human reproduction **ORIGINAL ARTICLE** *Reproductive biology*

Endometrial receptivity revisited: endometrial transcriptome adjusted for tissue cellular heterogeneity

Marina Suhorutshenko[1,2,†], Viktorija Kukushkina[3,†], Agne Velthut-Meikas[1,4], Signe Altmäe[1,5], Maire Peters[1,2], Reedik Mägi[3], Kaarel Krjutškov[1,6], Mariann Koel[1,7], Francisco M. Codoñer[8], Juan Fco. Martinez-Blanch[8], Felipe Vilella[9], Carlos Simón[9,10,11], Andres Salumets[1,2,12,13,*], and Triin Laisk[1,2,3]

[1]Competence Centre on Health Technologies, 50410 Tartu, Estonia [2]Department of Obstetrics and Gynecology, Institute of Clinical Medicine, University of Tartu, 50406 Tartu, Estonia [3]Estonian Genome Center, Institute of Genomics, University of Tartu, 51010 Tartu, Estonia [4]Department of Chemistry and Biotechnology, Tallinn University of Technology, 12618 Tallinn, Estonia [5]Department of Biochemistry and Molecular Biology, Faculty of Sciences, University of Granada, 18071 Granada, Spain [6]Research Program of Molecular Neurology, Research Programs Unit, University of Helsinki, FI-00014 Helsinki, Finland [7]Department of Cell Biology, Institute of Molecular and Cell Biology, University of Tartu, 51010 Tartu, Estonia [8]Lifesequencing SL, 46980 Valencia, Spain [9]Igenomix Foundation/INCLIVA, 46010 Valencia, Spain [10]Research Department, Igenomix SL, 46980 Valencia, Spain [11]Department of Pediatrics, Obstetrics and Gynecology, Valencia University, 46010 Valencia, Spain [12]Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, FI-00014 Helsinki, Finland [13]Department of Biomedicine, Institute of Biomedicine and Translational Medicine, University of Tartu, 50412 Tartu, Estonia

*Correspondence address. Competence Centre on Health Technologies, 50410 Tartu, Estonia. E-mail: andres.salumets@ut.ee

STUDY QUESTION: Does cellular composition of the endometrial biopsy affect the gene expression profile of endometrial whole-tissue samples?

SUMMARY ANSWER: The differences in epithelial and stromal cell proportions in endometrial biopsies modify the whole-tissue gene expression profiles and affect the results of differential expression analyses.

WHAT IS ALREADY KNOWN: Each cell type has its unique gene expression profile. The proportions of epithelial and stromal cells vary in endometrial tissue during the menstrual cycle, along with individual and technical variation due to the method and tools used to obtain the tissue biopsy.

STUDY DESIGN, SIZE, DURATION: Using cell-population specific transcriptome data and computational deconvolution approach, we estimated the epithelial and stromal cell proportions in whole-tissue biopsies taken during early secretory and mid-secretory phases. The estimated cellular proportions were used as covariates in whole-tissue differential gene expression analysis. Endometrial transcriptomes before and after deconvolution were compared and analysed in biological context.

PARTICIPANTS/MATERIAL, SETTING, METHODS: Paired early- and mid-secretory endometrial biopsies were obtained from 35 healthy, regularly cycling, fertile volunteers, aged 23–36 years, and analysed by RNA sequencing. Differential gene expression analysis was performed using two approaches. In one of them, computational deconvolution was applied as an intermediate step to adjust for the proportions of epithelial and stromal cells in the endometrial biopsy. The results were then compared to conventional differential expression analysis. Ten paired endometrial samples were analysed with qPCR to validate the results.

MAIN RESULTS AND THE ROLE OF CHANCE: The estimated average proportions of stromal and epithelial cells in early secretory phase were 65% and 35%, and during mid-secretory phase, 46% and 54%, respectively, correlating well with the results of histological evaluation ($r = 0.88$, $P = 1.1 \times 10^{-6}$). Endometrial tissue transcriptomic analysis showed that approximately 26% of transcripts ($n = 946$) differentially expressed in receptive endometrium in cell-type unadjusted analysis also remain differentially expressed after adjustment for biopsy cellular composition. However, the other 74% ($n = 2645$) become statistically non-significant after adjustment for biopsy cellular composition,

---

†The authors consider that the first two authors should be regarded as joint First Authors.

underlining the impact of tissue heterogeneity on differential expression analysis. The results suggest new mechanisms involved in endometrial maturation, involving genes like *LINC01320*, *SLC8A1* and *GGTA1P*, described for the first time in context of endometrial receptivity.

