Single cell RNA sequencing and lineage tracing confirm mesenchyme to epithelial transformation (MET) contributes to repair of the endometrium at menstruation

Phoebe M Kirkwood[1], Douglas A Gibson[1], Isaac Shaw[1], Ross Dobie[1], Olympia Kelepouri[1] Neil C Henderson[1,2] and Philippa TK Saunders[1, 3]

Affiliations

[1]Centre for Inflammation Research, University of Edinburgh, Edinburgh Bioquarter, Edinburgh EH16 4TJ

[2]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

[3]Corresponding Author.

Professor Philippa TK Saunders.

Email: p.saunders@ed.ac.uk

Highlights

- Single cell RNA sequencing identified a population of PDGFRβ+ mesenchymal cells with a unique transcriptomic signature that arises in response to induction of a menses-like wound
- The repair-specific mesenchymal cells express genes considered specific to both mesenchymal and epithelial lineages indicative of an intermediate phenotype and a mesenchymal to epithelial transition (MET)
- *in silico* trajectory analysis revealed that repair-specific cells originate from the fibroblast cell clusters, were distinct from perivascular cells, and had a predicted trans-differentiation trajectory towards definitive epithelial cells
- *in vivo* lineage tracing in adult mice provides definitive evidence that PDGFRα+ endometrial fibroblasts, and not NG2+ perivascular cells, undergo MET and can become incorporated into the luminal epithelium of the post repair tissue

ABSTRACT

The human endometrium experiences repetitive cycles of tissue wounding characterised by piecemeal shedding of the surface epithelium and rapid restoration of tissue homeostasis. In this study we used a validated mouse model of endometrial repair in combination with three transgenic lines of mice to investigate whether epithelial cells that become incorporated into the newly formed luminal epithelium have their origins in one or more of the mesenchymal cell types present in the stromal compartment of the cycling endometrium. Using scRNAseq we identified a novel population of PDGFRb+ cells that arose in the endometrium in response to endometrial breakdown/repair. These cells expressed genes usually considered specific to epithelial cells and in silico trajectory analysis suggested they arose from stromal fibroblasts and were in transition to becoming epithelial cells. To confirm our hypothesis we used a lineage tracing strategy to compare the fate of stromal fibroblasts (PDGFRa+) and stromal perivascular cells (NG2+). We demonstrate for the first time that stromal fibroblasts can undergo a mesenchyme to epithelial transformation and become incorporated into the re-epithelialised luminal surface of the repaired tissue. This study is the first to discover a novel population of wound-responsive, plastic endometrial stromal fibroblasts that contribute to restoration of tissue integrity during endometrial repair. These findings form a platform for comparisons both to endometrial pathologies which involve a fibrotic response (Ashermans syndrome, endometriosis) as well as other mucosal tissues which have a variable response to wounding.

INTRODUCTION

Efficient wound repair and restoration of tissue homeostasis is essential for reproductive and general health. Repetitive injury, inflammation and fibrosis resulting in disordered tissue architecture is a major cause of morbidity and mortality due to organ failure (Greenhalgh et al., 2015). The endometrium, a complex hormone responsive multicellular tissue exhibits remarkable resilience in its ability to respond to the monthly wound of 'menstruation' without fibrosis and with a response that is both rapid and scar-free (Garry et al., 2009, Ludwig and Spornitz, 1991). A critical 'trigger' for menstruation in women is the rapid fall in circulating concentrations of progesterone that occurs as a result of the involution of the ovarian corpus luteum in a non-fertile cycle (Maybin and Critchley, 2011). Menstruation is considered an inflammatory event with a 'cascade' of changes including an increase in synthesis of pro-inflammatory factors including prostaglandins, cytokines, and chemokines which is accompanied by infiltration of immune cells of the myeloid lineage (neutrophils, monocyte/macrophages) (Evans and Salamonsen, 2012). Increased production of prostaglandins is associated with arteriole vasoconstriction which results in local hypoxia, stabilisation of HIF1alpha and increased expression of genes such as *VEGF* (Critchley et al., 2006b). Whilst we have a good appreciation of the mechanisms that contribute to the tightly regulated breakdown and shedding of a portion of the endometrium at menses gaps remain in our understanding of the processes that contribute to rapid, fibrosis/scar-free restoration of an intact luminal epithelium over the tissue surface. It is also important to separate the mechanisms responsible for menstrual repair, which occurs at a time in the cycle when ovarian-derived steroids are low, from endometrial regeneration/proliferation which occurs in response to rising levels of follicle-derived oestrogens during the proliferative phase of the cycle (Maybin and Critchley, 2012).

Several different mechanisms have been proposed as contributing to the luminal repair process including, activation of progenitor/stem cells in the stroma and/or epithelium, epithelial cell

proliferation/migration and transformation of stromal cells into epithelial cells (mesenchyme to epithelial transition, MET) each of which are discussed briefly below.

The ability of the endometrium to rapidly regenerate at the end of each menstrual cycle has prompted researchers to explore a role for putative stem/progenitor cells residing in either the stroma and/or the epithelial compartments in endometrial repair (Gargett et al., 2016, Cousins et al., 2021). The Gargett group have identified populations of endometrial mesenchymal stem cells; eMSC) with clonogenic and broad multilineage potential in both endometrial tissue and menstrual fluid and characterised them as PDGFRB+/CD146+/ SUSD2+ (Gargett et al., 2016). Notably in human endometrial tissue sections these cells were identified as having a perivascular location and a putative pericyte identity. A role for progenitor cells residing in the basal glandular epithelium has also been proposed with a recent review highlighting the different marker proteins employed to establish their identity including SOX9 and SSEA1 (Cousins et al., 2021). Another source of putative stem cells that has been considered as contributing to endometrial repair processes and/or endometrial pathology is the bone marrow (Kaitu'u-Lino et al., 2012) although evidence for the precise role of the bone marrow in this context is somewhat limited (Deane et al., 2019, Aghajanova et al., 2010).

During the first two days of menstruation rapid shedding of the luminal portion of the tissue (functional layer) leaves a denuded (basal) stroma interspersed with glandular 'stumps'. This phenomenon was beautifully documented in the elegant studies of Ludwig and Spornitz (Ludwig and Spornitz, 1991). In the 1970's Ferenczy proposed that new epithelial cells arose from proliferation of intact surface epithelium bordering denuded stroma that was not lost during menstruation and the basal glands that were exposed when the functional layer was shed (Ferenczy, 1976b, Ferenczy, 1976a). He noted that the re-epithelialisation took about 48 hours and did not appear to involve mitosis of the epithelial cells. These mechanisms of epithelial repair were later endorsed by Ludwig and Spornitz who also observed fibrin mesh formation and in the early days of menstruation growth of an epithelial monolayer in spirals consistent with circle-structures arising from the endometrial glands (Ludwig and Spornitz, 1991). When Garry and colleagues used a more dynamic approach involving pressure-controlled, continuous flow hysteroscopy they found evidence that loss and repair of the surface occur synchronously at different places within the endometrial cavity a process they called 'piecemeal' (Garry et al., 2009). They reported epithelial cells of the glands underwent apoptosis and were shed with the decidual. Apoptosis of glandular epithelial cells in carefully dated human endometrium has also been documented using staining for cleaved caspase 3 and shown to be high in both late secretory and menstrual phases (Armstrong et al., 2017) questioning their ability to act as the source of new luminal cells.

There is also evidence that cells within the stromal compartment may change their identity to adopt an epithelial phenotype (mesenchyme to epithelial transition, MET) in parallel with epithelial cell proliferation and migration and thereby contribute to the luminal surface. This was first proposed in the 1960's by Baggish *et al* who evaluated sections of menstruating human endometria and noted hyperchromatic stromal cells between gland stumps that appeared to take part in re-epithelisation by a process simulating metaplasia (Baggish et al., 1967). In the 1970's the involvement of the stroma was challenged (Ferenczy, 1976b) but has been supported by more recent studies (Garry et al., 2009). Specifically in 2009 the authors reported that the denuded basalis endometrium was rapidly covered with a fibrinous mesh within which new surface epithelial cells were present which appeared to arise from the underlying stroma consistent with some cells responding to the menstrual 'wound' by undergoing MET.

To complement studies on human tissue samples and studies in primates (Brenner et al., 2002, Critchley et al., 2006a) we have used a highly reproducible mouse model that recapitulates key features of human menstruation in response to withdrawal of progesterone including vascular breakdown (vaginal bleeding), spatial and temporal hypoxia, cell apoptosis, expression of matrix metalloproteinases and influx of myeloid immune cells (Armstrong et al., 2017, Maybin et al., 2018, Cousins et al., 2016a, Cousins et al., 2016b, Cousins et al., 2014). Importantly, as in the human endometrium (Garry et al., 2009), restoration of an intact luminal epithelial layer occurs within two days (Cousins et al., 2014). Consistent with some of the reports on repair processes in human endometrium we found evidence for epithelial cell migration and proliferation of residual epithelial cells (Ki67+) but no significant proliferation of basal glands (Cousins et al., 2014). One of the key observations we made when exploring the architecture of the tissue in regions that were denuded of epithelium was the existence of stromal cells that co-expressed vimentin (stromal cell marker) and cytokeratin (epithelial cell marker) which prompted us to support the suggestion made by others that some stromal cells might be changing their phenotype via MET (Cousins et al., 2014). Other groups have also used similar mouse models to explore the dynamic changes during a menstrual-like event. Kaitu'u-Lino *et al* used a mouse model in combination with pulse labelling of cells with BrdU, and reported that a population of epithelial progenitor cells residing in the basal glands might contribute to postmenstrual repair (Kaitu'u-Lino et al., 2010, Kaitu'u-Lino et al., 2012).

In the current study we have used single cell RNA sequencing and lineage tracing to specifically address which, if any, of the stromal mesenchymal cell types we previously identified in the mouse endometrium (Kirkwood et al., 2021) might contribute to the rapid restoration of the intact luminal epithelial cell layer observed in our mouse model (Cousins et al., 2014, Cousins et al., 2016a). To the best of our knowledge this is the first time that inducible-Cre lineage tracing approaches in adult mice have been used to explore MET in a model of menstruation.

MATERIALS AND METHODS

Mouse lines

In *Pdgfrb*-BAC-eGFP reporter mice, eGFP expression is driven by the regulatory sequences of the *Pdgfrb* gene and is therefore expressed by all cells in which the promoter is active. Details of these transgenic mice and their use for studies on the endometrium of cycling mice have been previously described (Kirkwood et al, 2021). These mice were originally obtained from GENSAT and deposited in MMRRC-STOCK Tg(Pdgfrb-EGFP)JN169Gsat/Mmucd, 031796-UCD). Founder stocks were obtained from Professor Neil Henderson (CIR, University of Edinburgh).

For lineage tracing experiments mice from two lines that expressed Cre recombinase in response to a Tamoxifen inducible promoter were cross bred with the Ai14 reporter mouse line in house. Briefly, Ai14 is a Cre reporter allele designed to have a loxP-flanked STOP cassette preventing the transcription of a CAG promoter-driven red fluorescent protein variant (tdTomato)- all inserted into *the Gt(ROSA)26Sor* locus. Ai14 mice express robust tdTomato fluorescence following cre-mediated recombination. These mice were originally obtained from the Jackson Laboratory [STOCK *Gt(ROSA)26Sor*$^{tm14(CAG\text{-}tdTomato)Hze}$]. Founder stocks were obtained from Dr David Ferenbach and Dr Bryan Conway (University of Edinburgh).

We had previously identified expression of NG2*(Cspg4)* and *Pdgfra* as specific to endometrial perivascular and fibroblast populations of mesenchymal cells respectively (Kirkwood et al., 2021). We therefore selected mice which would express CRE recombinase in response to induction of the promoters of these genes as appropriate for segregating cells representing these

two subpopulations in our mouse model of endometrial repair. i*Pdgfra*Cre-ER^{T2} mice express a tamoxifen-inducible Cre recombinase (CreER^{T2}) under the control of the mouse *Pdgfra* promoter/enhancer. The generation of i*Pdgfra*Cre-ER^{T2} mice has been described previously (Chung et al., 2018). Founder stocks were a gift to Professor Neil Henderson from Professor Brigid LM Hogan and Dr Christina E Barkauskas (Duke University, USA) . NG2Cre-ER^{TM} BAC transgenic mice express a tamoxifen-inducible Cre recombinase (CreER^{TM}) under the control of the mouse NG2 (*Cspg4*) promoter/enhancer. The generation of NG2Cre-ER^{TM} mice has been previously described (Zhu et al., 2011). Founder stocks were obtained from the Jackson Laboratory [STOCK B6.Cg-Tg(Cspg4-cre/Esr1*)BAkik/J].

Tamoxifen induction of tdTomato transgenes

Tamoxifen (Sigma, Cat. #T5648-1G, TMX) was dissolved in sesame seed oil to a concentration of 20mg/ml and left overnight on a roller at 37°C (light protected). For TMX administration, in initial screening experiments an intraperitoneal (IP) injection of 100µl was given to each mouse (standard 20g weight) for 3 consecutive days (TMX dose 100mg/kg). In later experiments the method was refined and TMX was administered via oral gavage (daily, 3x 100µl/day). To optimise an appropriate TMX wash out period prior to the mouse model of 'menstruation' (endometrial shedding and repair), uterine weights and the stromal to epithelial cell ratio (as determined by flow cytometry analysis) were analysed in mice treated with TMX or vehicle (sesame oil) and uterine tissues collected 24h, 1 week, 4 weeks and 8 weeks later. After 4 weeks following the final tamoxifen administration tissue analysis was comparable to controls (Vehicle treated, no TMX). To validate timing so that the mouse model of 'menstruation' still worked, the decidualisation response (≥1 uterine horn decidualised) and corresponding uterine weights were analysed in mice in which reporter expression was induced with TMX and uterine tissues collected 24, 48 and 72h following the withdrawal of progesterone (onset of 'menses) and compared to those where no TMX administration had occurred. No significant differences were observed indicating that a tamoxifen induction 4 weeks prior to the start of the mouse model of 'menstruation' had no deleterious effects on the model.

