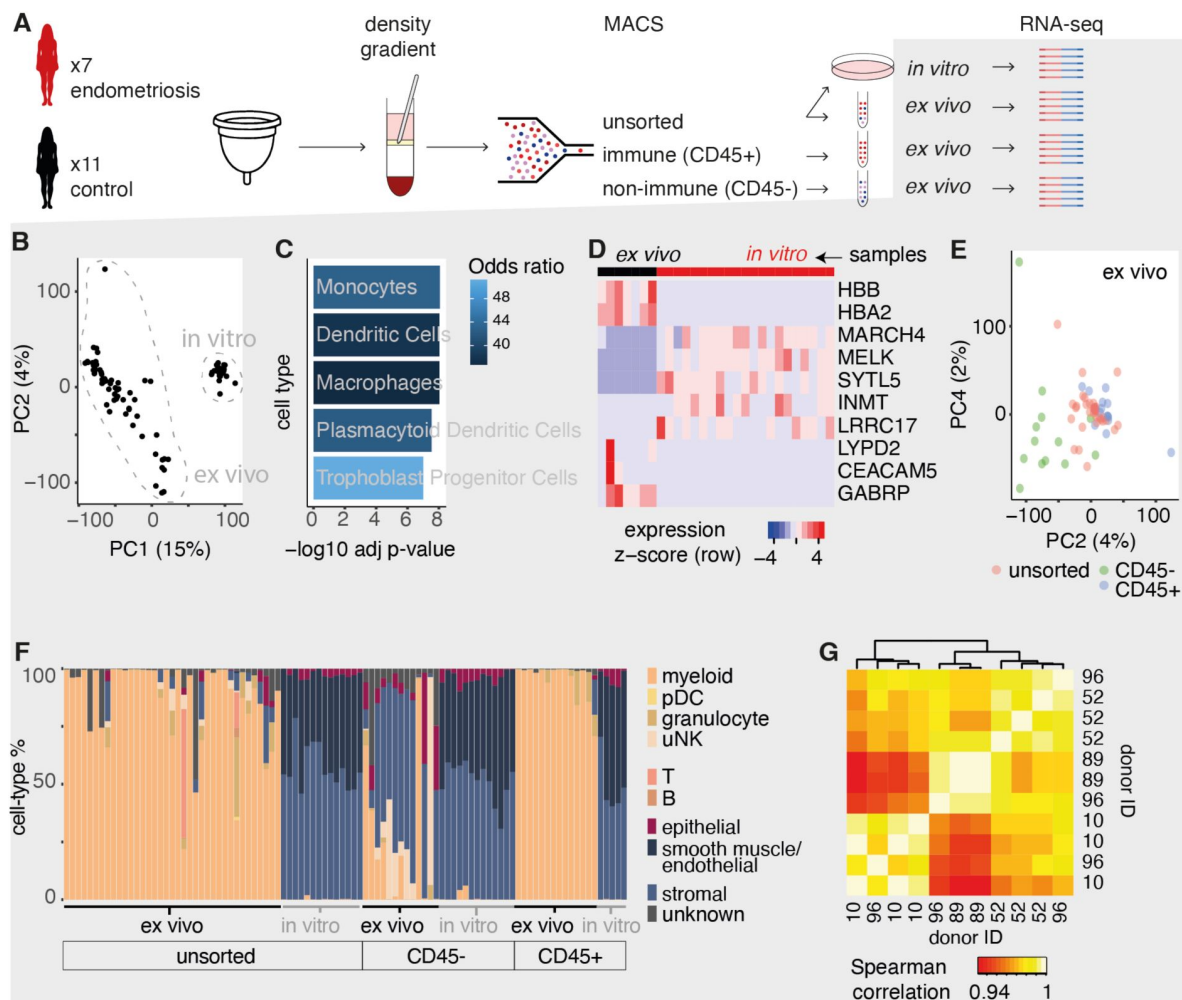


**Table 1**

**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





**Figure 4**

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



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

## Donor effect on transcriptomic profiles

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

## Identification of potential transcriptional biomarkers of endometriosis in MF

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

## Differentially expressed genes in CD45-stromal-like cells

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

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

The endogenous opioid peptide precursor proenkephalin (*PENK*) has been previously reported to be up-regulated in whole-endometrium ([31](#), [32](#)) of endometriosis patients when compared to healthy controls. In contrast, *PENK* expression was found to be down-regulated in ectopic stroma when compared to eutopic stroma of patients ([33](#)). The discrepancy in the direction of regulation may have to do with the specific cell-types assayed, as well as the cycle phase at which endometrial samples were obtained, as *PENK* expression is modulated by the menstrual cycle ([31](#)). Interestingly, a decrease in opioid peptides has been previously suggested to be involved in the maintenance of chronic inflammation in endometriosis ([34](#)).

