



OPEN

DATA DESCRIPTOR

A spatial transcriptomics dataset of the endometrium from repeated implantation failure patients

Yan Zhang^{1,4}, Junlin Pan^{1,4}, Zhitao Liu¹, Meimei Zheng¹, Ruijuan Liu¹, Lingshan Lei², Haiping Liu¹✉ & Peng Wang³✉

Repeated implantation failure (RIF) presents a significant challenge to *in vitro* fertilization (IVF), and abnormalities in the endometrium often play a crucial role in RIF. To enhance and deepen our understanding of tissue context in the endometrium underlying RIF, we applied spatial transcriptomics (ST) sequencing using the 10x Visium platform to 8 endometrial tissues from 4 normal individuals (CTR) and 4 RIF patients during the mid-luteal phase in this study. We achieved 10131 high-quality spots, with a median detected gene number of 3156. Seven distinct cellular niches (Niche 1–7) with specific characteristics were identified in our data. Deconvolution of ST data was performed by integrating with a public single-cell RNA (scRNA) dataset, and it was found that unciliated Epithelia were the dominant components in our data. To the best of our knowledge, this is the first spatial transcriptomics atlas of endometrial tissue in normal and RIF conditions, and the spatial data can provide a valuable source for further studies investigating the RIF mechanism.

Background & Summary

Repeated implantation failure (RIF) is defined as the failure of good-quality embryos to achieve a clinical pregnancy after transfer in multiple *in vitro* fertilization (IVF) treatment cycles^{1,2}. It is estimated that RIF occurs in approximately 5–15% of couples undergoing IVF^{3,4}. The causes of RIF are complex and can be broadly categorized into maternal and embryonic factors⁵. Compared to embryo quality, the assessment of the maternal factors, particularly the endometrium, is far more limited. Therefore, thoroughly exploring and understanding endometrial functions is critical for improving the success rate of IVF treatment^{6–8}.

Abnormal gene expression in the endometrium is frequently associated with RIF. Bulk RNA sequencing and microarray analyses focused on the endometrium have identified numerous dysregulated biomarkers in patients with RIF, including key mRNA and long noncoding RNA (lncRNA) hub genes^{9–11}. Furthermore, since endometrial receptivity is highly correlated with reproductive outcomes⁷, some of these identified biomarkers can be utilized in endometrial receptivity array (ERA) testing¹². Among the differentially expressed genes (DEGs), some biomarkers were identified related to immune factors significantly contributing to RIF through abnormal immune response¹³, immune infiltration¹¹, imbalances of immune cell types^{14,15}, and endometrial disruption¹⁶. However, there is notable inconsistency in the results from these studies, and the underlying biological mechanisms remains poorly understood¹⁷.

Single-cell RNA (scRNA) and spatial transcriptomics sequencing offer significant advantages in revealing cell type heterogeneity, spatial cell locations and communications, and cell-specific gene expression^{18–20}. Recently, many studies utilizing scRNA-seq focused on various aspects of the endometrium, including constructing normal or disordered endometrial tissue cell maps^{21–24}, elucidating cellular compositions and gene expression dynamics throughout the menstrual cycle²⁵, and exploring cell-cell crosstalk between the embryo and the endometrium, particularly the role of trophoblast on embryo implantation^{25–27}. Additionally, several single-cell RNA studies focused on RIF and identified abnormal gene expression patterns reported^{28–30}. To further investigate the communications among these cells, spatial transcriptomics has been employed in

¹Reproductive Medicine Center, The 960th Hospital of the PLA Joint Logistics Support Force, Jinan, 250031, China.

²Wuhan Frasergen Bioinformatics Co. Ltd., Wuhan, 430070, China. ³National Clinical Research Center for Mental Disorders & Beijing Key Laboratory of Mental Disorders, Beijing Anding Hospital & Advanced Innovation Center for Human Brain Protection, Beijing, 100088, China. ⁴These authors contributed equally: Yan Zhang, Junlin Pan. ✉e-mail: liuhaipingweihai@126.com; wangpeng@cibr.ac.cn

Sample	Group	Age (years)	BMI (kg/m ²)	Smoking Status	Previously transferred high-quality embryos	Parity
CTR1	CTR	25	18.14	None	0	1
CTR2	CTR	24	20.25	None	0	1
CTR3	CTR	32	24.37	None	0	2
CTR4	CTR	31	20.32	None	0	2
RIF1	RIF	31	22.18	None	3	0
RIF2	RIF	33	24.47	None	5	0
RIF3	RIF	25	21.64	None	3	0
RIF4	RIF	29	20.35	None	3	0

Table 1. Clinical characteristics of subjects included in the control and RIF groups. CTR means normal groups, and RIF means repeated implantation failure patients.

some studies. For instance, spatial and single-cell transcriptomics have been used to characterize the spatial gene expression features of the normal human endometrium throughout the menstrual cycle³¹ or *in vivo* and *in vitro*³². Additionally, other studies aim to create a spatiotemporal atlas of the human maternal-fetal interface³³. While these advancements offer a more comprehensive understanding of the endometrium, there remains a notable lack of spatial transcriptomics studies to identify cell type profiles, cell co-locations and gene expression atlas in the context of RIF.

Therefore, we performed 10x Visium Spatial Transcriptome (ST) Sequencing on 8 samples from 4 normal individuals and 4 patients with RIF in this study. After performing quality control and unsupervised clustering, we divided the spots into 7 niches with specifically expressed genes. Furthermore, we characterized the cellular compositions in our spots by integrating the public scRNA data, which reveals that the dominant cell type is epithelial cells. In summary, to our knowledge, this is the first study employing ST technology in RIF. We believe that the data source from this study will provide valuable references for further study on the molecular mechanism and treatment biomarkers of RIF.