Mouse model of 'menstruation' (endometrial shedding and repair)

All animal procedures were carried out in accordance with the Animal Welfare and Ethical Review Body (AWERB) and under licensed approval from the UK Home Office. The mouse model of 'menstruation' (Figure 1A(i)) was conducted according to standard protocols that have been described in detail in previous papers (Cousins et al., 2016a, Cousins et al., 2016b, Cousins et al., 2014). Briefly, mice between 8-12 weeks of age were ovariectomised (d0) and given daily injections of β-oestradiol (E2) in sesame oil (100ng/100µl, d7-9). A progesterone (P)-secreting pellet was placed sub-cutaneously on d13; mice also received daily injections of E2 (5ng/100µl; d13-15). A critical step in the model is the transformation of stromal fibroblasts into cells with a decidual phenotype which was achieved using introduction of a small amount of oil via the vagina (d15). Removal of a progesterone-secreting pellet 4 days later (d19) results in a rapid fall in progesterone culminating in breakdown and shedding of the decidual mass leaving parts of the endometrial lining with a surface denuded of epithelial cells (24h; (Cousins et al., 2014)). Within 3 days (72h after pellet removal) the integrity of the luminal epithelial layer is fully restored and gross uterine architecture resembles that of mice following ovariectomy (Figure 1A (ii)).

In those studies using mice in which transgene expression was induced using TMX the protocol was modified to ensure there were no agonist or antagonist impacts of the drug on the endometrium (Gielen et al., 2008) which could alter responsiveness to exogenous E2. Pilot

studies were conducted and a 4w wash out period applied to all TMX-treated mice before they were used in the menstruation model.

Tissue recovery and processing

To preserve and detect fluorescent signal from transgenes, uteri were fixed in 4% (w/v) PFA for 2h at 4°C, rinsed thoroughly in PBS and stored overnight in 18% (w/v) sucrose at 4°C before embedding in OCT medium and storage (-80°C). Frozen tissue sections (5µm) were cut and mounted onto Xtra adhesive pre-cleaned micro slides (Surgipath, Leica Biosystems) and air-dried at room temperature for 30 minutes prior to staining.

Analysis of tissue sections for expression of fluorescent transgenes and proteins associated with specific lineages

Analysis of transgene expression was complemented by co-staining with selected antibodies according to established protocols (Cousins et al., 2014). Briefly, for immunofluorescence, tissue sections were washed in PBS to remove residual OCT and incubated with 3% (v/v) hydrogen peroxide solution in methanol for 30 min at room temperature, washed (1x5min wash in PBS containing 0.05% (v/v) Tween20 and 2x5min washes in PBS) and further incubated with 20% (v/v) normal goat serum for 30 min at room temperature. Sections were washed and incubated overnight at 4°C with primary antibody at an optimised dilution in NGS; antibodies against EPCAM (epithelial cells), CD146 (endothelial/perivascular cells), PDGFRα (fibroblasts) and NG2 (perivascular cells) were used in this study (Details in Supplementary Table 1). Following a further wash, sections were incubated with an HRP-conjugated secondary antibody at an optimised dilution in NGS for 1 h at room temperature followed by a 10 min incubation with Opal Polaris 480 solution (Ayoka Biosciences). After staining sections were counterstained with DAPI (Sigma, D9542), overlaid with Permaflour (Immunotech) and mounted with coverslips (VWR Prolabo) before imaging on a Zeiss LSM 510 Meta Confocal microscope (Zeiss).

Cell isolation and purification using FACS

Tissue processing for flow cytometry analysis and FACS was performed as previously described (Cousins et al. 2016, Kirkwood et al, 2021). Briefly, whole uterine horns were dissociated and incubated with collagenase (10mg/ml) and DNase (10mg/ml) for 30 min at 37°C. Tissues were further dispersed and washed in FACS buffer (PBS Ca2-Mg2-: 5% charcoal stripped foetal calf serum (CSFCS), 2mM EDTA) and subsequently strained through 70µm and 40µm cell strainers. Cell suspensions were centrifuged at 400rcf for 5 min and cell pellets re-suspended in 1ml ACK lysing buffer (Gibco, Cat. No. A10492-01) for 1 min. Suspensions were then washed (as above) and incubated for 30 min on ice with optimised dilutions of fluorescently-conjugated antibodies as detailed in Supplementary Table 2. Cell suspensions were washed and re-suspended in PBS at 4°C and analysed using a BD 5L LSR Fortessa and BD FACSDiva software (BD Biosciences). To exclude dead cells, DAPI was added prior to flow cytometry analysis. Cells were sorted using a FACS Aria II instrument and BD FACSDiva software (BD Biosciences). Data analysis was performed using FlowJo analysis software (FlowJo LLC, Oregon, USA) and GraphPad Prism.

Single cell RNA sequencing- 10x Genomics

Endometrial mesenchymal cells (GFP+) were isolated from *Pdgfrb*-BAC-eGFP mice 24 and 48h after removal of a progesterone pellet (times chosen to model active shedding/breakdown and active repair respectively; Figure 1A). For identification of definitive endometrial epithelial cell populations GFP-EPCAM+ cells were isolated from *Pdgfrb*-BAC-eGFP mice 48h after removal of a progesterone pellet (time chosen to reflect complete re-epithelialisation) and cycling/control mice. Cells were recovered from 4 mice at each stage using FACS; 25,000

cells/sample giving a total of 100,000 cells for each population of interest for downstream application. Isolated cell suspensions were counted, and viability confirmed to be >85% using a TC20™ Automated Cell Counter (BioRad, Cat. No. 1450102).

Cells were partitioned into nanolitre-scale Gel bead-in-Emulsions (GEMs) containing unique 10x barcodes using the 10x Chromium™ Controller (10x Genomics, USA) and cDNA libraries generated as previously described in (Kirkwood et al., 2021). Sequencing was carried out on the Illumina NovaSeq platform using bespoke 10x parameters, according to the standard protocols at the facility (Edinburgh Genomics: http://genomics.ed.ac.uk/, Edinburgh).

Data files generated for mesenchymal cells (GFP+) from *Pdgfrb*-BAC-eGFP mice 24 and 48h after removal of a progesterone pellet were compared to cycling/control data files generated in previous studies which are available at GEO Series accession number GSE160772 (Kirkwood et al, 2021).

Bioinformatic workflows

Raw sequencing data files were processed using 10x Cell Ranger following standard protocols (version 2.0.1; https://www.10xgenomics.com). Genome alignment, filtering, barcode and UMI counting were performed using the 'cellranger-count' function and 'refdata-cellranger-mm10-1.2.0' transcriptome as supplied by 10x genomics. For downstream QC, clustering and gene expression analysis the *Seurat* R package (V3) (Stuart and Satija, 2019) was utilised with R version 4.0.2 in line with analyses performed as in (Kirkwood et al., 2021). Briefly, the data was filtered based on standard QC metrics: number of unique genes detected in each cell, nFeature_RNA >200 & <5000; the percentage of reads that map to the mitochondrial genome (<10%); and the percentage of reads that map to ribosomal proteins (<10%). The *Scrublet* python package was used to further identify and remove putative doublets from scRNA seq data (Wolock et al., 2019). Following filtering, data was normalised and scaled and PCA analysis used to detect highly variable features/genes. Unsupervised clustering based on the first 20 principal components of the most variably expressed genes was performed using a graph based approach ('FindNeighbours', 'FindClusters'; resolution= 0.2). Clusters were visualised using the manifold approximation and projection (UMAP) method and identified by analysing the expression of canonical cell markers.

Differential gene expression analysis was performed ('FindAllMarkers') to identify genes expressed by each cell cluster when compared to all other clusters, using the non-parametric Wilcoxon rank sum test and p-value threshold of <0.05. Over-represented functional annotations in the differentially expressed genes were detected using the *clusterProfiler* package (Yu et al., 2012) using core functions to interpret data in the context of biological function, pathways and networks. In silico trajectory analysis was performed using *Monocle3* (Qiu et al., 2017). Monocle3 works on the basis that cells in different functional 'states' express different sets of genes, and if/as cells move between these states, they undergo a process of transcriptional re-configuration where some genes are activated whilst others are silenced. Monocle3 uses an algorithm to learn if there is a sequence of gene expression changes between single cells that connect clusters through a putative differentiation trajectory.

RESULTS

Single cell RNA sequencing identifies novel mesenchymal cell subpopulations unique to 'repairing' tissue in a mouse model of endometrial breakdown and repair (menstruation)

To identify whether changes in the transcriptome of mesenchymal cells occurred during endometrial repair we recovered tissues from mice expressing the *Pdgfrb*-BAC-eGFP transgene during the normal cycle (controls) and at 24 and 48h after progesterone withdrawal

using a well validated model of endometrial shedding/repair (simulated 'menstruation'; (Cousins et al., 2016b)) (Figure 1A(i, ii)). Population restricted cell sorting was used to isolate CD31-/CD45-/GFP+(PDGFRβ+) cells prior to single cell sequencing. In cycling mice we have previously reported GFP+ cells are restricted to the stromal compartment of the endometrium (Kirkwood et al., 2021).

scRNAseq was used to generate a dataset for each timepoint and these were then integrated for downstream cluster and differential gene expression analysis. This strategy resulted in identification of 8 transcriptionally distinct clusters of GFP+ cells (Figure 1B). The expression of canonical phenotypic gene markers for mesenchymal, perivascular and fibroblast cell populations previously identified in control/cycling mouse endometrium was used to assign putative identities to each cluster (Kirkwood et al., 2021). This strategy identified 3 perivascular populations (V, P1, P2; expression of *Mcam*, *Cspg4*, *Rgs4*, *Acta2*) and 5 fibroblast subpopulations (F1-5; expression of *Pdgfra*, *Cd34*, *Fbln1*/2, *Mfap4*/5, *Col1a1*). Importantly, cells in all 8 clusters expressed mesenchymal markers *Pdgfrb*, *Vim*, *Des* and *Thy1* (Figure 1C, Figure S1A(iii)) while no clusters expressed *Pecam1* (CD31) or *Ptprc* (CD45) (Supplementary Figure 1A(i-ii)) consistent with the cell isolation strategy employed.

Analysis of the distribution of these subpopulations in the different datasets (Figure 1D) confirmed the presence of 5 mesenchyme subpopulations previously identified in cycling mouse uterine tissue (V, P1, F1-3; (Kirkwood et al., 2021)). Notably in tissue recovered at 24 and 48h three transcriptionally distinct repair-specific cell clusters were identified: P2 and F4 in 24h tissues (Figure 1D; green and purple circles respectively) and F5 present in 48h tissues (Figure 1D; pink ellipse).

Gene expression analysis identified a unique repair-associated cluster that expressed genes associated with epithelial cell phenotype

Differential gene expression analysis revealed a very high degree of similarity between P1 (previously identified as pericytes) (Kirkwood et al., 2021) and P2 which was specific to 24h tissue (Figure 2A). Unbiased GO analysis of the P2 transcriptome identified functions associated with blood vessel formation and immune cell signalling, characteristic of pericytes (Supplementary Figure 1D(i)) suggesting this is not a phenotypically distinct cell population but one with some changes in gene expression in response to changes in the local environment.

We noted that cluster F5 (specific to 48h tissue) expressed a large number of both mitochondrial genes and genes associated with cell death including *Casp3*, *Casp9*, *Trp53* and *Bax* (Supplementary Figure 1B-C). GO analysis of the F5 transcriptome identified functions associated with DNA damage. These data suggest that F5 represents a cluster of dead/apoptotic cells and it was therefore excluded from further downstream analyses.

Cluster F4 (24h tissue) was of particular interest as it had a transcriptomic signature unique to those of the mesenchymal cell populations (F1-3) present in cycling endometrium (Kirkwood et al, 2021) (Figure 2A; red box). Unbiased GO analysis enriched for biological functions that are associated with endometrial tissue repair including response to wounding, reproductive structure development, response to oxidative stress and regulation of vasculature development. Additional functions that were identified included epithelium migration, morphogenesis of a branching epithelium and regulation of epithelial cell migration (Figure 2B).

Cluster F4 was the only mesenchymal cell cluster that expressed markers known to be associated with both definitive mesenchymal (*Pdgfrb*, *Vim*) and epithelial (*Epcam*, *Krt18*) cell types (Figure 2C). Comparison across all mesenchymal clusters of the expression of both known canonical mesenchymal/fibroblast markers including *Thy1*, *Cd34*, *Cdh11, Col1a1, Col3a1, Fbln1, Fbln2, Mfap4* and *Mfap5* and epithelial markers such *as Krt8, Krt19, Muc1,*

Muc4, Cd24a, Car2, Sdc4, Alcam and *Cdh1* revealed F4 to be the only cluster that expressed genes associated with both cell types (Figure 2D; red box). We considered this was evidence that these cells may be of an intermediate phenotype and represent cells undergoing a mesenchymal to epithelial transition (MET) and therefore a did a further analysis of annotated gene sets from the MSigDB (Subramanian et al., 2005) that are known to be associated with the regulation of MET/EMT. This analysis identified that F4 also expressed key transcription factors associated with MET/EMT including *Cemip, Cited1, Ctnnb1, Grem1, Hmga2, Klf4* and *Twist2* (Figure 2E).