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



HAND2 Antisense RNA 1 (*HAND2-AS1*), whose expression is coordinated with *HAND2* in human endometrial stromal cells, was also previously found to be reduced in ectopic lesions from patients compared controls, primarily expressed by stromal cells as opposed to epithelial, and that its silencing impairs stromal cell decidualization (20 [4](#), 26 [4](#), 27 [4](#)).

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

Relatively fewer genes were found to be up-regulated in endometriosis patients when compared to controls. Notably, the most significant up-regulated gene was fibronectin 1 (*FN1*). A genome-wide meta-analysis and a subsequent targeted study have implicated single nucleotide polymorphisms in *FN1* with moderate to severe endometriosis (37 [4](#), 38 [4](#)), and relaxed fibronectin is being investigated as a potential target for imaging endometriotic lesions (39 [4](#)). At the protein level, studies comparing eutopic and ectopic endometrium from patients have yielded contradictory results (39 [4](#)). However, fibronectin levels in plasma and peritoneal fluid have consistently reported elevated levels in endometriosis patients when compared to healthy women (40 [4](#), 41 [4](#)).

### Differentially expressed genes in CD45+ immune cells

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

This was intriguing, as the production of potentially pro-inflammatory cytokines, such as IL-6, has been previously found to be up-regulated in cultured and fresh stromal cells from ectopic and eutopic endometrium from patients compared to controls (13 [4](#), 42 [4](#)). To explore the activation of inflammatory pathways in specific cell-types from endometriosis patients and controls, we used AUCell (43 [4](#)) to score the activity of the MsigDB “inflammatory response” pathway in each individual cell in the dataset from Shih et al. (13 [4](#)) (Methods).

In agreement with our findings, we confirmed that the activity of the inflammatory signaling pathway was significantly downregulated in myeloid (the most abundance cell-type in CD45+ *ex vivo* samples) and CD4+ T cells, and in a subtype of stromal cells of endometriosis patients. In contrast, the pathway was upregulated in uNK cells and two subtypes of CD8+ T cells of endometriosis patients (Fig. 5C [4](#)). A similar pattern emerged when looking at the expression of specific inflammatory ligand receptors (Fig. 5D [4](#)).

At the opposite end of the spectrum, we identified several genes significantly more highly expressed in endometriosis patients versus healthy controls, many of which had also previously shown to have increased expression in the eutopic endometrium of endometriosis patients compared to healthy controls (e.g. *CALD1* (44 [4](#)) and *SEMA3A* (45 [4](#)), which has been proposed to be involved in macrophage recruitment to endometriotic lesions (46 [4](#)), or increased expression in the eutopic endometrium of severe compared to mild endometriosis patients (*TEAD2* (47 [4](#))). Unlike the downregulated genes, these genes were typically not immune-cell specific, but rather connected to tissue invasiveness, epithelial to mesenchymal transition or myc targets (Fig. 5B [4](#)). C-myc expression has been previously reported to be altered in the eutopic endometrium of endometriosis patients when compared to controls which, together with a reduction in cell death by apoptosis, is thought to play a role in facilitating the invasive features of endometriotic endometrium (48 [4](#)).

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

## Differentially expressed genes in all cell and sample types

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

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

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

## Discussion

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

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

In patients with endometriosis, our MF analysis revealed distinctive gene expression patterns. Notably, the differential expression of genes such as *FN1*, *PDE4B*, *PENK*, *MLLT11* and *HAND2-AS1* in stromal-like cells of endometriosis patients aligns with previous findings and underscores their potential roles in the disease's pathogenesis. Furthermore, the dysregulation of immune-related genes in endometriosis patients suggests an altered immune landscape in endometriosis, possibly contributing to the disease's inflammatory component. The identification of *MTRNR2L1* and *HBG2* as potential biomarkers that are robust to cell culture and cell-type enrichment is intriguing. These findings necessitate further research to validate and understand the roles of these biomarkers in endometriosis.