LARGE-SCALE DATA: The RNA-seq data presented in this study is deposited in the Gene Expression Omnibus database with accession number GSE98386.

LIMITATIONS REASONS FOR CAUTION: Only dominant endometrial cell types were considered in gene expression profile deconvolution; however, other less frequent endometrial cell types also contribute to the whole-tissue gene expression profile.

WIDER IMPLICATIONS OF THE FINDINGS: The better understanding of molecular processes during transition from pre-receptive to receptive endometrium serves to improve the effectiveness and personalization of assisted reproduction protocols. Biopsy cellular composition should be taken into account in future endometrial 'omics' studies, where tissue heterogeneity could potentially influence the results.

STUDY FUNDING/COMPETING INTEREST(S): This study was funded by: Estonian Ministry of Education and Research (grant IUT34-16); Enterprise Estonia (EU48695); the EU-FP7 Eurostars program (NOTED, EU41564); the EU-FP7 Marie Curie Industry-Academia Partnerships and Pathways (SARM, EU324509); Horizon 2020 innovation program (WIDENLIFE, EU692065); MSCA-RISE-2015 project MOMENDO (No 691058) and the Miguel Servet Program Type I of Instituto de Salud Carlos III (CP13/00038); Spanish Ministry of Economy, Industry and Competitiveness (MINECO) and European Regional Development Fund (FEDER): grants RYC-2016-21199 and ENDORE SAF2017-87526. Authors confirm no competing interests.

Key words: endometrial receptivity / endometrial transcriptome / RNA-seq / deconvolution / cell type-specific gene expression / window of implantation

Introduction

The endometrium is a unique self-renewing tissue that undergoes histologically and functionally distinguished phases of growth and differentiation during the menstrual cycle. This is necessary to synchronize the endometrium with oocyte maturation, fertilization and embryo implantation, which takes place in the mid-secretory phase of menstrual cycle, corresponding to the window of implantation (WOI) (Harper, 1992). To understand and characterize the receptive phenotype of endometrium, initially a histological dating method was developed (Noyes et al., 1975). With 'omics' technologies, the understanding of molecular processes behind endometrial receptivity has evolved rapidly (Altmäe et al., 2014). Based on this knowledge, the first gene expression-based clinical endometrial receptivity test was introduced in 2011 (Díaz-Gimeno et al., 2011) and has been successfully applied to detect the most appropriate day for embryo transfer in women undergoing IVF (Ruiz-Alonso et al., 2013). Yet, while previous endometrial transcriptome studies, using either microarray or RNA sequencing (RNA-seq) technology, have identified numerous transcripts differentially expressed during the mid-secretory phase, the overlap between the reported significantly altered genes among studies has remained modest and the exact mechanisms of receptivity are still not fully understood (Gómez et al., 2015).

However, given the phase-dependent changes in tissue cellular composition during the menstrual cycle, the gene expression differences identified in the whole-tissue sample may be partially consequent on these changes eclipsing the true alterations in expression levels. This phenomenon is further exaggerated by inter-sample differences in cellular heterogeneity levels, likely due to biological or technical issues, related to the means used to obtain the tissue biopsy. Therefore, for studies reporting differential expression, it is unclear whether the observed changes in transcript levels reflect variations in cellular subpopulations in the whole-tissue biopsy, altered gene expression in these subpopulations, or a combination of the two. In support of this idea, distinct gene expression patterns of the two dominant endometrial cell types, epithelial and stromal cells, during the WOI were first shown by a study using laser microdissection microscopy to separate endometrial glandular epithelial and stromal cells (Evans et al., 2012), and confirmed by our recent study using transcriptomics of fluorescence-activated sorted cells (Altmäe et al., 2017). However, in addition to a costly and time-consuming cell sorting procedure, cell isolation entails a loss of a more general view on tissue complexity and may affect both the viability and genome activity of isolated cells (Hayman et al., 2006; Strell et al., 2018). Although the above-mentioned issues support the whole-tissue transcriptome analysis, ignoring the cellular heterogeneity in whole-tissue analysis may cause misinterpretation of the genomic data, particularly for highly heterogenic tissues (Shen-Orr and Gaujoux, 2013). As an alternative to cell-type isolation techniques, computational decomposition (or deconvolution) of gene expression profiles from heterogeneous tissue has been proposed to adjust for differences stemming from variation in biopsy cellular composition, and has been successfully used for gene expression profiling of blood, lymphoid and tumour tissues (Li et al., 2017; Pan et al., 2017; Schelker et al., 2017; Chen et al., 2017b).