Cells that co-express both mesenchymal and epithelial cell markers can be detected in the uterus during endometrial tissue repair

To validate this finding and analyse the topography of putative MET cells in the mouse uterus, tissue sections *Pdgfrb*-BAC-eGFP uterine tissues were co-stained with antibody directed against EPCAM. Cells that expressed both GFP and EPCAM were identified in uterine tissues undergoing active tissue repair (24h; Figure 3A(i); white arrows) specifically in regions where (a) the stromal surface was denuded after decidual shedding and (b) a new epithelial cell layer was being re-instated. Some GFP+EPCAM+ cells persisted in the luminal layer at 48h but were not detected in 72h tissue (Figure 3A ii, iii). Consistent with previous data in control/cycling endometrium expression of the GFP reporter was exclusive to the stromal compartment and epithelial cells were EPCAM+ (yellow arrows, Figure 3A(iv) (Kirkwood et al., 2021).

Analysis of the same tissues using flow cytometry confirmed the presence of a new population of cells that expressed both GFP and EPCAM at 24h with a reduction in their number at 48h and levels that appeared comparable to cycling/control tissues in the fully remodelled tissue (72h) (Figure 3B-C). These data are in agreement with the results of scRNAseq analysis where *Epcam* was detected in the F4 cluster in the 24h dataset (Figure 2C) with analysis of the tissue sections suggesting the F4 cluster may represent a mixture of cells in the stroma in the process of undergoing MET (Figure 3Aa) as well as those which are already incorporated into the luminal epithelium (Figure 3Aa) where they are still detectable at 48h (Figure 3Aii). Whilst promising studies on cell fate using mice from the *Pdgfrb*-BAC-eGFP 'knock-in' line has limitations as expression of GFP is limited to cells in which the *Pdgfrb* is active and immunohistochemistry has revealed that the PDGFRB protein is only present in endometrial mesenchyme (see Figure 1 in (Kirkwood et al., 2021)). With high expression of *Pdgfrb* in the mesenchyme and some expression persisting in F4 (Figure 2C) we speculated that presence of GFP in the luminal epithelial cells at 24 and 48h represented mesenchyme cells that underwent a rapid MET. It was not possible to determine whether these cells did/or did not persist at 72h as absence of GFP might be due to protein turnover and therefore additional studies with new mouse lines were necessary to resolve this question.

Comparison between fibroblast and epithelial cell scRNA clusters confirmed the repair-specific cells (F4) had a unique transcriptome.

To compare the transcriptome of the F4 cluster (putative MET mesenchyme/epithelium) to differentiated endometrium epithelial cells scRNAseq datasets were generated by from CD31-/CD45-/GFP-/EPCAM+ epithelial cells obtained from the *Pdgfrb*-BAC-eGFP mouse uterus (control/cycling and 48h). Bioinformatic analysis combining GFP+ datasets (cycling/24h/48h) and GFP-EPCAM+ epithelial cell data (cycling/48h) identified 16 distinct cell clusters (Figure 4A, C). Based on expression of *Pdgfrb* (Fiture 4Bii) and canonical markers (Figure 4D) 7 of these had a mesenchyme phenotype : three perivascular (V, P1-2), four fibroblast (F1-4). Nine of the clusters expressed classical markers of epithelial cells including Epcam (E1-9) (Figure 4A, B(i, ii), C). Notably, the previously identified repair-specific fibroblasts (F4) formed a cluster which appeared in close proximity to both fibroblast and epithelial cell clusters but

remained distinct from definitive perivascular cells (V, P1, P2). Doublet analysis was also conducted to confirm this was a unique population (Supplementary Figure 2A).

A high degree of similarity was observed between the transcriptomic signature of F4 and the definitive epithelial cell clusters E1 and E2 (Figure 4C, red box, Supplementary Figure 2B). Extended analysis of canonical markers of mesenchymal (perivascular/fibroblast) and epithelial cell types confirmed that cluster F4 uniquely expressed genes associated with both fibroblast and epithelial lineages (Figure 4D, Figure S2A(i-iv)). Additional analysis of genes associated with different epithelial subtypes present in the endometrium inferred putative epithelial phenotypes to each of the clusters: E1/2/3 luminal/differentiating, E4 proliferative, E5/6 glandular/ciliated, E7 secretory, E8/9 basal (Figure 4E). Notably the repair-specific mesenchyme cluster F4 also expressed genes associated with both luminal (*Fxyd4, Gsto1, Plac8, Ctsl, Lgals3, Cited2*) and differentiating (*Ctnnb1, Krt8, Krt18, Plk2, Basp1, Ptgs2, Olfm1*) epithelial cell phenotypes with no expression of genes considered definitive for glandular epithelium (*Foxa2*, *Cxcl15*).

In silico trajectory analysis infers a putative differentiation trajectory between fibroblasts and epithelial cells

Based on the pattern of gene expression and clustering of scRNA datasets generated from the purified mesenchyme and epithelial cells populations we made a preliminary assessment that the F4 population were an intermediate cell type. To investigate this further we used the Monocle3 package to perform *in silico* lineage tracing identify whether an inferred differentiation trajectory existed between the different cell clusters.

Monocle3 analysis segregated the scRNAseq data into two partitions: one that encompassed the perivascular subsets; and one that contained both the fibroblast and epithelial subsets. We infer that this would be consistent with a pattern of genetic changes linking fibroblasts and epithelial cells but none between either fibroblasts or epithelial cells and the perivascular clusters (Figure 5A(i-ii)). Further interrogation of the Monocle data identified a 'root' within fibroblast cluster F2 with a branch moving through the repair-specific cluster F4 into a root in the epithelial clusters on the border of E1 and E2, with subsequent branches moving into other epithelial subsets (Figure 5A(iii)). This bioinformatic analysis suggests that a cellular transition from F1/2/3 through F4 into E1/2 may exist. No roots were placed within the F4 cluster (only a trajectory branch; Figure 5A(i-iii)) which further suggests this population does not have a terminally differentiated phenotype, rather it represents a transitional transcriptional state.

The results described above would be consistent with the presence of a new population of fibroblast origin that has a distinct transcriptional profile that exists as a transient/intermediate population in response to the altered environment that was induced by progesterone withdrawal. To confirm the origin and fate of these cells we adopted a new strategy using *in vivo* genetic fate-mapping based on induction of cre recombinase under the control of promoters specific to genes we have identified as specific for fibroblast (*Pdgfra*) and perivascular (*Cspg4*/NG2) cell clusters (Figure 5B, C; (Kirkwood et al., 2021))

Induction of tdTm transgenes using a iCRE strategy to separately target fibroblast and perivascular cell populations

To validate the *in silico* trajectory analysis we used an *in vivo* genetic fate-mapping strategy using the TMX-inducible cre recombinase system (Figure 5B). Based on our scRNAseq all mesenchymal cell populations express *Pdgfrb* but we were able to discriminate between fibroblasts (F1-F4) and perivascular cells based on expression of *Pdgfra* and *Cspg4/*NG2 respectively (Figure 5C). Importantly neither of these markers were expressed in any of the epithelial cell clusters. We specifically targeted these two mesenchymal cell types in adult

female mice using a tamoxifen induction strategy (*iPdgfra*-creER^{T2}/*Rosa26-tdTm;* iNG2-creER^{TM}/*Rosa26-tdTm* mice respectively).

Prior to lineage tracing studies conducted experiments to optimise the amount of TMX administered both to confirm long term expression of the tdTM protein transgenes but also that the endometrial tissue of the TMX-treated mice was still able to respond appropriately to E2 and the decidualization stimulus. Specifically, we compared the uterine response (weight, stromal-epithelial cell ratio, decidualization response) following 3x daily oral gavage with 20mg/ml TMX) 24h, 1w, 4w and 8w after the last dose (Supplementary Figure 3 D-H); controls had no TMX. In line with expectations treatment with TMX, which is known to act as an oestrogen receptor agonist in endometrial tissue (Gielen et al., 2008), increased uterine weights at 24h, this increase was still noticeable at 1w but had returned to baseline by 4w (Supplementary Figure 3D). The weight change was associated with an increase in EPCAM+ epithelial cells (Supplementary Figure 3E-G) which also returned to baseline by 4w. Further tests showed decidualization response and changes in weights following progesterone withdrawal were the same as controls if mice had been treated with TMX 4w previously (Supplementary Figure 3 H, I).

Subsequent experiments to confirm the identity of the stromal cells in both iCre models and to investigate their fate following induction of endometrial breakdown/repair were all conducted using 3x daily gavage with TMX followed by a 4w washout period (Figure 5G). In the *iPdgfra*-creER^{T2}/*Rosa26-tdTm* mouse uterus tdTm expression was detected in cells dispersed throughout the endometrial stromal compartment, but not in cells of the myometrium or luminal/glandular epithelium (Figure 5D(ii)). This pattern replicated the location of GFP^{dim} stromal fibroblasts in the *Pdgfrb*-BAC-eGFP mouse uterus (Figure 5D(i)) We confirmed co-localisation of tdTm and native protein PDGFRα (Figure 5E). We also confirmed that these tdTm+ cells did not express CD146 (Figure S3A(iii), C) and concluded that the tdTM transgene is specific to PDGFRα+ endometrial fibroblasts.

In the iNG2-creER^{TM}/*Rosa26-tdTm* mouse uterus tdTm was detected in groups of cells dispersed throughout the stromal compartment but not the myometrium or epithelium (luminal/ glandular) (Figure 5D(iii)). This pattern was similar to perivascular GFP^{bright} cells in the *Pdgfrb*-BAC-eGFP mouse uterus (Figure 5D(i); Kirkwood et al 2021), and therefore appeared to represent perivascular cells. To confirm this we co-localised the tdTm reporter with native NG2 protein (Figure 5F). We also confirmed they express the canonical pericyte/vSMC marker CD146 (Supplementary Figure 3A(ii), C) but not the endothelial cell marker CD31. Taken together, we concluded that induced tdTm expression in the iNG2-creER^{TM}/*Rosa26-tdTm* mouse uterus correctly targets NG2+ perivascular cells (pericytes/vSMC).

Lineage tracing in adult mice confirms MET by PDGFRα+ fibroblasts during endometrial tissue repair.

Following transgene induction in *iPdgfra*-creER^{T2}/*Rosa26-tdTm* mice tdTm was detected in cells throughout the stromal compartment in both controls (Figure 5D) and at 24h after progesterone withdrawal (Figure 6A). Notably in the 24h tissue a unique population of tdTm+ cells that co-expressed the epithelial cell marker EPCAM (tdTm+EPCAM+ cells) were detected in two regions one where the decidualised tissue had detached leaving the stromal surface denuded (Figure 6A(i-iii); white arrows) and the second population having a phenotype that resembled a new epithelial layer which in some cases was continuous with tdTm-EPCAM+ cells (Figure 6A(iv); white and yellow arrows respectively). In other regions epithelial cells that did not express tdTM were present (Figure 6A(v); yellow arrows).

For independent validation of the tdTm staining of tissue sections flow cytometry was used to quantify the numbers of *tdTm+* cells that expressed the mesenchymal marker CD90 and/or the epithelial marker EPCAM in controls (no repair) and in the repair model at 24, 48 and 72h (Figure 6B; red boxes). In controls tdTm+ cells were all CD90+ (Figure 6Bi) A transient population of tdTm+CD90+EPCAM+ cells were detected at 24h (red box) which represented ~20% of the reporter+ population (Figure 6C). Notably whilst some of the tdTM+ cells continued to co-express EPCAM at 48 or 72h tissue these cells were CD90- (Figure 6C iii-iv).

When identical analysis methods were applied to tissue recovered from the iNG2-creER^{TM}/*Rosa26-tdTm* mice cells with a tdTm+CD90+ phenotype were identified in controls and at 24, 48 and 72h after induction of repair (Figure 6D) however using sensitive flow cytometry no tdTM+EPCAM+ cells were detected nor were any tdTm+ cells present in the epithelial cells at any timepoint (Figure 6D. F; Supplementary Figure 4A).

Lineage tracing confirms that epithelial cells originating from PDGFRα+ fibroblasts are retained within the repaired luminal epithelium.

To understand if the stromal fibroblasts we had detected undergoing MET in the *iPdgfra*-creER^{T2}/*Rosa26-tdTm* persisted in the luminal epithelium once the layer was fully restored (48-72h) we conducted a detailed analysis of tissue sections and additional flow cytometry. We identified tdTm+ cells throughout the luminal epithelium in tissues at 48h adjacent to tdTm-epithelial cells (Figure 7A(i-iv); white and yellow arrows respectively). These two populations of cells were also detected in uterine tissues at 72h (Figure 7B(iv). Notably, tdTm expression was not identified in epithelial cell populations at the time of progesterone withdrawal (0h; Figure 7B(i); yellow arrows). Quantification of tdTM reporter protein in EPCAM+ epithelial cells revealed time dependent changes with proportions rising from ~6% at 24h to ~17% at 48h and ~15% at 72h (Figure 7C).