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

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

## Methods

### Sample collection for the healthy donor cohort

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



## Sample processing for the healthy donor cohort

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

## Flow cytometry staining and acquisition for the healthy donor cohort

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

## Generation of single cell transcriptomes

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

## Computational analysis of the single-cell RNA-seq data

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

To annotate cell types, we used two strategies: annotation of individually processed samples, and annotation of all samples integrated together. For the individually processed samples, counts were normalized using the SCT normalization approach of Seurat. Sample-specific UMAPs were constructed using a subset of genes exhibiting high cell-to-cell variation which were identified by modeling the mean-variance relationship. The top 3000 features were used to perform PCA analysis. To cluster the cells, a K-nearest neighbor (kNN) graph based on the euclidean distance in PCA space was first constructed using the first 30 PC components as input. Next, the Louvain algorithm was applied to iteratively group cells. We identified the cell types in each cluster using a combination of manual and automated approaches using known marker genes (**Fig. 1E**). First, clusters were assigned to known cell populations using cell type-specific markers obtained through the FindAllMarkers function. Multiple testing correction was performed using Benjamini-



Hochberg procedure. Second, the R package Garnett ([58](#)) in cluster extension mode was used to annotate cells in a semi-automated manner. Integration of samples was performed using reference-based canonical correlation analysis together with SCT normalization (Seurat), and marker based annotation was performed as above. A consensus of the three annotation strategies was used to annotate each sample.

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

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

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

## Sample collection and processing for the endometriosis cohort

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

## Generation of bulk RNA-seq for the endometriosis cohort

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

## Computational analysis of the RNA-seq for the endometriosis cohort

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

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

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

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

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

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

## Data Availability

All data produced in the present work are contained in the manuscript and available online at 10.5281/zenodo.11105267

<https://zenodo.org/records/11105267>

## Acknowledgements

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

## Additional information

### Author contributions

AG and PCS: designed the study.

GC: provided critical input on study design and volunteer recruitment.

CB, JR, MB, MZ and DA: collected, processed samples and performed experiments. AG, PCS and EA: analysed data.

AG: wrote the manuscript.

BD: provided resources and input on data interpretation. PCS: edited the manuscript.

All authors: read, commented and approved the manuscript.

## **Declaration/Conflict of interest**

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

## References

1. Donoso MB, Serra R, Rice GE, Gana MT, Rojas C, Khoury M, et al. (2019) **Normality Ranges of Menstrual Fluid Volume During Reproductive Life Using Direct Quantification of Menses with Vaginal Cups** *Gynecol Obstet Invest* **84**:390–5
2. Ivarsson MA, Stiglund N, Marquardt N, Westgren M, Gidlöf S, Björkström NK (2017) **Composition and dynamics of the uterine NK cell KIR repertoire in menstrual blood** *Mucosal Immunology* **10**:322–31
3. Alcayaga-Miranda F, Cuenca J, Luz-Crawford P, Aguila-Díaz C, Fernandez A, Figueroa FE, et al. (2015) **Characterization of menstrual stem cells: angiogenic effect, migration and hematopoietic stem cell support in comparison with bone marrow mesenchymal stem cells** *Stem Cell Res Ther* **6**
4. Rossignoli F, Caselli A, Grisendi G, Piccinno S, Burns JS, Murgia A, et al. (2013) **Isolation, Characterization, and Transduction of Endometrial Decidual Tissue Multipotent Mesenchymal Stromal/Stem Cells from Menstrual Blood** *BioMed Research International* **2013**:1–14
5. Faramarzi H, Mehrabani D, Fard M, Akhavan M, Zare S, Bakhshalizadeh S, et al. (2016) **The Potential of Menstrual Blood-Derived Stem Cells in Differentiation to Epidermal Lineage: A Preliminary Report** *World J Plast Surg.* **5**:26–31
6. Liu Y, Niu R, Yang F, Yan Y, Liang S, Sun Y, et al. (2018) **Biological characteristics of human menstrual blood-derived endometrial stem cells** *J Cell Mol Med* **22**:1627–39
7. Patel AN, Park E, Kuzman M, Benetti F, Silva FJ, Allickson JG (2008) **Multipotent Menstrual Blood Stromal Stem Cells: Isolation, Characterization, and Differentiation** *Cell Transplant* **17**:303–11
8. Sabbaj S, Hel Z, Richter HE, Mestecky J, Goepfert PA (2011) **Menstrual Blood as a Potential Source of Endometrial Derived CD3+ T Cells** *PLoS ONE* **6**
9. Cindrova-Davies T, Zhao X, Elder K, Jones CJP, Moffett A, Burton GJ, et al. (2021) **Menstrual flow as a non-invasive source of endometrial organoids** *Commun Biol* **4**
10. Filby CE, Wyatt KA, Mortlock S, Cousins FL, McKinnon B, Tyson KE, et al. (2021) **Comparison of Organoids from Menstrual Fluid and Hormone-Treated Endometrium: Novel Tools for Gynecological Research** *JPM* **11**
11. Warren LA, Shih A, Renteira SM, Seckin T, Blau B, Simpfendorfer K, et al. (2018) **Analysis of menstrual effluent: diagnostic potential for endometriosis** *Mol Med* **24**
12. Schmitz T, Hoffmann V, Olliges E, Bobinger A, Popovici R, Nößner E, et al. (2021) **Reduced frequency of perforin-positive CD8+ T cells in menstrual effluent of endometriosis patients** *Journal of Reproductive Immunology* **148**