In the current study, we utilized a computational approach to estimate cell-type fractions in biopsies and to deconvolute the endometrial whole-tissue gene expression profile. This approach helps to adjust for expression changes stemming from altered cellular composition of biopsied samples taken at different phases of the menstrual cycle, and to reveal the true gene expression changes irrespective of the cellular heterogeneity of the tissue analysed. The current study is the first using gene expression profiling adjusted for cellular heterogeneity to uncover the mechanisms of endometrial receptivity that otherwise would remain hidden in whole-tissue genomic analysis.

Materials and Methods

The study was approved by the Research Ethics Committee of the University of Tartu, Estonia (No 221/M-31) and Ethical Clinical Research

Committee of IVI Clinic, Valencia, Spain (No 1201-C-094-CS). Informed consent was signed by all women who entered the study.

Study participants

For RNA-seq, the endometrial tissue samples were obtained in Estonia and Spain from fertile-aged women, with normal body mass index (BMI) and self-reported regular menstrual cycles (Estonia n = 20, Spain n = 15; a total of 35 women), who had not received hormonal treatments for at least three months prior to the time of biopsy. The average age and BMI were 29.1 ± 3.6 years and 23.2 ± 2.9 kg/m^2 for Spanish women, and 30.2 ± 3.4 years and 23.2 ± 4.5 kg/m^2 for Estonian women, respectively. The women had at least one live-born child and no previous infertility records. As it has been shown that endometrial transcriptomic profile is not affected by a previous biopsy taken in the same menstrual cycle (Miravet et al., 2017), two endometrial biopsies were obtained using Pipelle catheter (Laboratoire CCD, France) within the same natural menstrual cycle; the first during the early secretory endometrial phase (ESE; two days after the luteinizing hormone surge [LH + 2] in Spain and LH + 1 to LH + 3 in Estonia) and the second sample during the mid-secretory endometrial phase (MSE; day LH + 7 in Spain and day LH + 7 to LH + 9 in Estonia). Additionally, 10 paired endometrial samples from healthy women (age 30.1 ± 3.1 years and BMI 24.1 ± 4.8 kg/m^2, non-smokers with proven fertility) from Estonia were used for qPCR validation of RNA-seq results. Endometrial sample timing was confirmed by histological examination according to Noyes' criteria (Noyes et al., 1975). The samples were stored at −80°C until use.

Total RNA extraction from endometrial tissue

For the total RNA extraction, up to 30 mg of the tissue was homogenized in the presence of QIAzol reagent (Qiagen, Germany) and mechanical lysis (0.5 mm diameter steel bead, TissueLyser LT (20–40 s)), and processed using miRNeasy Mini kit (Qiagen), following the manufacturer's protocol. miRNA content was decreased using Qiagen's MinElute protocol. DNase I treatment was performed on column using RNase-Free DNase Set (Qiagen). Purified RNA integrity number (RIN) and quantity were determined with Bioanalyzer 2100 RNA Nano 6000 kit (Agilent Technologies, USA).

mRNA sequencing

The endometrial total RNA samples with a concentration at least 200 ng/µl and RIN >8.0 were used for cDNA synthesis and library construction. Libraries were generated from 4 µg of total RNA using Illumina TruSeq Stranded technology (Illumina, USA), following the manufacturer's protocol. Libraries were normalized, pooled and sequenced using HiSeq™ 2000/2500 with a configuration of 100 cycles paired-end reads in the three sample groups: Estonian Est 1 cohort samples at Estonian Genome Center Core Facility (Tartu, Estonia), Est 2 cohort samples at Institute for Molecular Medicine Finland (FIMM, Helsinki, Finland) and Spanish samples at Lifesequencing S.L. (Valencia, Spain).