The presence of groups of tdTm+EPCAM+ adjacent to tdTm-EPCAM+ cells in the newly formed luminal epithelium at 48 h and their persistence at 72h is consistent with both MET by PDGFRα+ fibroblasts and retention/proliferation of existing epithelial cells both contributing to rapid repair. Importantly we never detected, an equivalent population of tdTm+EPCAM+ cells in iNG2-creER^{TM}/*Rosa26-tdTm* mice at any timepoint (Supplementary Figure 4A-E).

DISCUSSION

In this study we used a combination of single cell RNA sequencing and cell lineage tracing to investigate the contribution of the stromal mesenchyme to rapid restoration of the endometrial luminal epithelium at the time of menstruation. The human endometrium is unusual in its ability to experience cyclical episodes of breakdown/shedding (wounding) which occur in parallel with repair processes that leave the tissue intact without an obvious scar. We used a previously validated mouse model that mimics the key features of human menstruation in which we made some preliminary observations that aligned with studies suggesting MET of stromal cells might contribute to the rapid repair process (Cousins et al., 2014). In the work presented here we have identified a new population of repair-specific PDGFRb+ cells that responded to changes in the endometrial environment induced by progesterone withdrawal by upregulating expression of genes associated with MET/epithelial cell identify. To best understand how this population might influence our understanding of endometrial repair we designed further experiments to answer two specific questions. First – what was the originating cell type of this transient population? Second – what was its fate?

As the initial objective of the study was to explore whether we could substantiate our hypothesis that MET was occurring in stromal cells exposed by shedding of the decidual mass in our mouse model of menstruation (Cousins et al., 2014). Therefore we decided to use

Pdgfrb-BAC-eGFP reporter mice in our menstrual model as we had previously shown eGFP was only expressed in stromal cell populations during the oestrus cycle (Kirkwood et al., 2021). Using these mice revealed two new findings, first that a scRNAseq of GFP+ cells identified three new populations of mesenchyme cells in tissues recovered 24 or 48h after withdrawal of progesterone at a time when the tissue in the process of active breakdown/repair. One population was very closely associated with the previously identified pericyte cell population with sequence analysis suggesting it might represent a subgroup of cells that had some minor changes in function possibly associated with loss of some of the vascular compartment. A second population was only present at 48h and based on expression of several genes associated with apoptosis we concluded that it represented stromal cells that were not going to survive in repairing tissue and that cell death was one potential fate of the GFP+ cells. The most interesting population of GFP+ cells were unique to the 24h tissue with a unique transcriptome that included expression of both canonical mesenchymal and epithelial cell markers as well as a number of genes such as Ctnnb1 (beta-catenin) and Twist which have been implicated cell transitions between mesenchyme and epithelium. Notably it has previously been reported that mice with stabilisation of *Ctnnb1*, a gene that plays an important role in Wnt signalling, in the mesenchyme of the reproductive tract had smaller uteri, endometrial gland defects, and were infertile (Stewart et al., 2013). Further analysis using flow cytometry and immunhistochemistry confirmed the existence of GFP+EPCAM+ cells in the 24h samples with numbers declining at 48h.

These results backed up our previous observation that MET occurred in regions of tissue associated with endometrial repair and are in agreement with studies exploring the fate of mesenchyme cells after parturition (Huang et al., 2012, Patterson et al., 2013). Huang et al used a knockin strategy with cells expressing LacZ induced by *Amhr2-Cre*: in virgin females the reporter was specific to stromal cells but post-partum they found labelled cells in both luminal and glandular epithelium. In a third study using a SM22a-Cre driver expression of which was reported to be restricted to a subset of CD34+ mesenchymal cells reporter gene expression (GFP) was also detected in the epithelium (Yin et al., 2019). Whilst authors of all these papers concluded their results were consistent with stromal MET this interpretation has been questioned (Ghosh et al., 2020). Specifically, Ghosh et al conducted a comprehensive examination of embryonic and adult reproductive tracts using LacZ reporter lines driven by promoters for the genes *Amhr2, Sm22, Cspg4, Thy1* and *Pdgfrß* all of which are expressed in mesenchyme to explore whether epithelial cells expressing reporter proteins detected in adult endometrium arose from MET or were induced at a time when cells had meso-epithelial characteristics. In all cases they attributed epithelial expression of the reporter protein in adulthood to activation of the promoters during embryonic life ruling out MET in adult cycling mice (Ghosh et al., 2020).

The limitation of our study using *Pdgfrb*-BAC-eGFP reporter mice was that the knockin strategy used to generate the line meant GFP was only expressed in cells when the *Pgdfrb* promoter was active. In our previous study we had already established that the expression of GFP in these mice was identical to the native PDGFRB protein and was stroma-specific (Kirkwood et al., 2021). The comparison between the original data and those obtained from uteri recovered after induction of endometrial shedding/repair led us to conclude that the presence of GFP in the epithelial cells was most likely because the turnover of the GFP was slow enough for it to persist after the *Pgdfrb* promoter was no longer active. This would be consistent with reports that eGFP protein was engineered to have increased fluorescence and slower turnover *in vivo* (Zhao et al., 1999) and although the half-life of the protein varies in different cell models a paper exploring *in vivo* expression in cancer reported it to be ~15h (Danhier et al., 2015).

Therefore, to make a more robust case for MET of endometrial stromal cells we expanded our scRNAseq datasets by analysing the epithelial cell compartment (EPCAM+) of tissue recovered from cycling mice and those in which endometrial repair was well advanced (48h) and complemented this by undertaking lineage tracing studies using a strategy that induced expression of a tdTm reporter protein in adult mice thereby avoiding any confusion of results with MET that might occur during development or early life. Analysis of EPCAM+ cells resulted in identification of 9 subpopulations of epithelial cells based on cluster analysis and comparisons to genes previously identified as being expressed in different locations including the lumen (Yang et al., 2021) and glands (*Foxa2*, (Kelleher et al., 2017). When we combined our eGFP+ and EPCAM+ scRNAseq datasets three broad groups of cell types was present: epithelial, perivascular, fibroblast and epithelial with the repair-specific transient population of cells appearing to be a distinct fibroblast subpopulation which on trajectory analysis appeared to form a 'bridge' between the previously identified fibroblast clusters (Kirkwood et al., 2021) and the new epithelial datasets. Based on previous reports of the importance of endometrial progenitor populations in endometrial repair and identification of cells with stem-like properties in a perivascular location (Gargett et al., 2016) we had initially anticipated this cell type might be the one that was mobilised to undergo MET however the new data suggested this assumption was incorrect. To ensure we gained a full picture of the source of the putative MET cells we targeted fibroblasts using a Cre under the control of *Pdgfra* and perivascular cells with one under the control of the *Cspg4* (NG2) promoters respectively using a TAM induction schedule that we refined so that it had no impact on the response of the tissue to decidualisation and induction of endometrial injury. The results obtained were unequivocal and supported not only MET of fibroblasts but also their incorporation into the restored luminal epithelium in a layer that also included epithelial cells which had not arisen by MET but were likely to have migrated from residual epithelial cell remnants left at time of shedding. These new findings provide the evidence that the endometrium can use multiple mechanisms including recruitment of mesenchyme cells to rapidly restore its integrity a feature of the tissue that is important for its normal functioning and for fertility.

The generation of these scRNAseq datasets have allowed us to make some preliminary comparisons to results recently reported studies using this technique to explore cell heterogeneity in human endometrium and human endometrial cells (Wang et al., 2020, Queckborner et al., 2021, Cao et al., 2021, Garcia-Alonso et al., 2021). Queckborner and colleagues focused their analysis cells from 3 donors during in the proliferative phase of their menstrual cycle (6864 cells) using a protocol that enriched for stromal cells (Queckborner et al., 2021). As in our data they subtyped two populations of perivascular/pericyte cells with different expression levels of MY11D and CSPG4 both of which we have previously mapped to perivascular populations in the mouse endometrium (Kirkwood et al., 2021). They subtyped the stromal fibroblasts into 10 closely related clusters of fibroblasts and concluded that the markers for progenitor cells were not specific enough. Wang et al incorporated tissues from across the menstrual cycle in their analysis including 6 that were assigned to the menstrual phase according to their supplementary data these were identified as days 4-11 of the cycle which might be after the most rapid phase of repair (Wang et al., 2020). They reported a previously uncharacterised ciliated epithelial subtype and changes in luminal epithelium and stroma that were discrete to the window of implantation but nothing specific to repair mechanisms. Cao et al focused their attention on the perivascular cell population performing sequence analysis of CD140b+CD146+ (eMSC) cells isolated from menstrual phase (days 2-3) and secretory phases (n=3 women in each group (Cao et al., 2021). They identified 2 subclusters in cultured cells from the menstrual phase with the larger cluster having high expression of PDGFRB/PDGFRA consistent with a fibroblast identity. When they compared their data to the Wang dataset described above they noted significant changes in gene

expression patterns and concluded that the cell transcriptomes had been influenced by cell culture which may compromise attempts to compare between in vitro and in vivo data. The most recent study on human endometrial tissues reporting data from scRNAseq is that by Garcia-Alonso and colleagues published in December 2021. This very large study generated scRNA data on 98.5K cells from 15 individuals charcterised as being in the proliferative and secretory phases and identified 14 cell clusters which was complemented by analysis of cells in tissue organoids (Garcia-Alonso et al., 2021). The major focus of this work was on the epithelial cell compartment as these cells are also represented in the organoids with novel findings related to SOX9 expression and WNT/NOTCH signalling but no comparable data to ours on changes in stromal cell identity at time of menstruation.

In summary, we describe novel data and new datasets that show MET can contribute to the rapid scar-free healing of the endometrial luminal epithelium when this occurs in response to an endometrial 'wound' at the time of menstruation. The cell type involved in this MET is a PDGFRA+ stromal cell and not the NG2+ pericytes. These data provide a platform for further analysis of the fibrosis-resistant phenotype of endometrial fibroblasts, comparisons to conditions where their function is abnormal (Asherman's syndrome, endometriosis) and comparison to other mucosal barrier tissues such as that found in the oral cavity as recently described by Williams et al (Williams et al., 2021). We postulate that manipulation of putative MET progenitors/specific cell types in the disease setting may provide novel strategies in the management and/or treatment of disorders associated with poor response to healing or excess fibrosis.

Acknowledgements.

PMK was supported by a MRC Transition Fellowship funded by a Doctoral Training Award to the University of Edinburgh (MR/N013166/1); DAG, IS and OK were funded by a MRC Programme Grant to PTKS (MR/N024524/1); NCH was supported by a Wellcome Trust Senior Research Fellowship in Clinical Science (219542/Z/19/Z). We thank staff in core facilities including flow cytometry and animal husbandry for their support and guidance, Arantza Esnal-Z for help with immunohistochemistry and John R Wilson-Kanamori for discussion about bioinformatic methods. We are most grateful to Professor Brigid Hogan (Duke University) for her generosity in providing founder stocks of the . i*Pdgfra*Cre-ER^{T2} mice.

Author contributions.

PMK, NCH, PTKS and DAG designed research; PMK, DAG, OK, IS performed research; PMK, DAG, JW-K analysed data; PMK, PTKS wrote the first draft; DAG, IS, RD, NCH revised the paper, PMK, NH and PTKS funded the research.

Competing interest statement.

The authors have no competing interests to report.

REFERENCES

AGHAJANOVA, L., HORCAJADAS, J. A., ESTEBAN, F. J. & GIUDICE, L. C. 2010. The bone marrow-derived human mesenchymal stem cell: potential progenitor of the endometrial stromal fibroblast. *Biol Reprod,* 82**,** 1076-87.

ARMSTRONG, G. M., MAYBIN, J. A., MURRAY, A. A., NICOL, M., WALKER, C., SAUNDERS, P. T. K., ROSSI, A. G. & CRITCHLEY, H. O. D. 2017. Endometrial apoptosis and neutrophil infiltration during menstruation exhibits spatial and temporal dynamics that are recapitulated in a mouse model. *Sci Rep,* 7**,** 17416.

BAGGISH, M. S., PAUERSTEIN, C. J. & WOODRUFF, J. D. 1967. Role of stroma in regeneration of endometrial epithelium. *Am J Obstet Gynecol,* 99**,** 459-65.

BRENNER, R. M., NAYAK, N. R., SLAYDEN, O. D., CRITCHLEY, H. O. & KELLY, R. W. 2002. Premenstrual and menstrual changes in the macaque and human endometrium: relevance to endometriosis. *Ann N Y Acad Sci,* 955**,** 60-74; discussion 86-8, 396-406.

CAO, D., CHAN, R. W. S., NG, E. H. Y., GEMZELL-DANIELSSON, K. & YEUNG, W. S. B. 2021. Single-cell RNA sequencing of cultured human endometrial CD140b(+)CD146(+) perivascular cells highlights the importance of in vivo microenvironment. *Stem Cell Res Ther,* 12**,** 306.

CHUNG, M. I., BUJNIS, M., BARKAUSKAS, C. E., KOBAYASHI, Y. & HOGAN, B. L. M. 2018. Niche-mediated BMP/SMAD signaling regulates lung alveolar stem cell proliferation and differentiation. *Development,* 145.

COUSINS, F. L., KIRKWOOD, P. M., MURRAY, A. A., COLLINS, F., GIBSON, D. A. & SAUNDERS, P. T. 2016a. Androgens regulate scarless repair of the endometrial "wound" in a mouse model of menstruation. *FASEB J,* 30**,** 2802-11.