Methods

Patients and samples enrollment. In this study, endometrial samples were collected after obtaining written consent from each participant. The research protocol was approved by the Medical Ethics Committee of The 960th Hospital of the PLA Joint Logistics Support Force. The study included four RIF subjects who had a history of failure to achieve clinical pregnancy after ≥ 3 embryo transfers of good-quality embryos³⁴. In comparison, the control group (referred to as the CTR group) consisted of 4 multiparous women without a history of miscarriage. All participants were confirmed free of uterine pathologies (including endometriosis and adenomyosis), endocrine, metabolic, auto-immune diseases and infection. The age range of the subjects was set to be ≤ 35 years, and the BMI threshold was set to be < 28 kg/m². Demographic details of the subjects recruited for spatial RNA sequencing in both the RIF and CTR groups are listed in Table 1. Combining transvaginal ultrasound with urinary LH dipstick testing, we were able to detect the Luteinizing Hormone (LH) surge (LH + 0). Endometrial samples of both CTR and RIF groups were collected at the fundal and upper part of the uterus using Pipelle endometrial biopsy during the LH + 7.

10x Visium spatial transcriptomics sequencing. Fresh endometrial tissues were obtained and rapidly frozen in isopentane pre-chilled with liquid nitrogen. The tissues were then stored at -80°C . Fresh frozen tissues were sectioned into slices. Initially, RNA quality was assessed to ensure a minimum RNA Integrity Number (RIN) larger than 7 to minimize the degradation of RNA molecules. The optimal tissue optimization time was determined based on fluorescence imaging strength. The tissue splices were carefully placed onto designated capture areas on the 10x Visium Spatial Tissue Optimization Slide. Each slide contained four capture areas, each measuring 6.5 mm*6.5 mm and containing about 5,000 spots equipped with barcode sequences. Standard methanol fixation, hematoxylin and eosin (H&E) staining, and brightfield microscopes were utilized to visualize tissue sections. The tissue was permeabilized to release mRNA molecules, which were then captured by the adjacent spots. mRNA was subjected to reverse transcription to generate cDNA. Subsequently, libraries were constructed according to the standard protocol. The Illumina NovaSeq 6000 platform was employed for sequencing using the PE150 model.

Alignment, quality control, and clustering of visium data. The Space Ranger count pipeline (version 2.0.0) automatically aligned spatial transcriptome data to the human reference genome (GRCh38-2020-A), detected tissue sections, and aligned fiducial across all slices. To construct the Seurat object, we used the Seurat (version 4.3.0) *Load10X_Spatial* function to import the spatial feature-spot expression matrices, spatial location of spots and tissue slices. Spots with gene count below 500 or mitochondrial gene percentage exceeding 20% were excluded for each slice. Spot expression data normalization was done using the *SCTransform* function for each slice, followed by merging all slices using *merge* function. Principal component analysis (PCA) was conducted using the top 30 principal components (PCs) and dimension reduction was performed with a resolution of 0.6. Differential gene expression analysis was performed among spatial clusters or niches using the *FindAllMarkers* function. Spots with similar gene expression profiles were grouped into a single niche.

Sample	Number of Reads	Sequencing Saturation (%)	Q30 Bases in Barcode (%)	Q30 Bases in RNA Read (%)	Q30 Bases in UMI (%)
CTR1	333,340,739	93.74	95.42	90.76	91.37
CTR1	301,378,472	93.68	95.79	91.08	91.57
CTR1	298,816,693	90.09	95.69	91.61	91.65
CTR1	334,052,677	91.54	95.63	92.03	91.4
RIF1	284,189,121	96.15	96.39	93.34	92.51
RIF2	337,594,141	94.95	95.35	90.63	91.26
RIF3	298,121,637	97.26	96.39	93.28	93.08
RIF4	279,198,231	95.15	95.54	90.74	91.69

Table 2. The detailed QC matrices of the spatial transcriptomics FASTQ files for the eight samples.

Sample	Number of Spots Under Tissue	Median Genes per Spot	Median UMI Counts per Spot	Reads Mapped to Genome (%)
CTR1	751	3,711	8,813	91.17
CTR1	820	3,825	9,761	94.64
CTR1	1,805	2,764	6,097	94.12
CTR1	2,018	2,664	5,837	95.11
RIF1	1,334	2,350	4,020	94.41
RIF2	1,367	3,346	7,306	92.46
RIF3	914	2,410	4,388	95.25
RIF4	1,154	3,486	8,546	91.47

Table 3. The detailed QC matrices of the spatial transcriptomics libraries for the eight samples.

Public single-cell data processing. Public single-cell data (GSE183837)²⁸ were downloaded from the GEO database. In total, 9 samples were obtained, containing 3 normal individuals and 6 RIF patients. The expression matrices from these samples were used for analysis. Several preprocessing steps were performed to ensure data quality; the poorly quality cells were first filtered out when they hit one of the following indices: the number of detected genes in the cell beyond the range 500–5000, UMI counts less than 800, or the mitochondrial gene percentage greater than 20%. In addition, the suspected double cells were removed using DoubletFinder (v2.0.3)³⁵. After quality control, one sample named RIF1 captured with too much cells (13917 cells) than the others (< 8200 cells per sample), so we removed this sample to mitigate potential technical variations. Finally, 31348 cells coming from 8 samples are retained for further analysis using Seurat (v4.3.0.1). The analysis comprised the selection of highly variable genes (HVGs), eliminating batch effects by Harmony³⁶, clustering cells, and annotating cell types or subtypes based on the expression of canonical markers.

Integration of spatial and single-cell data. The conditional autoregressive-based deconvolution (CARD) package (v1.1)³⁷ was used to deconvolve the mixture of cellular components within the spots of the spatial data. CARD employs a non-negative matrix factorization model to estimate the cell type proportions for each spot, and it requires the single-cell data information to implement the deconvolution process. We computed the cellular components by averaging percentages of the cell types in all spots for each sample separately.

Data Records

The raw and processed spatial transcriptomics data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under the accession number GSE287278³⁸.

Technical Validation

Quality matrix of the spatial transcriptomics data. 10x Genomics Visium sequencing data and H&E images were processed automatically by SpaceRanger (v2.0.0). When insight into FASTQ files, it shows that the read-pair number among all eight samples is about 3×10^8 , and the sequencing saturation is over 90%, thus our sequencing depth was sufficient to detect gene expression. The sequence data presents high sequencing accuracy with Q30 values for Barcode, UMI, and RNA Read all exceeding 90% (Table 2). Next, we summarized the quality matrix of the generated spatial libraries (Table 3). The number of spots under tissue shows some variation, ranging from 751 to 2018. The median genes and UMI Counts per spot among the samples were over 2000 and 4000, respectively. The rate of reads mapped to the genome was over 90% (Table 3).