RNA sequencing analysis and data preparation

The processing of raw RNA reads, that included quality control, adapter removing, trimming and mapping to human genome hg19 was performed as previously described (Altmäe et al., 2017). The RNA-seq data presented in this study is deposited in the Gene Expression Omnibus database with accession number GSE98386.

Transcriptome profiling of early secretory and mid-secretory endometrial samples

To rule out population- and/or batch-specific effects, differential expression (DE) analysis was carried out separately in three groups according to sequencing centre: Est 1, Est 2 and Spain. For group-level DE analyses, the edgeR software was used. Before DE analysis, transcripts with low or uneven expression patterns were filtered out, retaining transcripts with counts per million (CPM) ≥2 in at least half of the samples (±5%) for subsequent analysis. We used a paired analysis design, which takes into account the paired nature of the biopsies collected from each woman and thus increases statistical power, while at the same time reduces the effect of potential confounding factors.

Thereafter, results from all three participating groups were meta-analysed using the METAL tool (Willer et al., 2010) to identify transcripts that exhibit similar DE patterns in different groups. Genes with consistent effect direction over three groups with p-value lower than Bonferroni corrected p-value threshold (0.05 divided by the number of transcripts included in the meta-analysis) were considered statistically significantly differentially expressed.

Deconvolution analysis

For the deconvolution of endometrial whole-tissue gene expression profiles, single-cell tagged reverse transcription (STRT)-based RNA-seq data for bulk-RNA from endometrial stromal and epithelial cells was used. STRT data for both Estonian and Spanish samples was used to deconvolute the gene expression profiles of samples from respective groups. In totally, we used pre-analysed RNA-seq data for 19 stromal and 17 epithelial fractions from ESE and 24 stromal and 24 epithelial fractions from MSE. The detailed information on participants and sampling is presented in our previous study (Altmäe et al., 2017). Briefly, the cell populations were extracted by fluorescence-activated cell sorting (FACS) from endometrial biopsies of healthy women, and analysed using the STRTprep pipeline as previously described (Krjutškov et al., 2016) (Fig. 1). As the pipeline uses hg19 reference genome for mapping, and a different version of annotation file for aligning/counting, only the .bam files from the pipeline were used for further steps. The mapped reads were aligned on Ensembl v.75 annotation file, the same file used for whole-tissue RNA-seq data, and raw counts were obtained with HTSeq package script htseq-count v.0.6.1 (Anders et al., 2015). The normalized CPMs were obtained using R package edgeR v. 3.6.2 (Robinson et al., 2010) and the default normalization function instead the spike-in normalization.

For deconvolution, we used the R package DeconRNASeq v.1.10.0 (Gong and Szustakowski, 2013), which uses tissue and pure cell-type specific expression patterns to decompose mixed tissue to cell types. The package adopts a globally optimized non-negative decomposition algorithm through quadratic programming to estimate cell-type proportions in the heterogeneous tissue. As the input, it takes normalized expression levels from pure cell fractions for some set of transcripts and the mixed tissue normalized expression levels. Both the cell fraction-specific and mixed expression values should be in same units and normalized using the same method. In this study, normalized CPMs of all non-zero expressed transcripts from stroma and epithelium STRT-sequencing, and from endometrial biopsies were used as the input, and the median value of cell-specific expression per cell-type over sample group was used to estimate the proportions of stromal and epithelial cells in the heterogenous endometrial tissue (Fig. 1).