COUSINS, F. L., KIRKWOOD, P. M., SAUNDERS, P. T. & GIBSON, D. A. 2016b. Evidence for a dynamic role for mononuclear phagocytes during endometrial repair and remodelling. *Sci Rep,* 6**,** 36748.

COUSINS, F. L., MURRAY, A., ESNAL, A., GIBSON, D. A., CRITCHLEY, H. O. & SAUNDERS, P. T. 2014. Evidence from a mouse model that epithelial cell migration and mesenchymal-epithelial transition contribute to rapid restoration of uterine tissue integrity during menstruation. *PLoS One,* 9**,** e86378.

COUSINS, F. L., PANDOY, R., JIN, S. & GARGETT, C. E. 2021. The Elusive Endometrial Epithelial Stem/Progenitor Cells. *Front Cell Dev Biol,* 9**,** 640319.

CRITCHLEY, H. O., KELLY, R. W., BAIRD, D. T. & BRENNER, R. M. 2006a. Regulation of human endometrial function: mechanisms relevant to uterine bleeding. *Reprod Biol Endocrinol,* 4 Suppl 1**,** S5.

CRITCHLEY, H. O., OSEI, J., HENDERSON, T. A., BOSWELL, L., SALES, K. J., JABBOUR, H. N. & HIRANI, N. 2006b. Hypoxia-inducible factor-1alpha expression in human endometrium and its regulation by prostaglandin E-series prostanoid receptor 2 (EP2). *Endocrinology,* 147**,** 744-53.

DANHIER, P., KRISHNAMACHARY, B., BHARTI, S., KAKKAD, S., MIRONCHIK, Y. & BHUJWALLA, Z. M. 2015. Combining Optical Reporter Proteins with Different Half-lives to Detect Temporal Evolution of Hypoxia and Reoxygenation in Tumors. *Neoplasia,* 17**,** 871-881.

DEANE, J. A., ONG, Y., COUSINS, F. L. & GARGETT, C. E. 2019. Bone marrow-derived endometrial cells: transdifferentiation or misidentification? *Hum Reprod Update,* 25**,** 272-274.

EVANS, J. & SALAMONSEN, L. A. 2012. Inflammation, leukocytes and menstruation. *Rev Endocr Metab Disord,* 13**,** 277-88.

FERENCZY, A. 1976a. Studies on the cytodynamics of human endometrial regeneration. I. Scanning electron microscopy. *Am J Obstet Gynecol,* 124**,** 64-74.

FERENCZY, A. 1976b. Studies on the cytodynamics of human endometrial regeneration. II. Transmission electron microscopy and histochemistry. *Am J Obstet Gynecol,* 124**,** 582-95.

GARCIA-ALONSO, L., HANDFIELD, L. F., ROBERTS, K., NIKOLAKOPOULOU, K., FERNANDO, R. C., GARDNER, L., WOODHAMS, B., ARUTYUNYAN, A., POLANSKI, K., HOO, R., SANCHO-SERRA, C., LI, T., KWAKWA, K., TUCK, E., LORENZI, V., MASSALHA, H., PRETE, M., KLESHCHEVNIKOV, V., TARKOWSKA, A., PORTER, T., MAZZEO, C. I., VAN DONGEN, S., DABROWSKA, M., VASKIVSKYI, V., MAHBUBANI, K. T., PARK, J. E., JIMENEZ-LINAN, M., CAMPOS, L., KISELEV, V. Y., LINDSKOG, C., AYUK, P., PRIGMORE, E., STRATTON, M. R., SAEB-PARSY, K., MOFFETT, A., MOORE, L., BAYRAKTAR, O. A., TEICHMANN, S. A., TURCO, M. Y. & VENTO-TORMO, R. 2021. Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. *Nat Genet,* 53**,** 1698-1711.

GARGETT, C. E., SCHWAB, K. E. & DEANE, J. A. 2016. Endometrial stem/progenitor cells: the first 10 years. *Hum Reprod Update,* 22**,** 137-63.

GARRY, R., HART, R., KARTHIGASU, K. A. & BURKE, C. 2009. A re-appraisal of the morphological changes within the endometrium during menstruation: a hysteroscopic, histological and scanning electron microscopic study. *Hum Reprod,* 24**,** 1393-401.

GHOSH, A., SYED, S. M., KUMAR, M., CARPENTER, T. J., TEIXEIRA, J. M., HOUAIRIA, N., NEGI, S. & TANWAR, P. S. 2020. In Vivo Cell Fate Tracing Provides No Evidence for Mesenchymal to Epithelial Transition in Adult Fallopian Tube and Uterus. *Cell Rep,* 31**,** 107631.

GIELEN, S. C., SANTEGOETS, L. A., HANIFI-MOGHADDAM, P., BURGER, C. W. & BLOK, L. J. 2008. Signaling by estrogens and tamoxifen in the human endometrium. *J Steroid Biochem Mol Biol,* 109**,** 219-23.

GREENHALGH, S. N., CONROY, K. P. & HENDERSON, N. C. 2015. Healing scars: targeting pericytes to treat fibrosis. *QJM,* 108**,** 3-7.

HUANG, C. C., ORVIS, G. D., WANG, Y. & BEHRINGER, R. R. 2012. Stromal-to-epithelial transition during postpartum endometrial regeneration. *PLoS One,* 7**,** e44285.

KAITU'U-LINO, T. J., YE, L. & GARGETT, C. E. 2010. Reepithelialization of the uterine surface arises from endometrial glands: evidence from a functional mouse model of breakdown and repair. *Endocrinology,* 151**,** 3386-95.

KAITU'U-LINO, T. J., YE, L., SALAMONSEN, L. A., GIRLING, J. E. & GARGETT, C. E. 2012. Identification of label-retaining perivascular cells in a mouse model of endometrial decidualization, breakdown, and repair. *Biol Reprod,* 86**,** 184.

KELLEHER, A. M., PENG, W., PRU, J. K., PRU, C. A., DEMAYO, F. J. & SPENCER, T. E. 2017. Forkhead box a2 (FOXA2) is essential for uterine function and fertility. *Proc Natl Acad Sci U S A,* 114**,** E1018-E1026.

KIRKWOOD, P. M., GIBSON, D. A., SMITH, J. R., WILSON-KANAMORI, J. R., KELEPOURI, O., ESNAL-ZUFIAURRE, A., DOBIE, R., HENDERSON, N. C. & SAUNDERS, P. T. K. 2021. Single-cell RNA sequencing redefines the mesenchymal cell landscape of mouse endometrium. *FASEB J,* 35**,** e21285.

LUDWIG, H. & SPORNITZ, U. M. 1991. Microarchitecture of the human endometrium by scanning electron microscopy: menstrual desquamation and remodeling. *Ann N Y Acad Sci,* 622**,** 28-46.

MAYBIN, J. A. & CRITCHLEY, H. O. 2011. Progesterone: a pivotal hormone at menstruation. *Ann N Y Acad Sci,* 1221**,** 88-97.

MAYBIN, J. A. & CRITCHLEY, H. O. 2012. Steroid regulation of menstrual bleeding and endometrial repair. *Rev Endocr Metab Disord,* 13**,** 253-63.

MAYBIN, J. A., MURRAY, A. A., SAUNDERS, P. T. K., HIRANI, N., CARMELIET, P. & CRITCHLEY, H. O. D. 2018. Hypoxia and hypoxia inducible factor-1alpha are required for normal endometrial repair during menstruation. *Nat Commun,* 9**,** 295.

PATTERSON, A. L., ZHANG, L., ARANGO, N. A., TEIXEIRA, J. & PRU, J. K. 2013. Mesenchymal-to-Epithelial Transition Contributes to Endometrial Regeneration Following Natural and Artificial Decidualization. *Stem Cells Dev,* 22**,** 964-74.

QIU, X., MAO, Q., TANG, Y., WANG, L., CHAWLA, R., PLINER, H. A. & TRAPNELL, C. 2017. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods,* 14**,** 979-982.

QUECKBORNER, S., VON GROTHUSEN, C., BOGGAVARAPU, N. R., FRANCIS, R. M., DAVIES, L. C. & GEMZELL-DANIELSSON, K. 2021. Stromal Heterogeneity in the Human Proliferative Endometrium-A Single-Cell RNA Sequencing Study. *J Pers Med,* 11.

STEWART, C. A., WANG, Y., BONILLA-CLAUDIO, M., MARTIN, J. F., GONZALEZ, G., TAKETO, M. M. & BEHRINGER, R. R. 2013. CTNNB1 in mesenchyme regulates epithelial cell differentiation during Mullerian duct and postnatal uterine development. *Mol Endocrinol,* 27**,** 1442-54.

STUART, T. & SATIJA, R. 2019. Integrative single-cell analysis. *Nat Rev Genet,* 20**,** 257-272.

SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R., LANDER, E. S. & MESIROV, J. P. 2005. Gene set

enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A,* 102**,** 15545-50.

WANG, W., VILELLA, F., ALAMA, P., MORENO, I., MIGNARDI, M., ISAKOVA, A., PAN, W., SIMON, C. & QUAKE, S. R. 2020. Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle. *Nat Med,* 26**,** 1644-1653.

WILLIAMS, D. W., GREENWELL-WILD, T., BRENCHLEY, L., DUTZAN, N., OVERMILLER, A., SAWAYA, A. P., WEBB, S., MARTIN, D., GENOMICS, N. N., COMPUTATIONAL BIOLOGY, C., HAJISHENGALLIS, G., DIVARIS, K., MORASSO, M., HANIFFA, M. & MOUTSOPOULOS, N. M. 2021. Human oral mucosa cell atlas reveals a stromal-neutrophil axis regulating tissue immunity. *Cell,* 184**,** 4090-4104 e15.

WOLOCK, S. L., LOPEZ, R. & KLEIN, A. M. 2019. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst,* 8**,** 281-291 e9.

YANG, Y., ZHU, Q. Y. & LIU, J. L. 2021. Deciphering mouse uterine receptivity for embryo implantation at single-cell resolution. *Cell Prolif,* 54**,** e13128.

YIN, M., ZHOU, H. J., LIN, C., LONG, L., YANG, X., ZHANG, H., TAYLOR, H. & MIN, W. 2019. CD34(+)KLF4(+) Stromal Stem Cells Contribute to Endometrial Regeneration and Repair. *Cell Rep,* 27**,** 2709-2724 e3.

YU, G., WANG, L. G., HAN, Y. & HE, Q. Y. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS,* 16**,** 284-7.

ZHAO, X., JIANG, X., HUANG, C. C., KAIN, S. R. & LI, X. 1999. Generation of a destablilized form of enhanced green fluorescent protein. *Methods Enzymol,* 302**,** 438-44.

ZHU, X., HILL, R. A., DIETRICH, D., KOMITOVA, M., SUZUKI, R. & NISHIYAMA, A. 2011. Age-dependent fate and lineage restriction of single NG2 cells. *Development,* 138**,** 745-53.

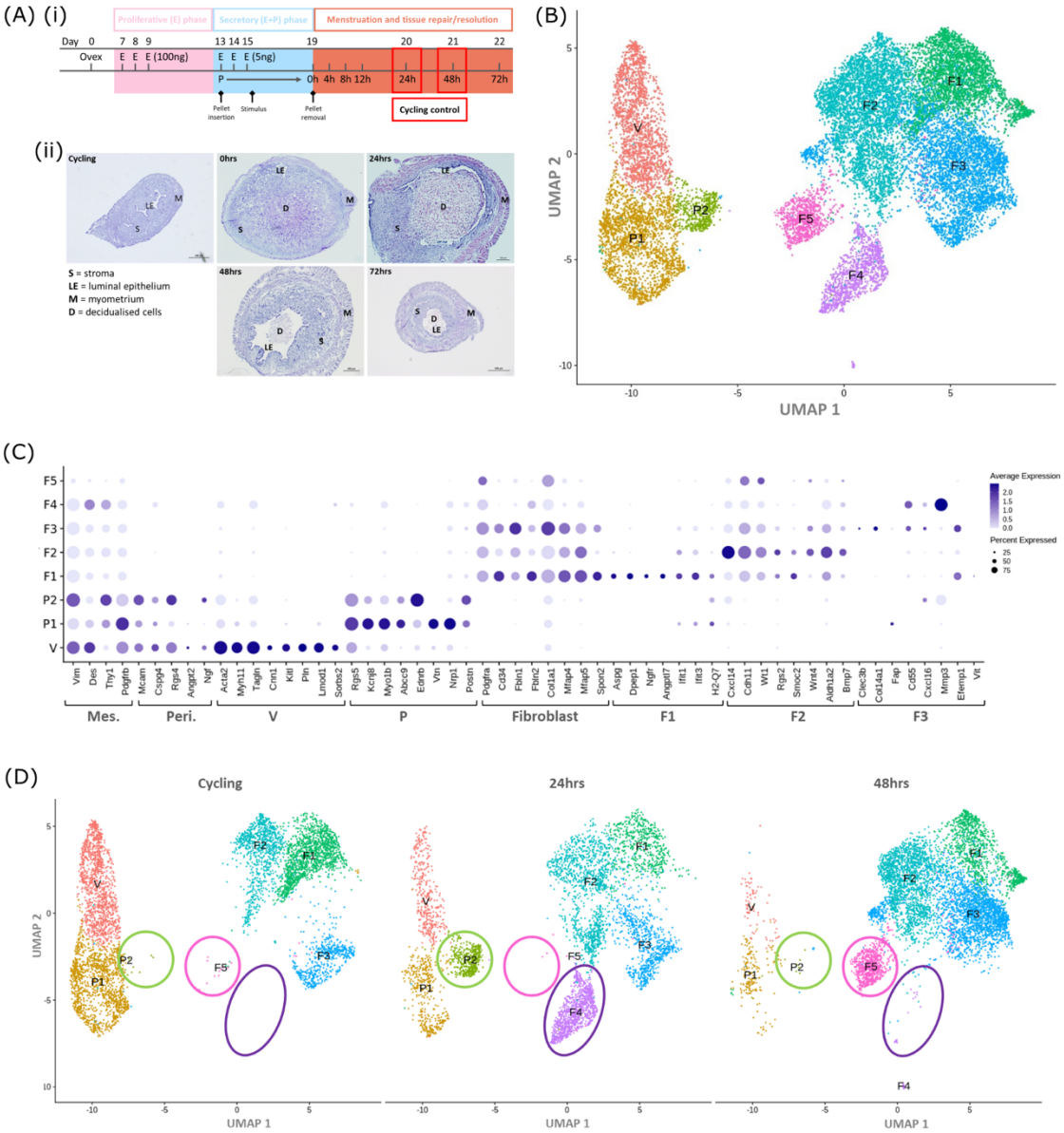

Figure 1. Single cell RNA sequence analysis identified mesenchymal cell populations unique to 'repairing' tissue in a mouse model of endometrial breakdown and repair (menstruation).