After applying screening criteria with gene counts greater than 500 and mitochondrial gene percentages less than 20%, a total of 10131 spots were retained. For these high-quality spots from all eight samples, the median feature count was 3156, the average median UMI count was 6860, the average median gene count was 3074, and the average median percentage of the UMI count belonging to the mitochondrial gene was 5.5%. There were no significant differences in the distributions of UMI counts (nCount_Spatial), gene counts (nFeature_Spatial) and mitochondrial genes percentage (percent_mito) among the control (CTR) and RIF samples (Fig. 1a). Moreover,

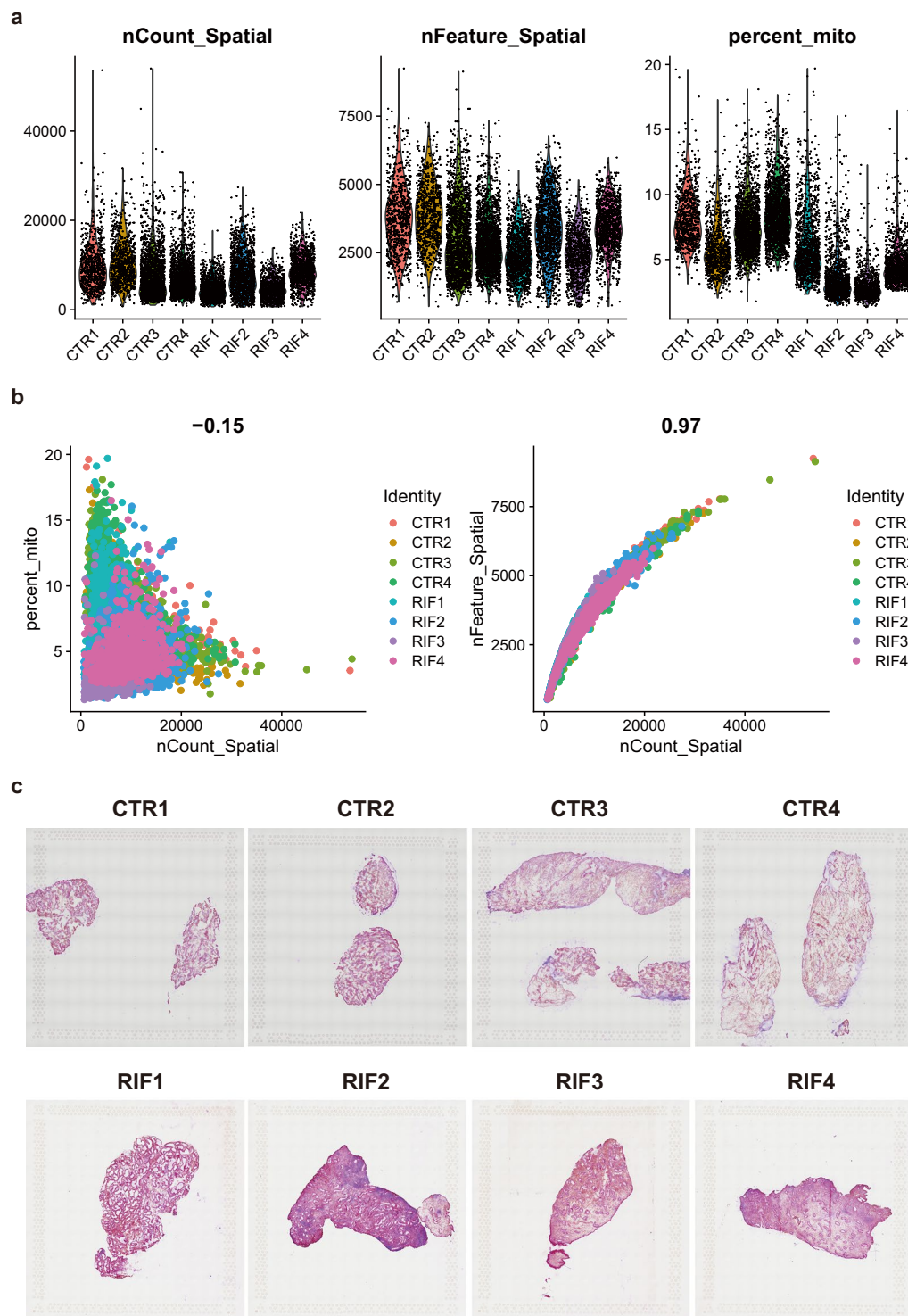


Fig. 1 Quality control of the spatial transcriptomics dataset of the endometrium from four normal samples and four repeated implantation failure patients. **(a)** A Violin diagram illustrating the distributions of the number of unique molecular identifiers (UMIs), (nCount_Spatial), the number of genes (nFeature_Spatial) and the percentage of mitochondrial genes (percent_mito) in each spot of 8 samples. **(b)** Scatter diagrams showing the relationships between the UMIs counts and the percentage of mitochondrial genes (left), and between the UMIs counts and the gene counts (right). **(c)** Hematoxylin and eosin (H&E) staining images for eight samples.

as expected, the relationship between nCount_Spatial and percent_mito is very weak ($R = -0.15$), and the correlation coefficient between nCount_Spatial and nFeature_Spatial is very high ($R = 0.98$) (Fig. 1b). At last, when insight into the niche H&E images, it was observed that there were some morphological differences among the eight samples (Fig. 1c).