13. Shih AJ, Adelson RP, Vashistha H, Khalili H, Nayyar A, Puran R, et al. (2022) **Single-cell analysis of menstrual endometrial tissues defines phenotypes associated with endometriosis** *BMC Med* **20**
14. Nayyar A, Saleem MI, Yilmaz M, DeFranco M, Klein G, Elmaliki KM, et al. (2020) **Menstrual Effluent Provides a Novel Diagnostic Window on the Pathogenesis of Endometriosis** *Front Reprod Health* **2**
15. Van Der Molen RG, Schutten JHF, Van Cranenbroek B, Ter Meer M, Donckers J, Scholten RR, et al. (2014) **Menstrual blood closely resembles the uterine immune micro-environment and is clearly distinct from peripheral blood** *Human Reproduction* **29**:303–14
16. Fraser IS, McCarron G, Markham R, Resta T (1985) **Blood and total fluid content of menstrual discharge** *Obstetrics and gynecology* **65**:194–8
17. Hosseini S, Shokri F, Tokhmechy R, Savadi-Shiraz E, Jeddi-Tehrani M, Rahbari M, et al. (2015) **Menstrual blood contains immune cells with inflammatory and anti-inflammatory properties** *J of Obstet and Gynaecol* **41**:1803–12
18. Brulport A, Bourdon M, Vaiman D, Drouet C, Pocate-Cheriet K, Bouzid K, et al. (2024) **An integrated multi-tissue approach for endometriosis candidate biomarkers: a systematic review** *Reprod Biol Endocrinol* **22**
19. Koks CAM, Dunselman GAJ, De Goeij AFPM, Arends JW, Evers JLH (1997) **Evaluation of a menstrual cup to collect shed endometrium for in vitro studies** *Fertility and Sterility* **68**:560–4
20. Wang W, Vilella F, Alama P, Moreno I, Mignardi M, Isakova A, et al. (2020) **Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle** *Nat Med* **26**:1644–53
21. Chevronnay HP Gaide, Selvais C, Emonard H, Galant C, Marbaix E, Henriët P (2012) **Regulation of matrix metalloproteinases activity studied in human endometrium as a paradigm of cyclic tissue breakdown and regeneration** *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1824**:146–56
22. Mutlu L, Hufnagel D, Taylor HS (2015) **The Endometrium as a Source of Mesenchymal Stem Cells for Regenerative Medicine** *1 Biology of Reproduction [Internet]* **92**
23. Bozorgmehr M, Gurung S, Darzi S, Nikoo S, Kazemnejad S, Zarnani AH, et al. (2020) **Endometrial and Menstrual Blood Mesenchymal Stem/Stromal Cells: Biological Properties and Clinical Application** *Front Cell Dev Biol* **8**
24. Wang X, Park J, Susztak K, Zhang NR, Li M (2019) **Bulk tissue cell type deconvolution with multi-subject single-cell expression reference** *Nat Commun* **10**
25. Chen ST, Shi WW, Lin YQ, Yang ZS, Wang Y, Li MY, et al. (2023) **Embryo-derive TNF promotes decidualization via fibroblast activation** *eLife* **12**
26. Liu L, Dong H, Guan Y, Fan T, Sun W, Bagchi IC, et al. (2023) **Regulation of HAND2 Expression by LncRNA HAND2-AS1 in Ovarian Endometriosis Involving DNA Methylation** *Journal of the Endocrine Society* **7**