(A)(i) *mouse model of endometrial tissue breakdown and repair,* ***(ii)*** *histological morphology of uterine tissues 0, 24, 48 and 72h progesterone withdrawal illustrating tissue breakdown, repair, remodelling and resolution.* ***(B)*** *UMAP visualisation: GFP+ mesenchymal cells isolated from Pdgfrb-BAC-eGFP mouse endometrium (cycling plus 24/48h after progesterone withdrawal) cluster into 8 distinct populations.* ***(C)*** *Dot plot: expression of canonical gene signatures associated with known cell types present in the endometrium: mesenchymal cells (Pdgfrb, Vim, Des, Thy1), perivascular cells (Mcam, Cspg4, Rgs4, Acta2) and fibroblasts (Pdgfra, Cd34, Fbln1/2, Mfap4/5, Col1a1) (dot colour: average expression per cluster; dot size: percent cluster expressing gene)* ***(D)*** *UMAP visualisation: mesenchymal clusters split by source/dataset identifies transient repair-specific subpopulations P2 and F4 in 24h tissues and F5 in 48h tissues.*

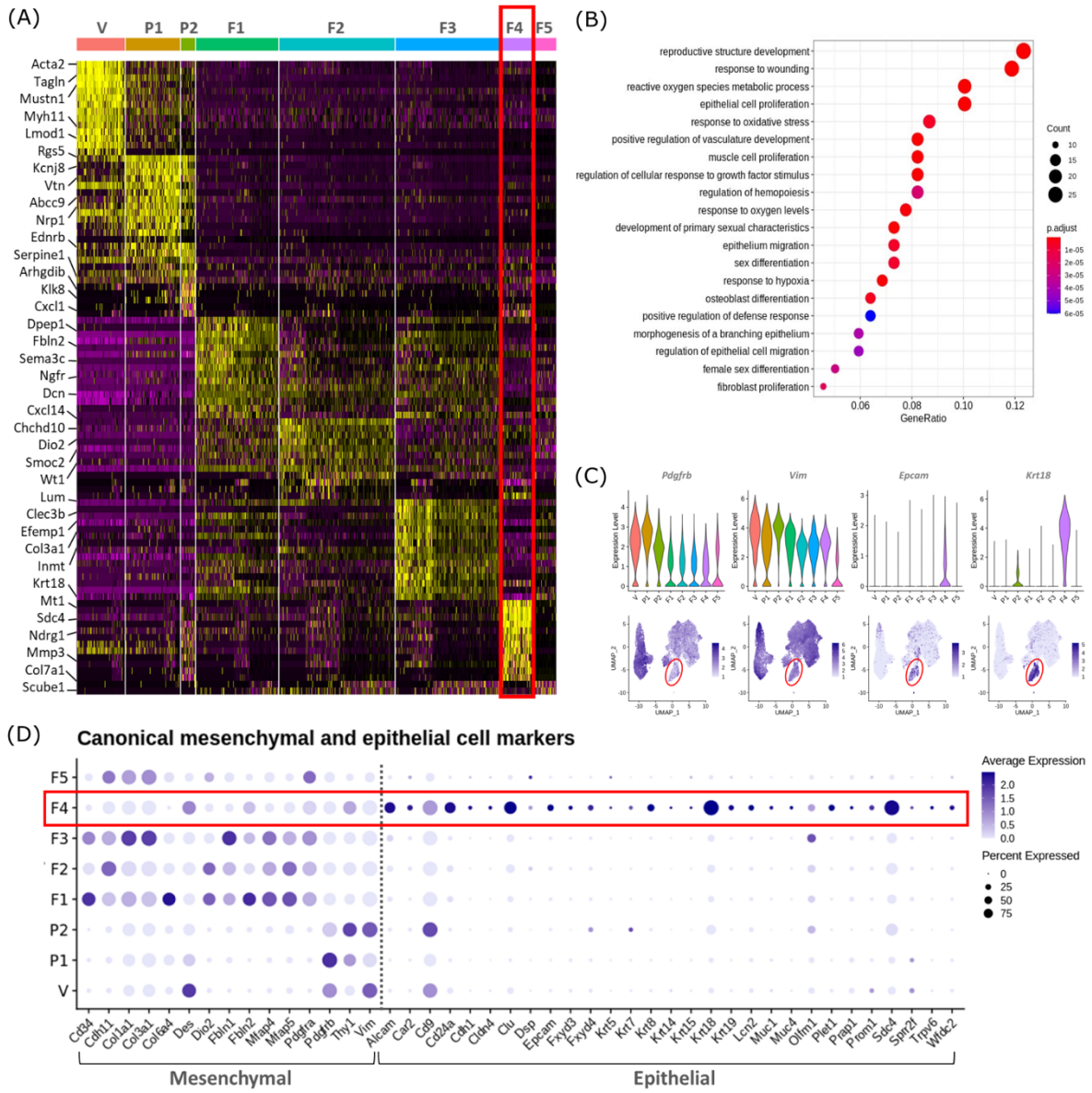
(A)
V
P1
P2
F1
F2
F3
F4
F5
Acta2
Tagln
Mustn1
Myh11
Lmod1
Rgs5
Kcnj8
Vtn
Abcc9
Nrp1
Ednrb
Serpine1
Arhgdib
Klk8
Cxcl1
Dpep1
Fbln2
Sema3c
Ngfr
Dcn
Cxcl14
Chchd10
Dio2
Smoc2
Wt1
Lum
Clec3b
Efemp1
Col3a1
Inmt
Krt18
Mt1
Sdc4
Ndrg1
Mmp3
Col7a1
Scube1
(B)
reproductive structure development
response to wounding
reactive oxygen species metabolic process
epithelial cell proliferation
response to oxidative stress
positive regulation of vasculature development
muscle cell proliferation
regulation of cellular response to growth factor stimulus
regulation of hemopoiesis
response to oxygen levels
development of primary sexual characteristics
epithelium migration
sex differentiation
response to hypoxia
osteoblast differentiation
positive regulation of defense response
morphogenesis of a branching epithelium
regulation of epithelial cell migration
female sex differentiation
fibroblast proliferation
0.06
0.08
0.10
0.12
GeneRatio
Count
10
15
20
25
p.adjust
1e-05
2e-05
3e-05
4e-05
5e-05
6e-05
(C)
Pdgfrb
Vim
Epcam
Krt18
Expression Level
UMAP_1
UMAP_2
(D)
Canonical mesenchymal and epithelial cell markers
F5
F4
F3
F2
F1
P2
P1
V
Cd34
Cdh11
Col1a1
Col3a1
Col6a4
Des
Dio2
Fbln1
Fbln2
Mfap4
Mfap5
Pdgfra
Pdgfrb
Thy1
Vim
Alcam
Car2
Cd9
Cd24a
Cdh1
Cldn4
Clu
Dsp
Epcam
Fxyd3
Fxyd4
Krt5
Krt7
Krt8
Krt14
Krt15
Krt18
Krt19
Lcn2
Muc1
Muc4
Olfm1
Plet1
Prap1
Prom1
Sdc4
Spnr2l
Trpv6
Wfdc2
Mesenchymal
Epithelial
Average Expression
2.0
1.5
1.0
0.5
0.0
Percent Expressed
0
25
50
75
(E)
MSigDB GSEA; genes associated with MET/EMT
F5
F4
F3
F2
F1
P2
P1
V
Bmp4
Bmp7
Cemip
Cited1
Ctnnb1
Ehf
Elf3
Fzd7
Gata3
Grem1
Hmga2
Irf6
Klf4
Pax2
Pax8
Smo
Twist2
Wnt4
Wt1
Notch1
Snai2
Sox9
Twist1
MET
EMT
Percent Expressed
0
25
50
75
Average Expression
2.0
1.5
1.0
0.5
0.0

Figure 2: Gene expression analysis identifies transcriptomic profiles of mesenchymal cells and the expression of genes associated with epithelial cell identity in repair-specific fibroblasts (cluster F4).

(A) *Heatmap (yellow, high; purple, low) displaying differentially expressed genes per cluster when compared to all other clusters (logFC >0.5, pvalue <0.05, Wilcoxon rank-sum test) top is colour coded and named by cluster; V= vascular smooth muscle cells (vSMCs), P1= pericytes 1, P2= pericytes 2, F1= fibroblasts 1, F2= fibroblasts 2, F3= fibroblasts 3, F4= fibroblasts 4, F5= fibroblasts 5). The expression of the top 5 exemplar genes in each cell cluster is illustrated.* ***(B)*** *Dot plot: GO enrichment terms relating to biological processes (BP) associated with the genetic signature of repair-specific cluster F4 (dot size: gene ratio, number of genes in data/number of genes associated with GO term; dot colour: p-value representing the enrichment score).* ***(C)*** *Gene expression plots: expression of canonical mesenchymal cell markers* ***(i)*** *Pdgfrb and* ***(ii)*** *Vim and canonical epithelial cell markers* ***(iii)*** *Epcam and* ***(iv)*** *Krt18. Note F4 fibroblasts (red box) express markers associated with both cell lineages.* ***(D)*** *Dot plot: expression of extended gene signatures associated with mesenchymal and epithelial cell lineages in the endometrium (dot colour: average expression per cluster; dot size: percent cluster expressing gene).* ***(E)*** *Dot plot: expression of genes associated with the regulation of MET/EMT taken from the MSigDB (dot colour: average expression per cluster; dot size: percent cluster expressing gene).*

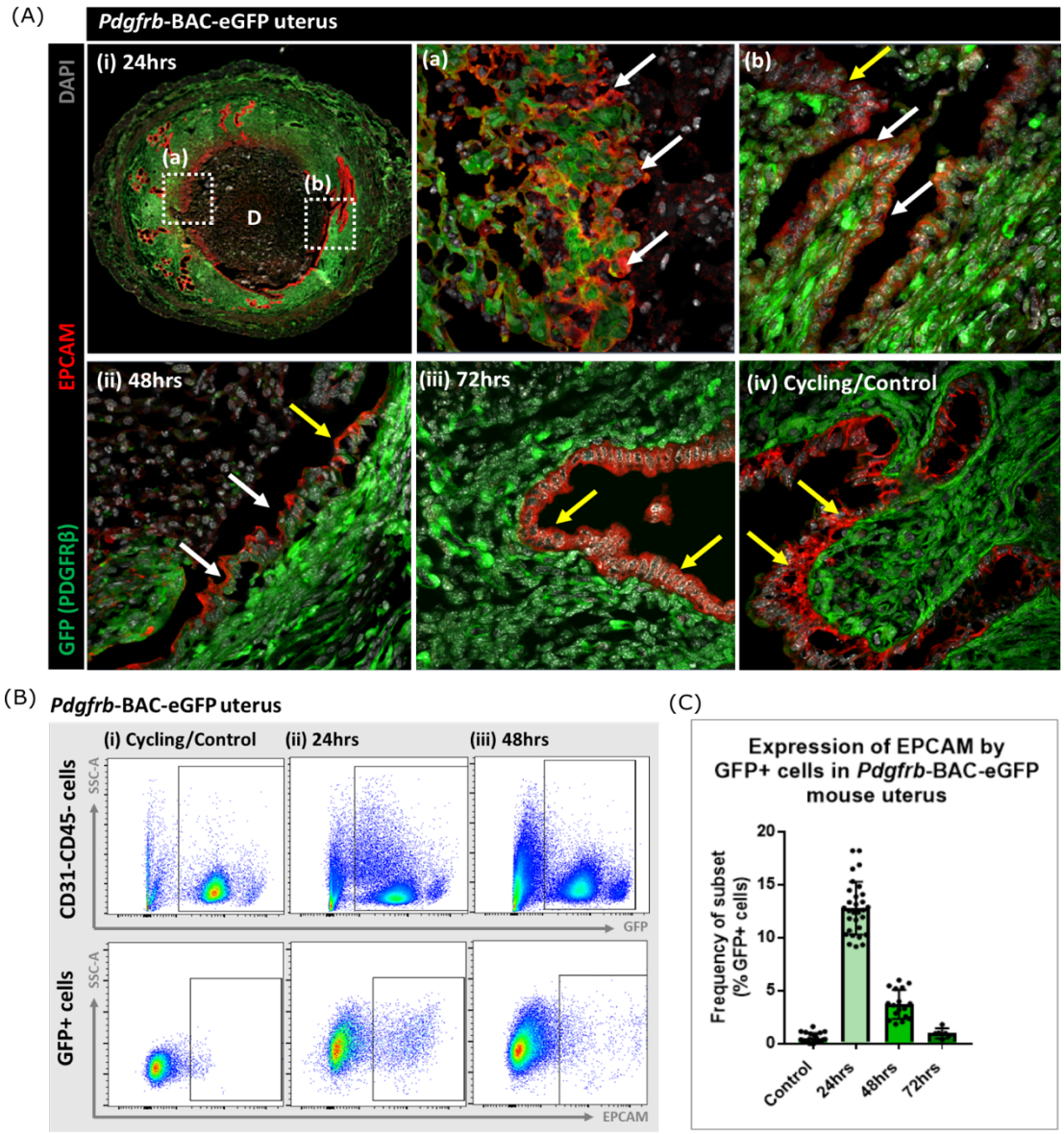

Figure 3. Analysis of Pdgfrb-BAC-eGFP uterine tissue during endometrial tissue repair and remodelling identifies transient expression of GFP in EPCAM+ epithelial cells.