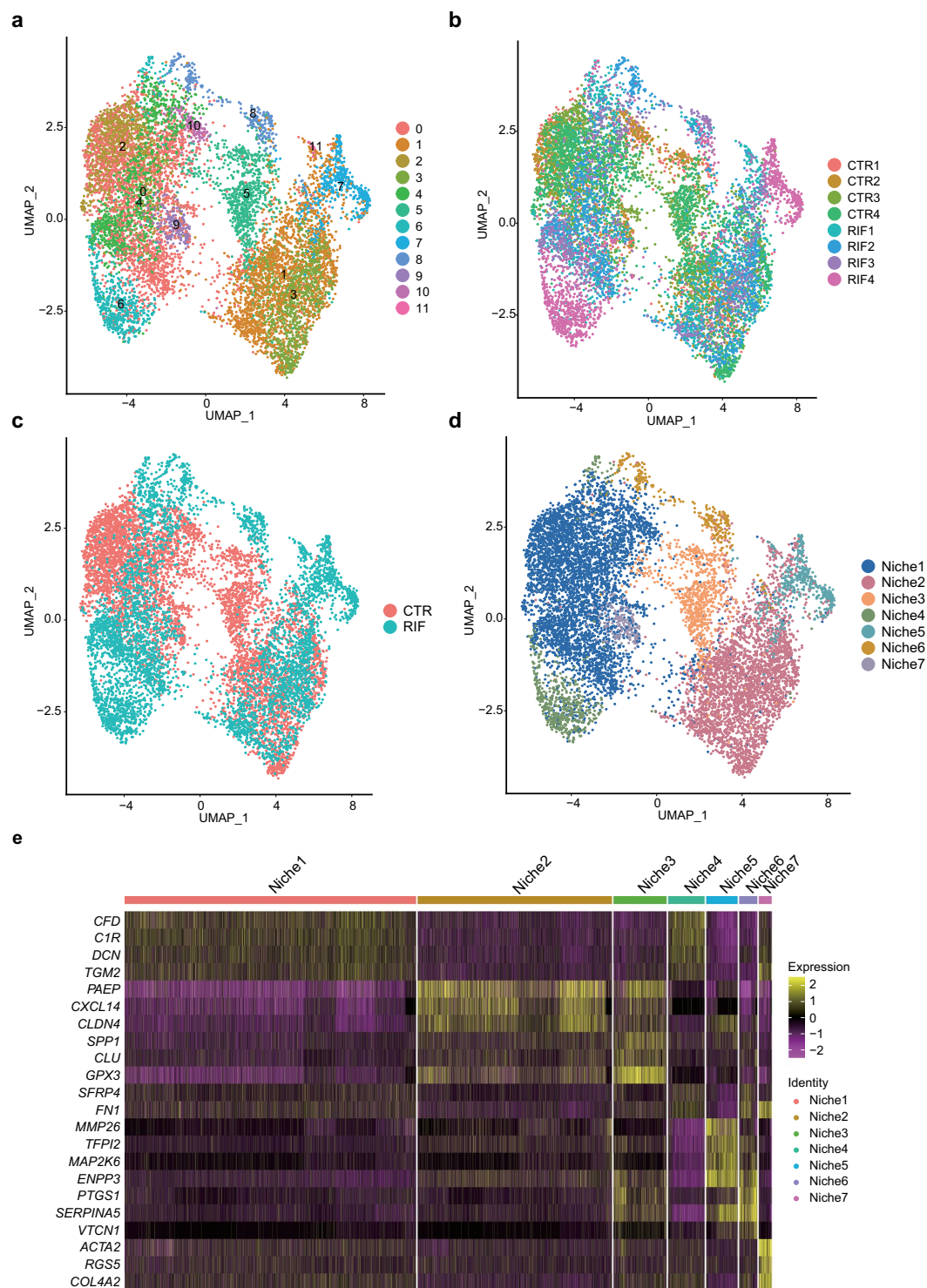


Fig. 2 Clustering of 10x Visium spots and constructing the spatial atlas. **(a)** Unsupervised clustering of 10131 spots into 12 clusters, and visualization on a non-linear dimension reduction method named 2D Uniform Manifold Approximation and Projection (UMAP). The 2D-UMAP plotting colored by sample **(b)** and group **(c)** information. **(d)** The 11 clusters were merged into 7 niches and shown using the UMAP plot. **(e)** The heatmap plot of the top 6 marker genes based on log₂(Fold Change) values for each niche.

Unsupervised clustering of the spatial transcriptomics data. To construct the spatial transcriptomics atlas, we clustered all spots from the RIF and CTR groups based on the gene expression profiles. Following spatial high variable genes identification, dimension reduction and unsupervised clustering, 12 clusters with distinct gene expression patterns were identified and the clusters were visualized using 2D UMAP plot (Fig. 2a). Additionally, sample and group information were visualized through using 2D UMAP plot, and batch effects