27. Huyen DV, Bany BM (2011) **Evidence for a conserved function of heart and neural crest derivatives expressed transcript 2 in mouse and human decidualization** *REPRODUCTION* **142**:353–68
28. Klemmt PAB, Carver JG, Kennedy SH, Koninckx PR, Mardon HJ (2006) **Stromal cells from endometriotic lesions and endometrium from women with endometriosis have reduced decidualization capacity** *Fertility and Sterility* **85**:564–72
29. Barragan F, Irwin JC, Balayan S, Erikson DW, Chen JC, Houshdaran S, et al. (2016) **Human Endometrial Fibroblasts Derived from Mesenchymal Progenitors Inherit Progesterone Resistance and Acquire an Inflammatory Phenotype in the Endometrial Niche in Endometriosis** *Biology of Reproduction* **94**
30. Bulun S, Monsavais D, Pavone M, Dyson M, Xue Q, Attar E, et al. (2012) **Role of Estrogen Receptor- $\beta$  in Endometriosis** *Semin Reprod Med* **30**:39–45
31. Burney RO, Talbi S, Hamilton AE, Vo KC, Nyegaard M, Nezhat CR, et al. (2007) **Gene Expression Analysis of Endometrium Reveals Progesterone Resistance and Candidate Susceptibility Genes in Women with Endometriosis** *Endocrinology* **148**:3814–26
32. Chen S, Chai X, Wu X (2022) **Bioinformatical analysis of the key differentially expressed genes and associations with immune cell infiltration in development of endometriosis** *BMC Genom Data* **23**
33. Lavogina D, Samuel K, Lavrits A, Meltsov A, Söritsa D, Kadastik Ü, et al. (2019) **Chemosensitivity and chemoresistance in endometriosis – differences for ectopic versus eutopic cells** *Reproductive BioMedicine Online* **39**:556–68
34. Velho RV, Taube E, Sehouli J, Mechsner S (2021) **Neurogenic Inflammation in the Context of Endometriosis—What Do We Know?** *IJMS* **22**
35. Proestling K, Husslein H, Hudson QJ, Witzmann-Stern M, Widmar B, Bagó-Horváth Z, et al. (2023) **MLLT11 Regulates Endometrial Stroma Cell Adhesion, Proliferation and Survival in Ectopic Lesions of Women with Advanced Endometriosis** *IJMS* **25**
36. Yang F, Wu Y, Hockey R, Doust J, Mishra GD, Montgomery GW, et al. (2023) **Evidence of shared genetic factors in the etiology of gastrointestinal disorders and endometriosis and clinical implications for disease management** *Cell Reports Medicine* **4**
37. Sapkota Y, Steinhorsdottir V, Morris AP, Fassbender A, Rahmioglu N, De Vivo I, et al. (2017) **Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism** *Nat Commun* **8**
38. Matalliotaki C, Matalliotakis M, Rahmioglu N, Mavromatidis G, Matalliotakis I, Koumantakis G, et al. (2019) **Role of FN1 and GREB1 gene polymorphisms in endometriosis** *Mol Med Report*
39. Trachsel B, Imobersteg S, Valpreda G, Singer G, Grabherr R, Ormos M, et al. (2024) **Relaxed fibronectin: a potential novel target for imaging endometriotic lesions** *EJNMMI Res* **14**
40. Lis-Kuberka J, Kubik P, Chrobak A, Pająk J, Chelmońska-Soyta A, Orczyk-Pawłowicz M (2021) **Fibronectin Molecular Status in Plasma of Women with Endometriosis and Fertility Disorders** *IJMS* **22**