(A) *Immunohistochemical analysis of GFP reporter protein (green) and epithelial cell marker EPCAM (red) in uterine tissues 24, 48 and 72h following progesterone withdrawal.* ***(i)*** *In 24h tissues cells that are GFP+EPCAM+ are detected in (a) regions of denuded stromal surfaces where the decidualised tissue has detached (white arrows) and (b) regions where a new epithelium has formed (white arrows).* ***(ii)*** *At 48h GFP+EPCAM+ cells were detected within the renewed luminal epithelium (white arrows) adjacent to GFP- EPCAM+ epithelial cells (yellow arrows).* ***(iii)*** *At 72h no GFP+ EPCAM+ cells were detected.* ***(iv)*** *In control/cycling tissues epithelial cells are EPCAM+ (yellow arrows), stroma is GFP+.* ***(B)*** *Flow cytometry (FC): analysis of EPCAM in Pdgfrb-BAC-eGFP uterine tissues at 24, 48 and 72h following progesterone withdrawal; note detection of a new population of GFP+EPCAM+ cells at 24h (boxed lower panel) with decreased numbers at 48h.* ***(C)*** *Bar plot: quantification of FC data analysing the expression of EPCAM by GFP reporter positive cells in the uterus (%) - control (n=21) 0.53±0.46%; 24h (n=30) 12.78±2.49%; 48h (n=17) 3.71±1.36%; 72h (n=6) 0.96±0.47%).*

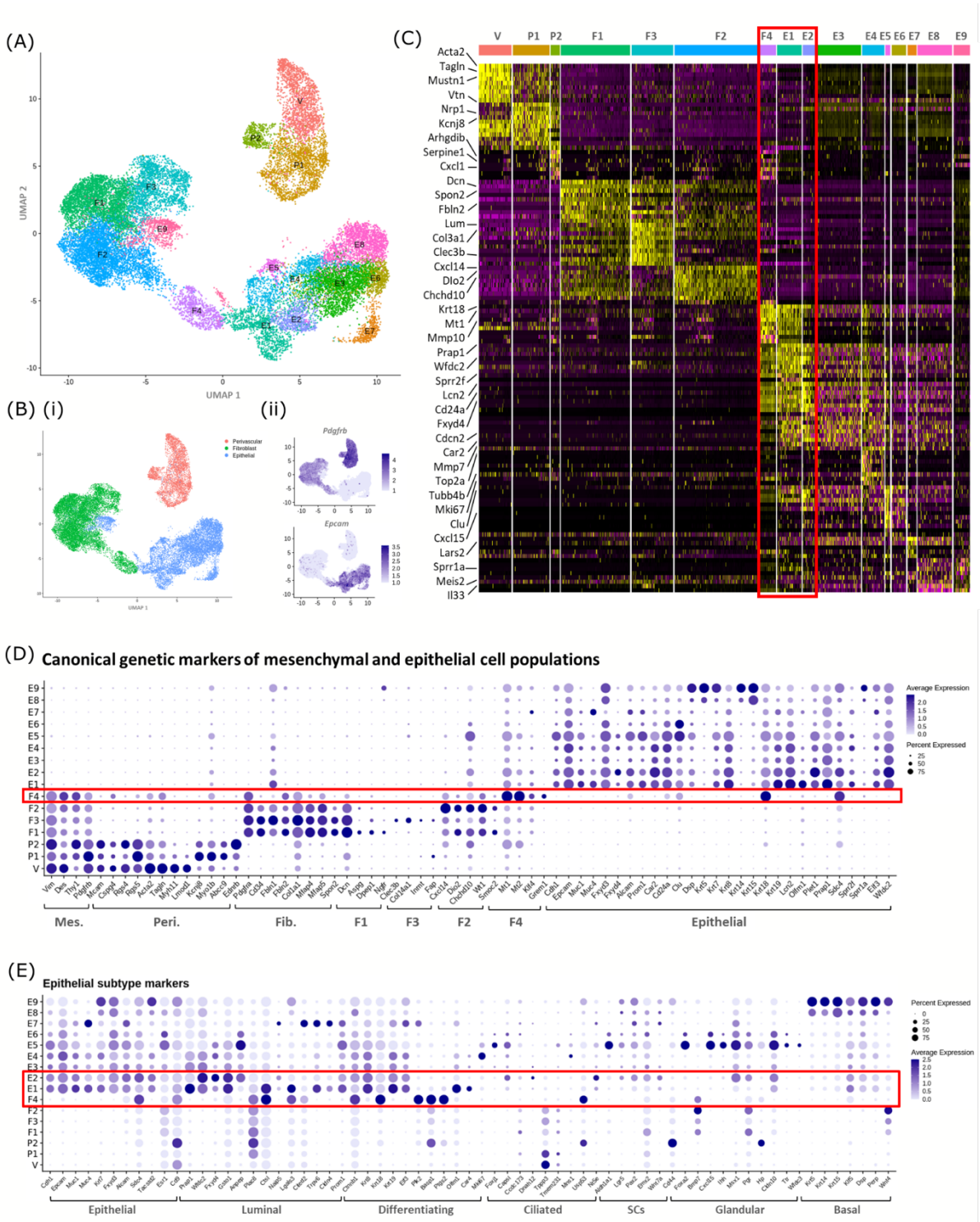

Figure 4: In silico analysis of scRNAseq reveals transcriptomic similarity between repair-specific mesenchymal cells and subpopulations of EPCAM+ endometrial epithelial cells.

(A) *UMAP visualisation: GFP+ mesenchymal cells and GFP-EPCAM+ epithelial cells isolated from Pdgfrb-BAC-eGFP mouse endometrium (cycling/24/48h and cycling/48h respectively) cluster into 16 distinct populations: 7 mesenchymal, 9 epithelial. Repair-specific*

fibroblasts (F4) cluster in close proximity to epithelial cell clusters but remain a distinct population of cells. ***(B) (i)*** *UMAP visualisation: 3 groups of cell populations are detected in the data: perivascular, fibroblast and epithelial cells (add colours)* ***(ii)*** *gene expression of Pdgfrb is restricted to mesenchymal populations (perivascular/fibroblast) while Epcam is restricted to epithelial populations (and repair-specific fibroblasts; F4), confirming the isolation strategy adopted.* ***(C)*** *Scaled heatmap (yellow, high; purple, low) displaying differentially expressed genes per cluster when compared to all other clusters (logFC >0.5, pvalue <0.05, Wilcoxon rank-sum test) top is colour coded and named by cluster; V= vascular smooth muscle cells (vSMCs), P1/2= pericytes 1/2, F1-4= fibroblasts 1-4, E1-9= epithelial cells 1-9). The expression of 3 exemplar genes in each cell cluster is displayed. F4 shows genetic similarity to E1/2 (red box).* ***(D)*** *Dot plot: expression of canonical genes associated with mesenchymal, perivascular, fibroblast and epithelial lineages, F4 is the only cluster that expressed genes from multiple lineages (red box) (dot colour: average expression per cluster; dot size: percent cluster expressing gene).* ***(E)*** *Dot plot: expression of gene signatures associated with known epithelial cell subtypes present in the endometrium as per the literature: canonical epithelial, luminal, differentiating, ciliated, SCs, glandular, basal epithelial cells (dot colour: average expression per cluster; dot size: percent cluster expressing gene). Red box F4 compared to E1, E2] highlighting expression of genes found in the luminal and differentiating epithelial cell populations.*

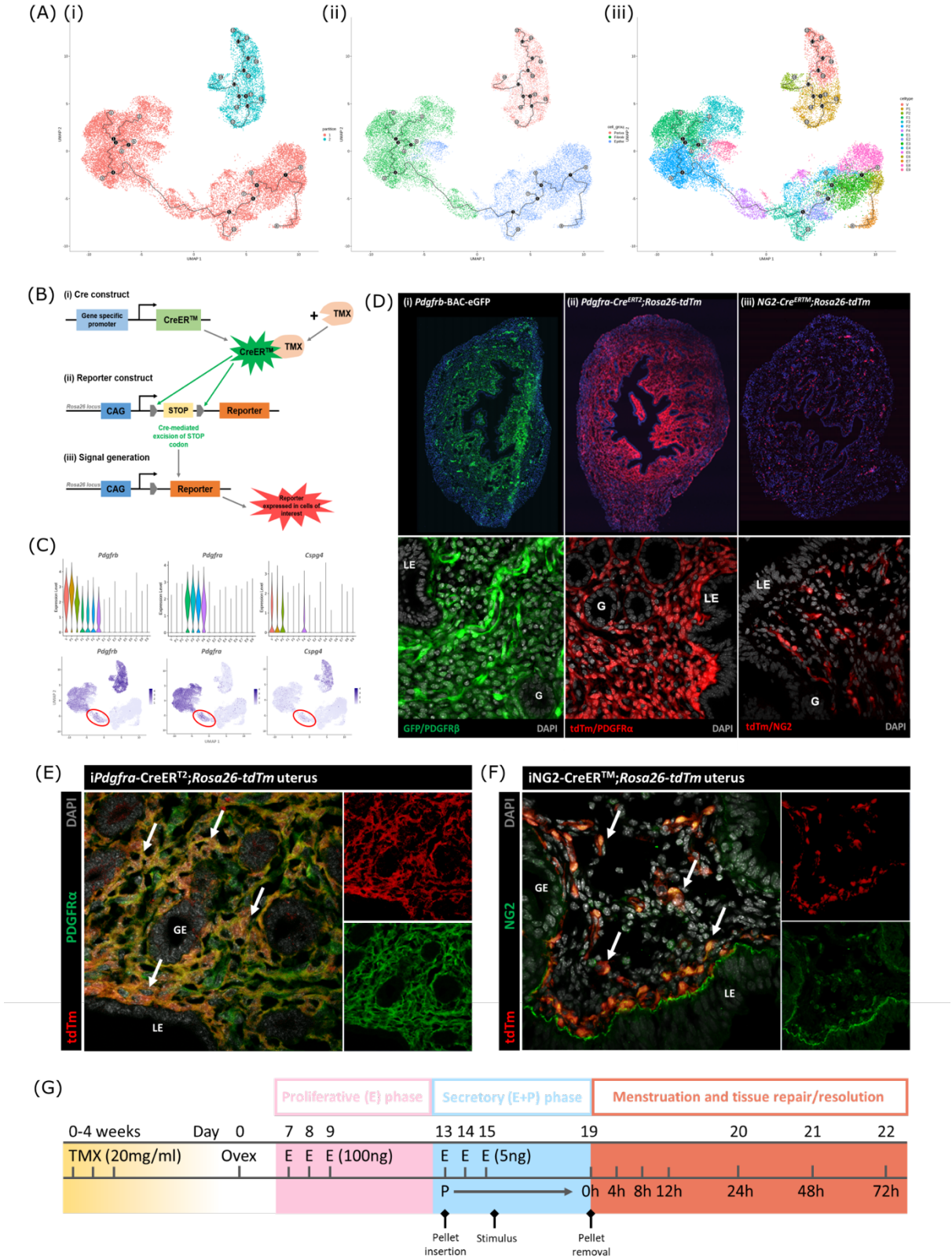

Figure 5: Lineage tracing strategy using inducible cre recombinase system to target mesenchymal subpopulations in endometrium of adult mice.