The stromal and epithelial cell proportions obtained from deconvolution for ESE and MSE samples were used in downstream DE analyses. The DE analyses were done the same way as described above, except this time cell proportions were used as covariates (Fig. 1). As the stroma and epithelium
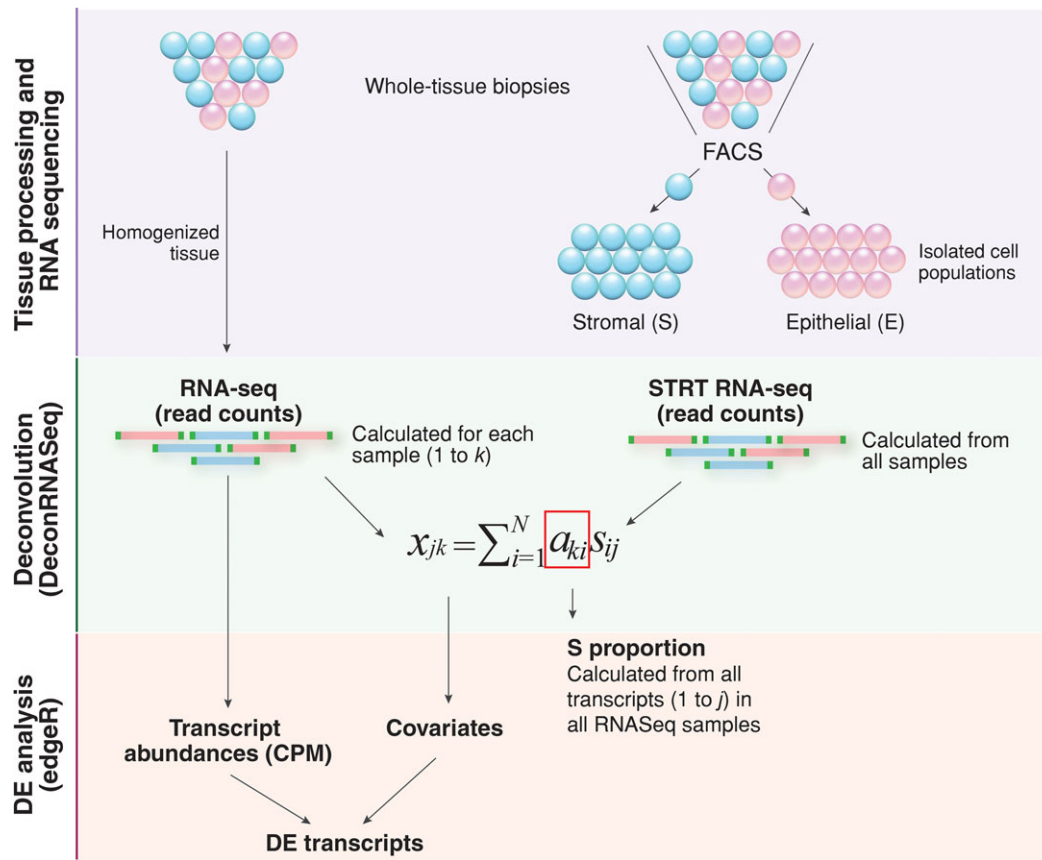
Figure 1 A simplified overview of the gene expression profile computational deconvolution of endometrial tissue. Endometrial biopsies were homogenized and total RNA extracted from whole-tissue samples was sequenced. In parallel, two cell types from endometrial biopsies were separated; epithelial cells were labelled with fluorochrome-conjugated mouse anti-human CD9 monoclonal antibody and stromal cells were simultaneously labelled with fluorochrome-conjugated mouse anti-human CD13 monoclonal antibody, followed by flow cytometric analysis and FACS cell sorting. Bulk-RNA full transcriptome analysis of FACS-sorted endometrial epithelial and stromal cells was performed with the RNA-seq method, following the single-cell tagged reverse transcription (STRT) protocol with modifications. Reference proportions of epithelium (E) and stroma (S) were calculated for each sample using DeconRNASeq algorithm. By using transcript measures from separate cell-types (FACS-sorted epithelial and stromal cell populations) and transcript measures from whole-tissue samples, the algorithm estimates the cell-type proportions over samples by fitting a non-negative least squares equation for each transcript. Further, it calculates the theoretical proportions of epithelium and stroma across the whole transcriptome. The expression level x_{jk} of gene j in a sample k is the average of expected expression level s across the cell types s_{ij}, weighted by the respective cell-type proportions a_{ki} ($i = 1 \ldots N$, N: the total number of cell types) (Adapted from (Gong and Szustakowski, 2013)). Average stroma proportions calculated from all whole-tissue samples used for RNAseq were then applied as covariates in differential expression (DE) analysis using edgeR.

proportions together add up to 1.0, only the stroma proportions were used as covariates. The meta-analysis for groups was done as described above.