41. Warzecha D, Załęcka J, Mańka G, Kiecka M, Lipa M, Spaczyński R, et al. (2022) **Plasma and Peritoneal Fluid Fibronectin and Collagen IV Levels as Potential Biomarkers of Endometriosis** *IJMS* **23**
42. Tseng JF, Ryan IP, Milam TD, Murai JT, Schriock ED, Landers DV, et al. (1996) **Interleukin-6 secretion in vitro is up-regulated in ectopic and eutopic endometrial stromal cells from women with endometriosis** *The Journal of Clinical Endocrinology & Metabolism* **81**:1118–22
43. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. (2017) **SCENIC: single-cell regulatory network inference and clustering** *Nat Methods* **14**:1083–6
44. Meola J, Hidalgo GDS, Silva JCRE, Silva LECM, Paz CCP, Ferriani RA (2013) **Caldesmon: New Insights for Diagnosing Endometriosis** *Biology of Reproduction* **88**
45. Liang Y, Wang W, Huang J, Tan H, Liu T, Shang C, et al. (2015) **Potential Role of Semaphorin 3A and Its Receptors in Regulating Aberrant Sympathetic Innervation in Peritoneal and Deep Infiltrating Endometriosis** *PLoS ONE* **10**
46. Wu J, Xie H, Yao S, Liang Y (2017) **Macrophage and nerve interaction in endometriosis** *J Neuroinflammation* **14**
47. Aghajanova L, Giudice LC (2011) **Molecular Evidence for Differences in Endometrium in Severe Versus Mild Endometriosis** *Reprod Sci* **18**:229–51
48. Johnson MC, Torres M, Alves A, Bacallao K, Fuentes A, Vega M, et al. (2005) **Augmented cell survival in eutopic endometrium from women with endometriosis: Expression of c-myc, TGF-beta1 and bax genes** *Reprod Biol Endocrinol.* **3**
49. Kisovar A, Becker CM, Granne I, Southcombe JH (2023) **The role of CD8+ T cells in endometriosis: a systematic review** *Front Immunol* **14**
50. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. (2015) **Tissue-based map of the human proteome** *Science* **347**
51. Juárez-Barber E, Corachán A, Carbajo-García MC, Faus A, Vidal C, Giles J, et al. (2024) **Transcriptome analysis of adenomyosis eutopic endometrium reveals molecular mechanisms involved in adenomyosis-related implantation failure and pregnancy disorders** *Reprod Biol Endocrinol* **22**
52. Girling JE, Lockhart MG, Olshansky M, Paiva P, Woodrow N, Marino JL, et al. (2017) **Differential Gene Expression in Menstrual Endometrium From Women With Self-Reported Heavy Menstrual Bleeding** *Reprod Sci* **24**:28–46
53. Kobayashi H, Yamashita Y, Iwase A, Yoshikawa Y, Yasui H, Kawai Y, et al. (2012) **The ferroimmunomodulatory role of ectopic endometriotic stromal cells in ovarian endometriosis** *Fertility and Sterility* **98**:415–422
54. Franzén O, Gan LM, Björkegren JLM (2019) **PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data** *Database*
55. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. (2017) **Massively parallel digital transcriptional profiling of single cells** *Nat Commun* **8**

- 56. Fleming SJ, Chaffin MD, Arduini A, Akkad AD, Banks E, Marioni JC, et al. (2023) **Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender** *Nat Methods*
- 57. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. (2021) **Integrated analysis of multimodal single-cell data** *Cell* **184**:3573–3587
- 58. Pliner HA, Shendure J, Trapnell C (2019) **Supervised classification enables rapid annotation of cell atlases** *Nat Methods* **16**:983–6
- 59. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. (2005) **Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles** *Proc Natl Acad Sci USA* **102**:15545–50
- 60. Alpern D, Gardeux V, Russeil J, Mangeat B, Meireles-Filho ACA, Breyse R, et al. (2019) **BRB-seq: ultra-affordable high-throughput transcriptomics enabled by bulk RNA barcoding and sequencing** *Genome Biol* **20**
- 61. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. (2013) **Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool** *BMC Bioinformatics* **14**
- 62. Love MI, Huber W, Anders S (2014) **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2** *Genome Biol* **15**
- 63. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P (2015) **The Molecular Signatures Database Hallmark Gene Set Collection** *Cell Systems* **1**:417–25

## Author information

### Petra C Schwalie

Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, Roche Innovation Center Basel, Pharma Research and Early Development, Basel, Switzerland  
ORCID iD: [0000-0002-6004-8095](https://orcid.org/0000-0002-6004-8095)