UMAP visualisation of trajectory analysis. ***(i)*** *Monocle3 separated data into two 'partitions' placing fibroblasts and epithelial cells into the same trajectory.* ***(ii)*** *Monocle3 revealed a putative differentiation trajectory between fibroblasts (green) and epithelial cells (blue) but*

with no link to perivascular cells (pink). ***(iii)*** *The trajectory between fibroblasts and epithelial cells runs through the repair-specific fibroblast cluster F4.* ***(B)*** *Schematic: representation of induced reporter expression in cells of interest using the tamoxifen inducible cre recombinase system.* ***(C)*** *Gene expression plots (scRNAseq data):* ***(i)*** *Pdgfrb all mesenchymal cell clusters* ***(ii)*** *Pdgfra fibroblasts,* ***(iii)*** *Cspg4/NG2 perivascular cells no expression of these genes in epithelial cell populations.* ***(D)*** *Expression of reporter proteins:* ***(i)*** *GFP+ mesenchyme in Pdgfrb-BAC-eGFP mouse endometrium (see (Kirkwood et al., 2021));* ***(ii)*** *tdTm detected in fibroblasts in iPdgfra-Cre^{ERT2};Rosa26-tdTm mouse endometrium;* ***(iii)*** *tdTm detected in perivascular cells in iNG2-Cre^{ERTM};Rosa26-tdTm mouse endometrium. Epithelial cells (LE/G) do not express reporter proteins any of these mice (LE= luminal epithelium, G= epithelial glands).* ***(E)*** *Expression of tdTm reporter and native PDGFRα protein in iPdgfra-Cre^{ERT2};Rosa26-tdTm mouse endometrium, co-localisation of tdTm and PDGFRα detected throughout (white arrows)****.*** ***(F)*** *Expression of tdTm reporter and native NG2 protein in iNG2-Cre^{ERTM};Rosa26-tdTm mouse endometrium, co-localisation of tdTm and NG2 detected throughout (white arrows).* ***(G)*** *Schematic: mouse model of endometrial tissue breakdown and repair modified to include a 3x injections with TMX followed by a 4w washout period prior to ovariectomy.*

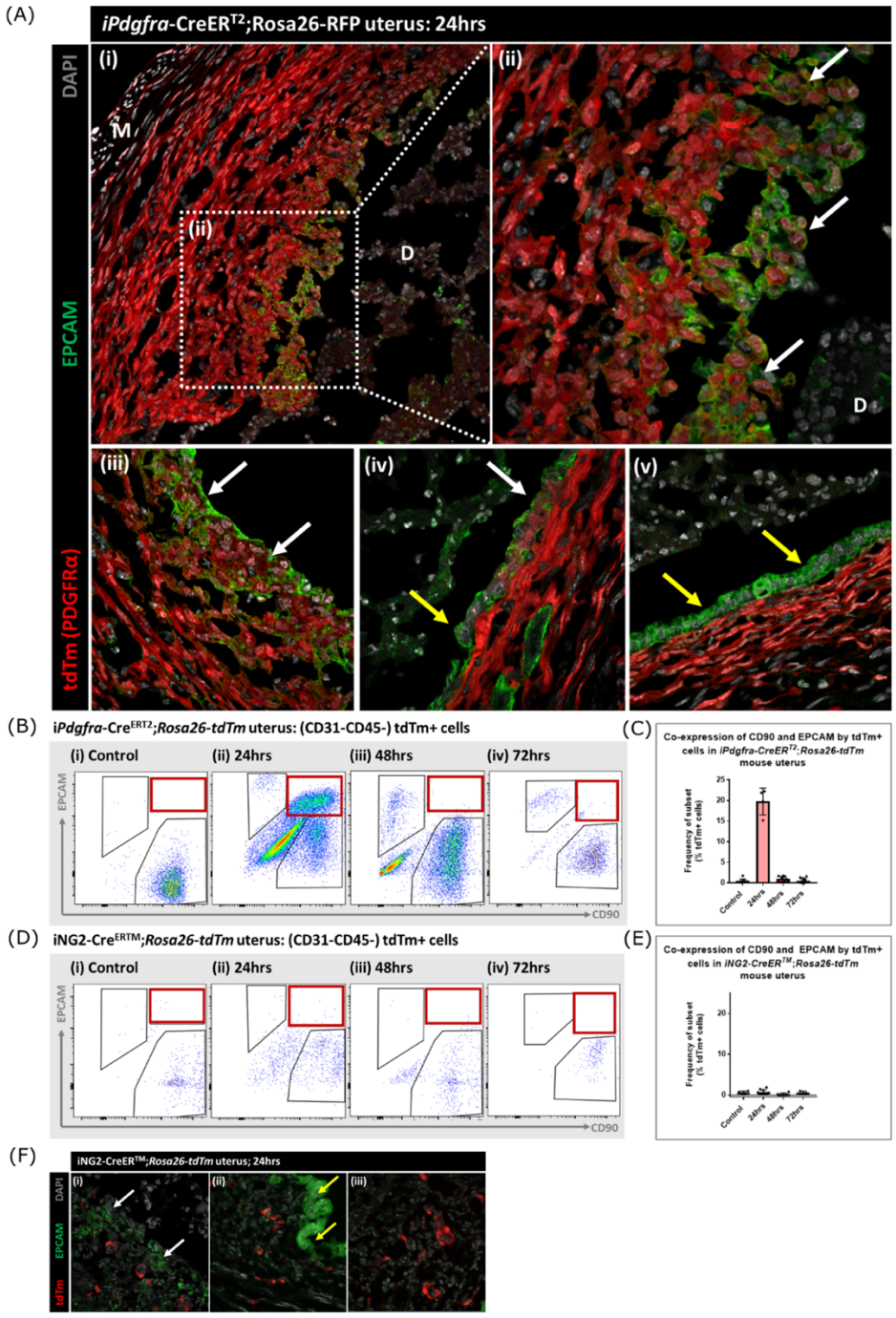

Figure 6: Lineage tracing of Pdgfra+ cells identifies a population that undergoes MET. (A) *Analysis of tdTm reporter protein and canonical epithelial cell marker EPCAM in iPdgfra-Cre^{ERT2};Rosa26-tdTm uterine tissues 24h following progesterone withdrawal when the tissue is undergoing simultaneous decidual breakdown and repair.* ***(i-iii)*** *tdTm+EPCAM+ cells can be detected in regions where the decidualised tissue is detached and the underlying stromal surface left exposed (white arrows) and* ***(iv)*** *in regions where a new epithelial layer is formed (white arrows) adjacent to tdTm- epithelial cells (yellow arrows).* ***(v)*** *Regions of epithelium where no tdTm+EPCAM+ cells were detected exist in other regions of the tissue (yellow arrows).* ***(B)*** *FC: analysis of mesenchymal marker CD90 and epithelial marker EPCAM by tdTm+ cells in iPdgfra-Cre^{ERT2};Rosa26-tdTm uterine tissues 24/48/72h following progesterone withdrawal, detection of tdTm+CD90+EPCAM+ cells in 24h tissues, red box.* ***(C)*** *Bar plot:*

quantification of tdTm+CD90+EPCAM+ cells at 24/48/72h in iPdgfra-Cre^{ERT2};Rosa26-tdTm uterine tissue, calculated as a frequency of tdTm+ cells: Control (n=9): 0.45±0.51%; 24h (n=4): 19.76±3.24%; 48h (n=9): 0.96±0.64%; 72h (n=13): 0.48±0.54%. ***(D)*** *FC: analysis of mesenchymal marker CD90 and epithelial marker EPCAM by tdTm+ cells in iNG2-Cre^{ERTM};Rosa26-tdTm uterine tissues 24/48/72h following progesterone withdrawal, no detection of tdTm+CD90+EPCAM+ cells in 24h tissues, red box.* ***(E)*** *Bar plot: quantification of tdTm+CD90+EPCAM+ cells in 24/48/72h iNG2-Cre^{ERTM};Rosa26-tdTm uterine tissues, , calculated as a frequency of tdTm+ cells: Control (n=4): 0.58±0.34%; 24h (n=12): 0.64±0.55%; 48h (n=6): 0.18±0.29%; 72h (n=3): 0.61±0.32%.* ***(F)*** *IHC: analysis of tdTm reporter protein and canonical epithelial cell marker EPCAM in iNG2-Cre^{ERTM};Rosa26-tdTm uterine tissues 24h following progesterone withdrawal when the tissue is undergoing simultaneous decidual breakdown and repair. tdTm+EPCAM- and tdTm-EPCAM+ cells can be detected in* ***(i)*** *regions where the decidualised tissue is detached and the underlying stromal surface left exposed (white and yellow arrows respectively),* ***(ii)*** *regions where a new epithelial layer is formed and (iii) regions of residual epithelium. No co-localisation of tdTm and EPCAM is detected in any region throughout the tissue.*

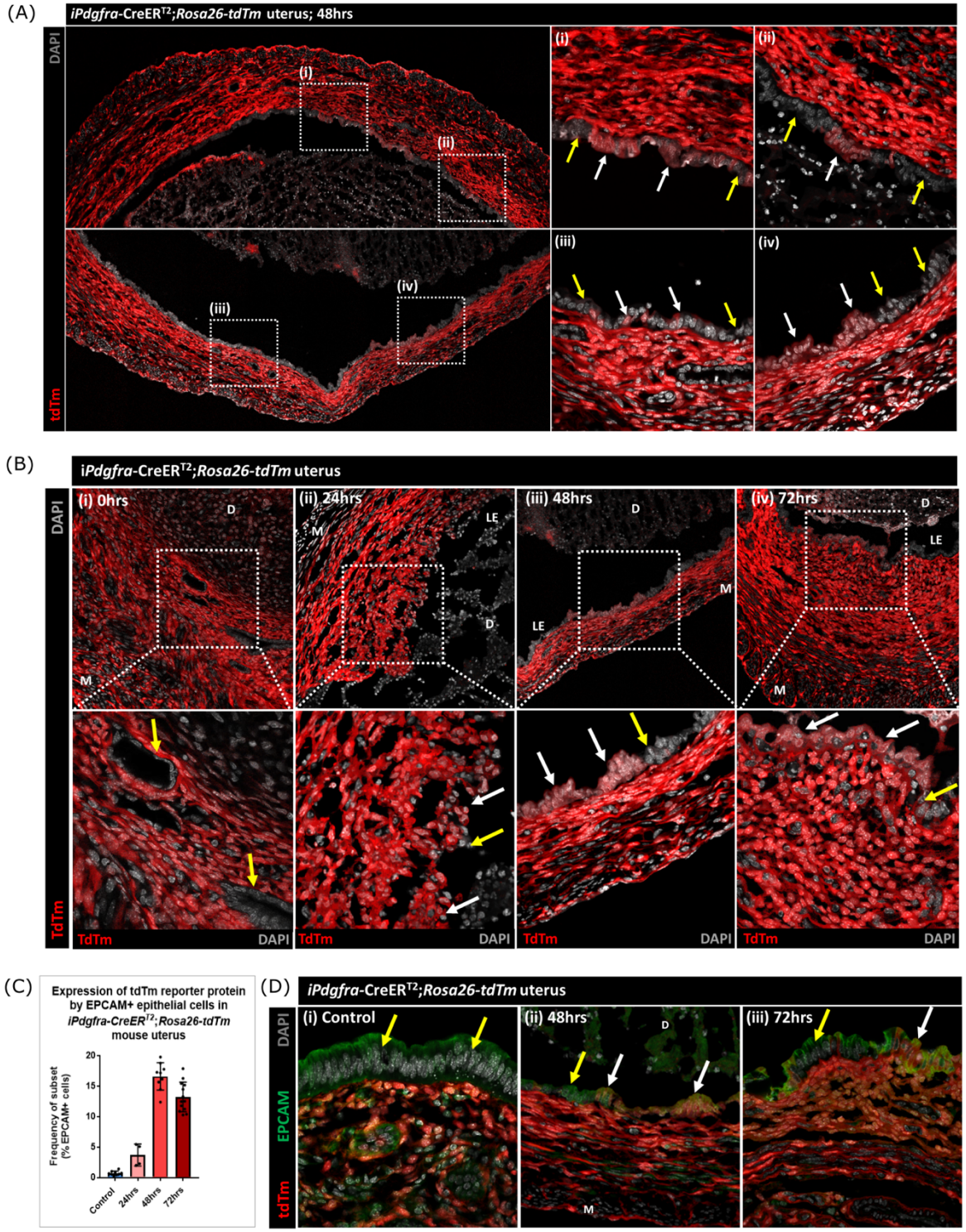

Figure 7: Lineage tracing studies confirm that epithelial cells derived from PDGFRα+ fibroblasts persist in the post repair luminal epithelium.

(A) *Analysis of tdTm reporter protein expression in iPdgfra-Cre^{ERT2};Rosa26-tdTm uterine tissues 48h following progesterone withdrawal when the tissue is remodelling and the luminal epithelium is re-instated.* ***(i-iv)*** *tdTm is expressed in stromal fibroblasts throughout the stromal tissue and in some cells in the luminal epithelium (white arrows) adjacent to tdTm- epithelial cells (yellow arrows).* ***(B)*** *IHC: analysis of tdTm reporter protein expression in iPdgfra-*

Cre^{ERT2};Rosa26-tdTm uterine tissues 0, 24, 48 and 72h following progesterone withdrawal, tdTm+ cells are detected within the renewed luminal epithelium at both 48 and 72h (white arrows) indicative of long term survival in the new epithelial layer. ***(C)*** *Bar plot: quantification of FC data analysing EPCAM+ cells that express tdTm in iPdgfra-Cre^{ERT2};Rosa26-tdTm uterine tissues 24, 48 and 72h following progesterone withdrawal, calculated as a frequency of EPCAM+ cells: Control (n=9): 0.75±0.35%; 24h (n=4): 3.73±1.81%; 48h (n=9): 16.63±2.24%; 72h (n=13): 13.26±2.42%.* ***(D)*** *IHC: analysis of tdTm reporter and epithelial marker EPCAM expression in iPdgfra-Cre^{ERT2};Rosa26-tdTm uterine tissues 48 and 72h following progesterone withdrawal.* ***(ii-iii)*** *tdTm+EPCAM+ cells can be detected in the luminal epithelium at 48/72h (white arrows) adjacent to tdTm-EPCAM+ epithelial cells (yellow arrows) and are not detected in* ***(i)*** *steady state control tissues.*