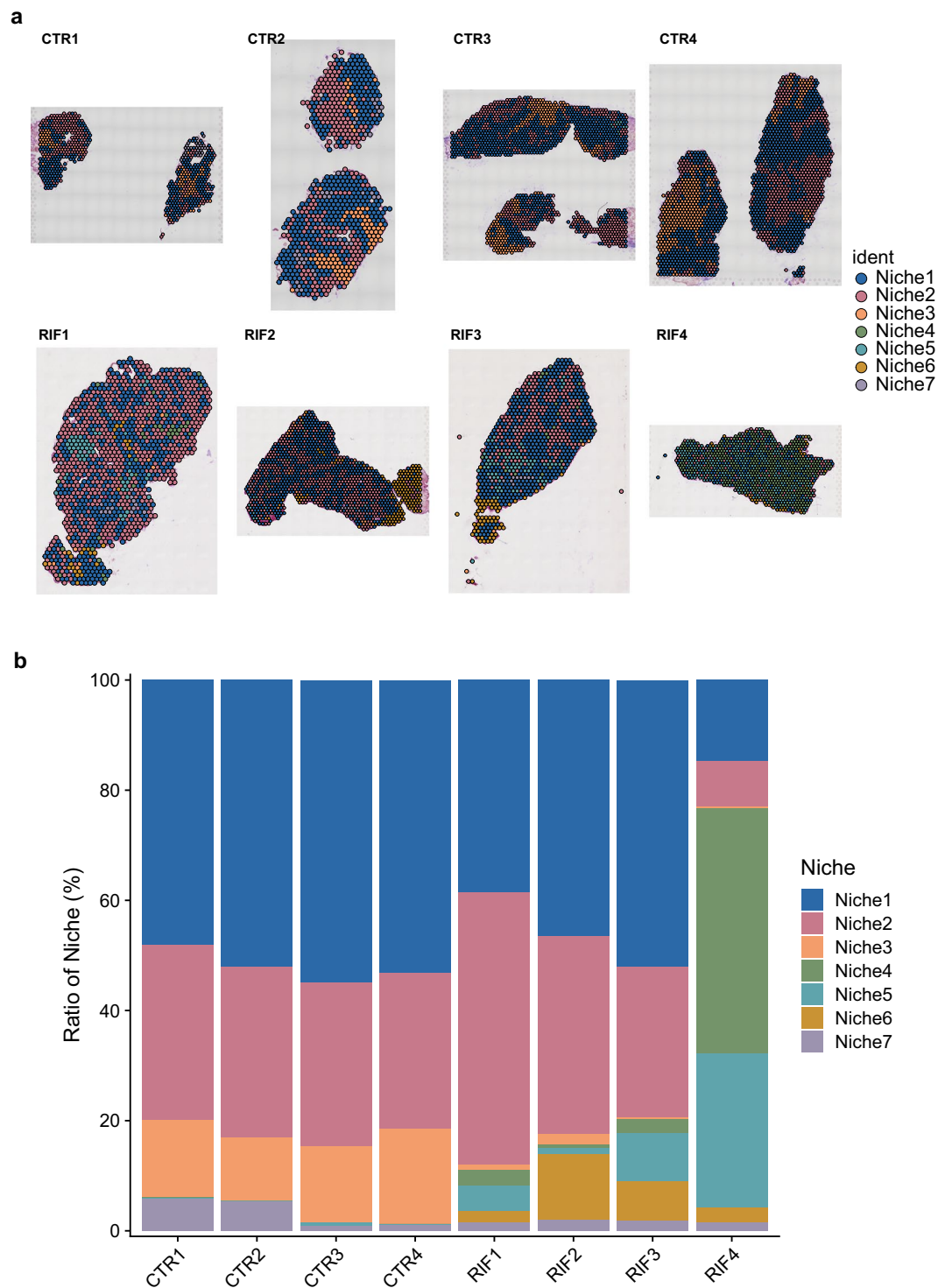


Fig. 3 The characteristics of the seven niches in the spatial data. **(a)** The spatial distributions of the seven niches in the splices from the eight samples. **(b)** The ratios of spots belonging to the niches among the eight samples.

were not observed in our data, as evidenced by spot distribution patterns of most clusters contained contributions from multiple samples (Fig. 2b,c). We merged clusters into 7 niches based on similar gene expression patterns (Fig. 2d), and visualized the top 6 marker genes for each niche via a heatmap plot (Fig. 2e). In detail, Niche1 was characterized by high expression of *CFD*, *C1R*, and *DCN*. Niche2 exhibited high expression of *PAEP*, *CXCL14* and *CLDN4*. Niche3 was characterized by specific expression of *SPP1*, *CLU* and *GPX3*. Niche4 showed a specific high expression of *SFRP4*. Niche5 was characterized by specific expression of *MMP26*, *TAPI2*, *MAP2K6*, and *ENPP3*. *VTCN1*, *PTGS1*, *LGR5* and *SERPINA5* are the top markers of Niche6. Niche7 was primarily marked by smooth muscle (SMCs) cell markers such as *ACTA2*, *RGS5* and *COL4A2*, indicating that the spots in Niche7 mainly are

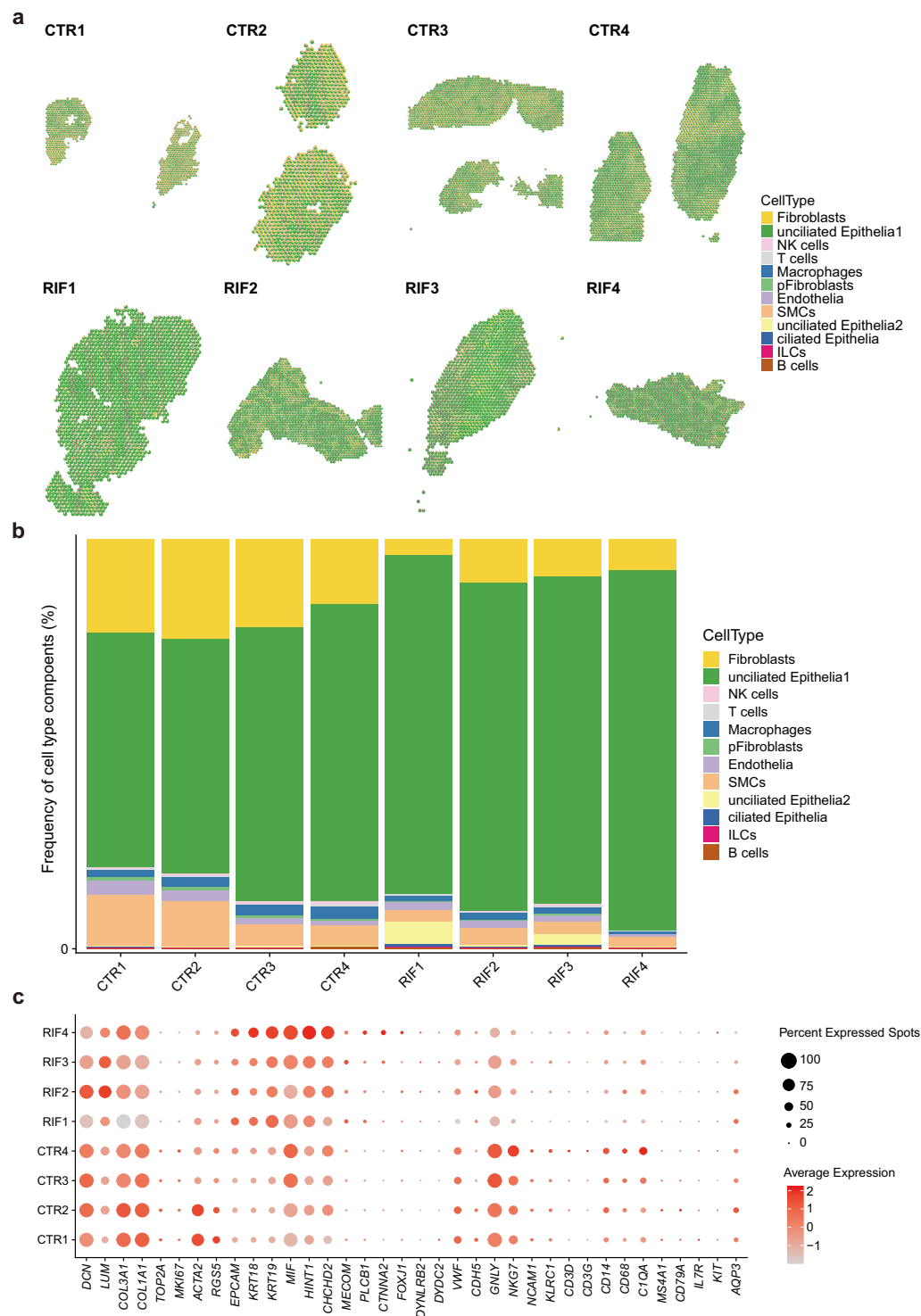


Fig. 4 The cell type spatial atlas of the endometrium. **(a)** The pie charts of cell-type compositions of ST data deconvolved by CARD for each sample. **(b)** The averaged cell-type composition frequencies of the intra-sample spots among the eight samples. **(c)** A bubble plot showing the expression patterns of the marker genes among the spatial data of the eight samples.