### Cemsel Baflligil

Wellcome Trust Sanger Institute, Wellcome Genome Campus, Cambridge, UK

### Julie Russeil

Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

### Magda Zachara

Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

### Marjan Biocanin

Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

**Daniel Alpern**

Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

**Evelin Aasna**

German Cancer Research Center (DKFZ), Somatic Evolution and Early Detection Heidelberg, Germany

**Bart Deplancke**

Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

**Geraldine Canny**

University College Dublin, Research Office, Dublin, Ireland

**Angela Goncalves**

Wellcome Trust Sanger Institute, Wellcome Genome Campus, Cambridge, UK, German Cancer Research Center (DKFZ), Somatic Evolution and Early Detection Heidelberg, Germany  
ORCID iD: [0000-0002-3890-7658](https://orcid.org/0000-0002-3890-7658)

**For correspondence:** [a.goncalves@dkfz.de](mailto:a.goncalves@dkfz.de)

**Editors**

Reviewing Editor

**Axelle Brulport**

Senior Editor

**Diane Harper**

University of Michigan-Ann Arbor, Ann Arbor, United States of America

**Reviewer #1 (Public review):**

Summary:

The characteristics of endometrium health are an increasing topic in women's health issues, especially in the context of endometriosis. In this respect, having access to information is hampered by the inaccessibility of the uterine tissue. The authors propose here using the menstrual fluid (easily accessible by non-invasive methods) as an access door towards getting relevant information.

Overall, the paper is divided into two parts:

(1) The comparison between menstrual fluid samples and biopsies of the endometrium.

1. As a proof of concept, the authors then compared 11 controls and 7 endometriosis cases in this way, from different severity stages.

Strengths:

In Figure 1, general features of the 15 samples are presented (volume/number of cells/hematopoietic cells - cd45 labeling). The authors then used single-cell RNA-seq to characterize the different samples. Through having access to endometrium biopsies, they were able to compare the profiles obtained.

In the MF samples from the second part of the paper - aiming at comparing endometriosis and controls - one question is raised about the effect of culture. The authors compared freshly isolated and cultured tissues (ex vivo vs in vitro) by bulk RNA seq. Biases induced by the culture procedure were identified. Deconvolution was applied to strengthen this observation, with an important increase of seemingly stromal and unknown cells, especially in the unsorted cells and the CD45+ cells.

Interestingly, since the authors got successive samples from the same donor, they could evaluate the consistency of the samples and reveal indeed an overall stability of the molecular profile of the samples in a given patient.

The authors then attempted - quite originally - to characterize biomarkers in two major cell compartments that they studied - CD45- (stromal-like) and CD45+ (immune cells).

Weaknesses:

A potential problem is the justification of the a priori mix of cell types of three different phenotypes (CD45+, CD45- EPCAM+, and CD45- EPCAM-) from each patient before moving to the scRNAseq. It is not clear to me why this has been done, I guess that using directly the samples would supposedly bias the result. But in this case, why is it supposed that three categories are enough (immune cells, epithelial cells, and stromal cells)? I suppose that other markers could characterize other subtypes of the cells, and take into account the possibility of other cell types, for instance, connected to pain sensitivity, such as neuron precursor. Hence, the justification of the organized mixes should be much more detailed in my opinion.

It is a bit unclear to me when the biopsies were collected in the cycle of the donor patients.

The description of these markers that are deregulated is presented as a list, and connected with existing publications, which could rather be presented in discussion than in the results. The authors do tend to demonstrate that the Menstrual Fluid is a good proxy to analyse the endometrium health status of the women affected with endometriosis.

The identification of MTRNR2L1 seems to be a major discovery of the paper, as well as in a lesser measure HBG2, and it is a bit strange why these putative markers were not emphasized in the abstract. HBG2 was certainly identified previously in endometriosis endothelial cells but seems extremely variable from one sample to another - Geo profile (GDS3060, GDS3060 / 213515\_x\_at (inist.fr)).

Overall, the transcriptome analysis is a bit shallow, with no effort made to try to find potential transcription factors or miRNA that could activate/inhibit a series of modified genes; it could be relevant to identify such master genes or master regulators through bioinformatics analyses and wet-lab validations, to understand better the cascade of events.