SMCs. Furthermore, there are some divergent characterizations of the spatial distributions for these niches in different samples or groups (Fig. 3a). Finally, we quantified the niche frequency across all samples, which shows that Niche1 and Niche2 constitute the primary components in 7 samples, except RIF4 uniquely exhibited Niche4 as its predominant component (Fig. 3b). Besides, the frequencies of other Niches, such as 3, 5, and 6 also show some variations across the eight samples (Fig. 3b).

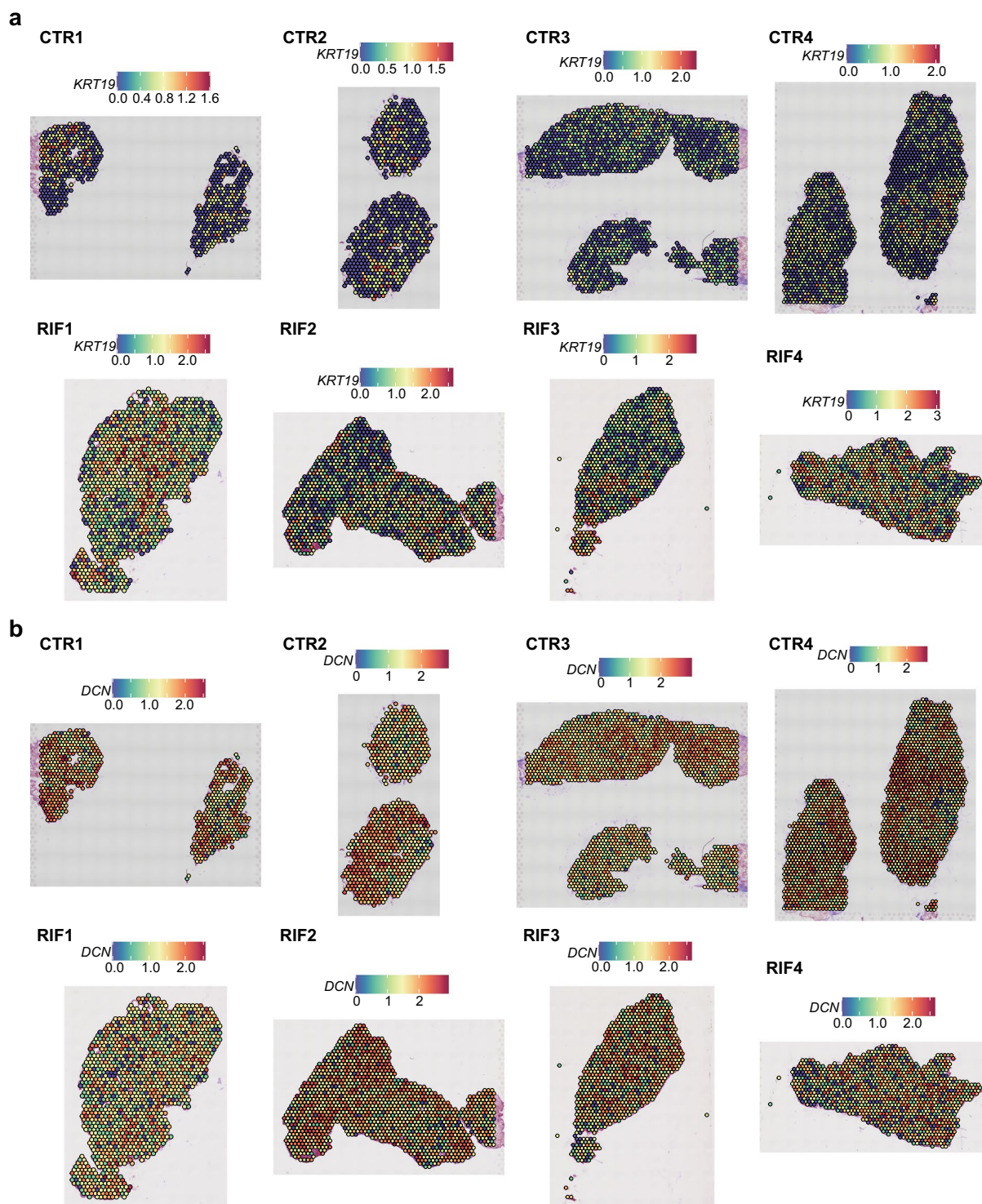


Fig. 5 The spatial expression of mark genes. **(a)** The spatial expression of the marker of epithelial cells (marked by *KRT19*) in H&E anatomic images among the eight samples. **(b)** The spatial expression of the marker of fibroblast (marked by *DCN*) in H&E anatomic images among the eight samples.