Another issue that was overlooked is the presence of 'stem-cells' in the MF obtained. Since endometriosis is supposed to occur from the implantation of uterine stem cells, this category could be a major topic of scrutiny, in terms of quantity in the MF, as well as in terms of their specific molecular properties.

<https://doi.org/10.7554/eLife.99558.1.sa1>

## Reviewer #2 (Public review):

Summary:

The authors provided further evidence that menstrual fluid (MF) can be used as a non-invasive source of endometrial tissue for studying its normal physiological state and when it

is abnormal such as in endometriosis. Single-cell RNA sequencing confirmed the presence of the major cell types -blood and tissue immune cells and endometrial stromal, epithelial, and vascular cells. The major new finding was that interindividual variation for the blood immune cells was minimal between multiple MF samples from an individual. A comparison between the ex vivo MF gene profile and cultured MF showed the expected attachment and culture of stromal (and a small number of epithelial) cells, but the immune cells failed to attach. Several differentially expressed genes between controls and endometriosis were suggested as potential biomarkers of the disease, however, these were a mitochondrial pseudogene and a hemoglobin subunit, both very unlikely related to endometriosis pathogenesis.

#### Strengths:

The Spearman correlation analysis between the control MF gene profiles of multiple samples from the same individual and its graphic presentation provided strong evidence that there is little variation between MF samples. Together with another study which showed similar findings for endometrial stem cells and a number of proteins in MF supernatant, this important data shows MF as a promising biofluid for pathology testing.

The bioinformatic analyses conducted by bioinformatic and computational experts are a major strength of the manuscript and in particular the comparison between MF and endometrial biopsy data obtained from published scRNAseq studies. This is an important finding, particularly if comparisons included late secretory and early proliferative stage biopsy tissue which would be most similar to shedding menstrual endometrium.

The inclusion of workflows in the Figures for the various studies and the use of symbols in the various panels is very helpful for the reader.

MF cell suspensions were enriched for stromal and epithelial cells to enable a detailed bioinformatic analysis of their respective gene profiles

#### Weaknesses:

Two patient cohorts from different institutions were used in the study and somewhat different methods were used to extract the cellular fraction from these cohorts for further study: (1) sample dilution and differential filtration to separate blood-derived immune cells from endometrial tissue then dissociated into single cells and separated into CD45+, CD45-EpCAM+ and CD45-EpCAM- cells, and (2) gradient density separation to generate unsorted, CD45+, CD45- and putative mesenchymal stem cells (MSC) CD45-CD105+ which were also cultured. In addition, questions on pelvic pain and proven fertility would have addressed the 2 key symptoms of endometriosis.

The use of CD105 to purify MSC from MF rather than well-characterised markers of clonogenic, self-renewing, and mesodermal differentiating endometrial MSC such as CD146+PDGFRB+ or SUSD2 (both mentioned in references 22 and 23) is a weakness. The ISCT markers are not specific and are also found on stromal fibroblasts of many tissues (Phinney and Sensebe Cytotherapy 2013; Demu et al Acta Haematologica 2016).

The UMAPs generated from the scRNAseq were at low resolution and more individual immune and endometrial cell types have previously been identified and reported in MF. More comparisons with these studies would also have enhanced the Discussion.

It was not always possible to work out how the data was reported in the gene expression tables (Supplementary Tables 2, 4-10) as they were not in adjusted P value order and sometimes positive log2 fold change values appeared amongst the negative log2FC. In some comparisons described, the adj P values were not significant but were described as up or down-regulated in the text.

The 2 DEGs highlighted in the endometriosis and control arm of the study appear as poor choices from many others that could have been chosen as MTRNR2L1 is a mitochondrial pseudogene and HBG2 is a hemoglobin subunit. Neither are likely indicators of endometriosis pathogenesis.

The manuscript format and organisation could be improved by reducing the discussion in the Results section and providing a more in-depth Discussion. More references need to be included in the Discussion and other work in the MF analysis field that supports - or not - the authors' findings or at least puts them into context, and should be included and referenced.

The potential to use MF as a non-invasive source of endometrial tissue for potential diagnosis is a very important avenue of research that is currently in its infancy and could have a major impact in the endometriosis research arena.

<https://doi.org/10.7554/eLife.99558.1.sa0>