Spatial cellular landscapes of the endometrium in RIF. Since the 10x Visium platform captured a mixture of several cells for each spot and does not achieve single-cell resolution, we integrated the spatial data with the publicly available single-cell data to deconvolute the cell components within each spot. First, after performing quality control of the public single-cell data, we obtained a total of 31,348 high-quality single cells from eight samples. Using unsupervised clustering with Seurat, these cells were classified into 24 clusters (Supplementary Fig. 1a), and further annotated into 12 cell types based on the expression patterns of classical known markers^{21–23,25,28} (Supplementary Fig. 1b,c). The atlas showed that the major cell types in endometrial tissues are non-immune

cells, including fibroblasts (*DCN*, *LUM*), proliferating fibroblasts (abbreviated as pFibroblasts, *TOP2A*, *MKI67*), smooth muscle cells (SMCs, *ACTA2*, *RGS5*), epithelial cells (*EPCAM*, *KRT18*, *KRT19*), and endothelial cells (*VWF*, *CDH15*). Furthermore, epithelial cells were further subclustered into three categories: unciliated epithelia1 (*MIF*, *HINT1*, *CHCHD2*), unciliated epithelia2 (*CTNNA2*) and ciliated Epithelia (*FOXJ1*, *DYNLRB2*, *DYDC2*). Immune cell types are all present with low frequencies in the tissue, such as T cells (*CD3D*, *CD3G*), NK cells (*GNLY*, *NKG7*, *NCAM1*, *KLRC1*), B cells (*MS4A1*, *CD79A*), macrophages (*CD14*, *CD68* and *CIQA*) and innate lymphoid cells (ILCs, *PTPRC*, *IL7R*, *KIT*, *AQP3*) (Supplementary Fig. 1c).

We then transferred the annotated cell type information from scRNA data to our spatial transcriptome data using CARD. We presented the spatial distributions of cellular components in each spot (Fig. 4a) and the sample-specific mean cellular frequencies across the within-sample spots (Fig. 4b). The results showed that all the cell types captured in the single cell data replicated in our data, and the most abundant cell component is unciliated epithelia1, with an average percentage of approximately 73.94% across all spots (Fig. 4b). On the other hand, unciliated epithelial1 cells constituted the first predominant population in 10,103/10,131 spots. Other cell types with relatively high average proportions include fibroblasts (14.27%), SMCs (5.15%), macrophages (2.05%), endothelia (1.66%) and unciliated epithelia2 (1.2%) (Fig. 4b). To ensure the accuracy of the prediction results, we also verified the expressions of the canonical known markers in our spatial data using bubble plot (Fig. 4c). The analysis revealed robust expression of fibroblast and epithelial cell markers with high percent expressed spots and average expression (Fig. 4c), and widespread spatial expression distributions (Fig. 5a,b), while other cell markers present lowly expressed (Fig. 4c). In summary, the expression patterns of these marks validated the spatial deconvolution results.

Data availability

The data spatial transcriptomics is available at the Gene Expression Omnibus (GEO) database under the accession number [GSE287278](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE287278)³⁸.

Code availability

All spatial transcriptomics sequencing data were processed using SpaceRanger software (<https://www.10xgenomics.com/cn/support/software/cell-ranger/downloads>) for spot-gene expression matrix construction, and followed by Seurat based on the official manual (https://satijalab.org/seurat/articles/get_started) for spots clustering. Besides, the process for scRNA data was also based on the official protocol of Cellranger and Seurat, similar to spatial data. Deconvolution of spatial transcriptomics data was based on the manual of CARD (<https://yma-lab.github.io/CARD/>).

Received: 29 January 2025; Accepted: 17 September 2025;

Published online: 29 October 2025

References

- Simon, A. & Laufer, N. Repeated implantation failure: clinical approach. *Fertil Steril* **97**, 1039–1043, <https://doi.org/10.1016/j.fertnstert.2012.03.010> (2012).
- Cimadomo, D., Craciunas, L., Vermeulen, N., Vomstein, K. & Toth, B. Definition, diagnostic and therapeutic options in recurrent implantation failure: an international survey of clinicians and embryologists. *Hum Reprod* **36**, 305–317, <https://doi.org/10.1093/humrep/deaa317> (2021).
- Pirtea, P. *et al.* Rate of true recurrent implantation failure is low: results of three successive frozen euploid single embryo transfers. *Fertil Steril* **115**, 45–53, <https://doi.org/10.1016/j.fertnstert.2020.07.002> (2021).
- Busnelli, A. *et al.* How common is real repeated implantation failure? An indirect estimate of the prevalence. *Reprod Biomed Online* **40**, 91–97, <https://doi.org/10.1016/j.rbmo.2019.10.014> (2020).
- Franasiak, J. M. *et al.* A review of the pathophysiology of recurrent implantation failure. *Fertil Steril* **116**, 1436–1448, <https://doi.org/10.1016/j.fertnstert.2021.09.014> (2021).
- Timeva, T., Shterev, A. & Kyurkchiev, S. Recurrent implantation failure: the role of the endometrium. *J Reprod Infertil* **15**, 173–183 (2014).
- Ruiz-Alonso, M. *et al.* The endometrial receptivity array for diagnosis and personalized embryo transfer as a treatment for patients with repeated implantation failure. *Fertil Steril* **100**, 818–824, <https://doi.org/10.1016/j.fertnstert.2013.05.004> (2013).
- Zou, Y. *et al.* The endometrial microbiota profile influenced pregnancy outcomes in patients with repeated implantation failure: a retrospective study. *J Reprod Immunol* **155**, 103782, <https://doi.org/10.1016/j.jri.2022.103782> (2023).
- Pathare, A. D. S. & Hinduja, I. Endometrial Expression of Cell Adhesion Genes in Recurrent Implantation Failure Patients in Ongoing IVF Cycle. *Reprod Sci* **29**, 513–523, <https://doi.org/10.1007/s43032-021-00708-x> (2022).
- Wang, Z. *et al.* ATF3 deficiency impairs the proliferative-secretory phase transition and decidualization in RIF patients. *Cell Death Dis* **12**, 387, <https://doi.org/10.1038/s41419-021-03679-8> (2021).
- Feng, X. *et al.* Identification of key genes and immune cell infiltration in recurrent implantation failure: A study based on integrated analysis of multiple microarray studies. *Am J Reprod Immunol* **88**, e13607, <https://doi.org/10.1111/aji.13607> (2022).
- Diaz-Gimeno, P. *et al.* The accuracy and reproducibility of the endometrial receptivity array is superior to histology as a diagnostic method for endometrial receptivity. *Fertil Steril* **99**, 508–517, <https://doi.org/10.1016/j.fertnstert.2012.09.046> (2013).
- Papuchova, H. *et al.* Endometrial HLA-F expression is influenced by genotypes and correlates differently with immune cell infiltration in IVF and recurrent implantation failure patients. *Hum Reprod* **37**, 1816–1834, <https://doi.org/10.1093/humrep/deac118> (2022).
- Kuroda, K. *et al.* Increasing number of implantation failures and pregnancy losses associated with elevated Th1/Th2 cell ratio. *Am J Reprod Immunol* **86**, e13429, <https://doi.org/10.1111/aji.13429> (2021).
- Berdiaki, A., Vergadi, E., Makrygiannakis, F., Vrekoussis, T. & Makrygiannakis, A. Repeated implantation failure is associated with increased Th17/Treg cell ratio, during the secretory phase of the human endometrium. *J Reprod Immunol* **161**, 104170, <https://doi.org/10.1016/j.jri.2023.104170> (2024).
- Diaz-Gimeno, P. *et al.* Predicting risk of endometrial failure: a biomarker signature that identifies a novel disruption independent of endometrial timing in patients undergoing hormonal replacement cycles. *Fertil Steril* **122**, 352–364, <https://doi.org/10.1016/j.fertnstert.2024.03.015> (2024).
- Walker, E. R., McGrane, M., Aplin, J. D., Brison, D. R. & Ruane, P. T. A systematic review of transcriptomic studies of the human endometrium reveals inconsistently reported differentially expressed genes. *Reprod Fertil* **4**, <https://doi.org/10.1530/RAF-22-0115> (2023).
- Jovic, D. *et al.* Single-cell RNA sequencing technologies and applications: A brief overview. *Clin Transl Med* **12**, e694, <https://doi.org/10.1002/ctm2.694> (2022).

19. Walker, B. L., Cang, Z., Ren, H., Bourgain-Chang, E. & Nie, Q. Deciphering tissue structure and function using spatial transcriptomics. *Commun Biol* **5**, 220, <https://doi.org/10.1038/s42003-022-03175-5> (2022).
20. Burgess, D. J. Spatial transcriptomics coming of age. *Nat Rev Genet* **20**, 317, <https://doi.org/10.1038/s41576-019-0129-z> (2019).
21. Krjutskov, K. *et al.* Single-cell transcriptome analysis of endometrial tissue. *Hum Reprod* **31**, 844–853, <https://doi.org/10.1093/humrep/dew008> (2016).
22. Queckborner, S. *et al.* Stromal Heterogeneity in the Human Proliferative Endometrium—A Single-Cell RNA Sequencing Study. *J Pers Med* **11**, <https://doi.org/10.3390/jpm11060448> (2021).
23. Mareckova, M. *et al.* An integrated single-cell reference atlas of the human endometrium. *Nat Genet* **56**, 1925–1937, <https://doi.org/10.1038/s41588-024-01873-w> (2024).
24. Liu, S. *et al.* Single-cell and spatial transcriptomic profiling revealed niche interactions sustaining growth of endometriotic lesions. *Cell Genom* **5**, 100737, <https://doi.org/10.1016/j.xgen.2024.100737> (2025).
25. Giudice, L. C. Multidimensional transcriptomic mapping of human endometrium at single-cell resolution. *Nat Med* **26**, 1513–1514, <https://doi.org/10.1038/s41591-020-1075-1> (2020).
26. Liu, D. *et al.* Primary specification of blastocyst trophectoderm by scRNA-seq: New insights into embryo implantation. *Sci Adv* **8**, eabj3725, <https://doi.org/10.1126/sciadv.abj3725> (2022).
27. Vento-Tormo, R. *et al.* Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353, <https://doi.org/10.1038/s41586-018-0698-6> (2018).
28. Lai, Z. Z. *et al.* Single-cell transcriptome profiling of the human endometrium of patients with recurrent implantation failure. *Theranostics* **12**, 6527–6547, <https://doi.org/10.7150/thno.74053> (2022).
29. Zhang, H., Zhang, C. & Zhang, S. Single-Cell RNA Transcriptome of the Human Endometrium Reveals Epithelial Characterizations Associated with Recurrent Implantation Failure. *Adv Biol (Weinh)* **8**, e2300110, <https://doi.org/10.1002/adbi.202300110> (2024).
30. Cao, D. *et al.* Time-series single-cell transcriptomic profiling of luteal-phase endometrium uncovers dynamic characteristics and its dysregulation in recurrent implantation failures. *Nat Commun* **16**, 137, <https://doi.org/10.1038/s41467-024-55419-z> (2025).
31. Wang, X. & Hawkins, S. M. Using advanced spatial and single-cell transcriptomics to characterize the human endometrium. *Nat Genet* **53**, 1628–1630, <https://doi.org/10.1038/s41588-021-00982-0> (2021).
32. Garcia-Alonso, L. *et al.* Mapping the temporal and spatial dynamics of the human endometrium *in vivo* and *in vitro*. *Nat Genet* **53**, 1698–1711, <https://doi.org/10.1038/s41588-021-00972-2> (2021).
33. Greenbaum, S. *et al.* A spatially resolved timeline of the human maternal-fetal interface. *Nature* **619**, 595–605, <https://doi.org/10.1038/s41586-023-06298-9> (2023).
34. Failure, E. W. Go. R. I. *et al.* ESHRE good practice recommendations on recurrent implantation failure. *Hum Reprod Open* **2023**, hoad023, <https://doi.org/10.1093/hropen/hoad023> (2023).
35. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329–337 e324, <https://doi.org/10.1016/j.cels.2019.03.003> (2019).
36. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289–1296, <https://doi.org/10.1038/s41592-019-0619-0> (2019).
37. Ma, Y. & Zhou, X. Spatially informed cell-type deconvolution for spatial transcriptomics. *Nat Biotechnol* **40**, 1349–1359, <https://doi.org/10.1038/s41587-022-01273-7> (2022).
38. NCBI GEO <https://identifiers.org/geo/GSE287278> (2025).

Acknowledgements

All authors gratefully acknowledge the participants who contributed to the study. This work was supported by the National Natural Science Foundation of China (81871134), Beijing Natural Science Foundation (7242062), Beijing Hospitals Authority Youth Programmer (QML20230311), and Wu Jieping Medical Foundation (320.6750.2024-6-69).

Author contributions

P.W. and H.P.L. designed and supervised the research. Y.Z. and J.L.P. performed the sample collection and sequencing, data analysis, and wrote the manuscript. Z.T.L., M.M.Z. and R.J.L. contributed to the sample collection and data analysis. L.S.L. participated in the data analysis. All authors have read the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-025-05995-6>.

Correspondence and requests for materials should be addressed to H.L. or P.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025