Correlation between estimated cellular proportions and histological dating

To test the accuracy of the estimated proportions of stromal and epithelial cells resulting from deconvolution, we correlated the epithelial fraction estimates with estimated endometrial cycle day (corresponding to a 28-day cycle) from histological analysis. For this, we used the histological evaluation data for 18 paired samples (nine from ESE and nine from MSE phase) for which detailed dating information was available. If the endometrial cycle date had been given as a range (e.g. days 21–23), the arithmetic mean was used for calculating the Pearson correlation coefficient.

Additionally, we assessed the proportion of epithelial and stromal tissue in the two time-points in histological images, and compared the numbers to those obtained from deconvolution analysis. The histological specimens were prepared from formalin-fixed paraffin-embedded endometrial tissue sections using standard hematoxylin and eosin staining protocol (Fischer et al., 2008). The proportions of epithelial and stromal cells in the specimen were robustly calculated by measuring the areas occupied by either cell type in histological images using Adobe Photoshop CC 2017. The size of the selected areas was converted to pixels, and the proportion of epithelial cells was calculated by dividing the area under epithelial cells (in pixels) to that of the entire specimen. The analysis included a total of 13 specimens (eight ESE and five MSE), since for some individuals, more than one image was available for evaluation. One to three optical fields were analysed per each sample and, in case of multiple measurements, the average proportions were calculated across evaluations.

Functional annotation of differentially expressed genes

Functional annotation was performed with Ingenuity Pathway Analysis (IPA) tool using all significant differentially expressed genes (DEGs) identified after deconvolution. Pathway analysis was performed using IPA with default settings, and pathways with a Benjamini-Hochberg p-value <0.05 were considered as significantly enriched. Networks of significantly DEGs were then algorithmically generated based on their connectivity.

Endometrial receptivity markers

To outline potential biomarker candidates that exhibit true expression change in MSE and could thus be detected by routine transcriptome analysis of whole-tissue biopsy, meta-analysis DEG lists identified before and after deconvolution analysis ($|log_2FC| \geq 1$, outranked Bonferroni p-value) were compared and overlapping transcripts were selected as marker candidates.

Further, our marker candidate list was compared with DEG lists reported in previous RNA-seq and microarray studies on human endometrium (Carson *et al.*, 2002; Kao *et al.*, 2002; Borthwick *et al.*, 2003; Riesewijk *et al.*, 2003; Mirkin *et al.*, 2005; Talbi *et al.*, 2006; Altmäe *et al.*, 2010; Tseng *et al.*, 2010; Díaz-Gimeno *et al.*, 2011; Hu *et al.*, 2014; Sigurgeirsson *et al.*, 2016). Genes previously identified as differentially expressed between MSE and ESE were referred to as 'known marker candidates', including transcriptomic markers commonly used in endometrial testing (Díaz-Gimeno *et al.*, 2011; Enciso *et al.*, 2018), and others were recognized as 'novel'.

In addition, the list of transcripts, significant in the unadjusted transcriptome and non-significant after deconvolution, were compared with gene expression profiles of isolated epithelial and stromal cells, obtained from STRT RNA-seq (original data in Altmäe *et al.*, 2017). Genes that exhibited no significant expression change on the cell-population level were considered as genes whose expression change in whole tissue is derived from cell-type proportion fluctuations.

Quantitative real-time PCR analysis

For technical validation of the results, reverse transcriptase quantitative PCR (RT-qPCR) was applied on selected endometrial marker candidates, using ten paired ESE and MSE total RNA samples from healthy volunteers different from these used in the RNA-seq study. The selection criteria for RT-qPCR validation were expression change $|log_2FC| \geq 1.5$ in all groups and read count $log_2CPM \geq 4$, based on our previous expertise.

For RT-qPCR, 1 μg of total RNA was reverse-transcribed using random hexamer primer (Thermo Scientific, USA) and cDNA was amplified and quantified using 7500 Fast Real-Time PCR System (Applied Biosystems, USA), using HOT FIREPol EvaGreen qPCR Mix Plus (Solis BioDyne, Estonia) and 250 nM of sense and antisense primers specific for nucleotide sequence of selected marker candidates or housekeeping genes, *TBP* and *SDHA*. Data were analysed using 7500 Software v2.0.5 (Applied Biosystems). Differences in gene expression levels were estimated using comparative Ct ($2^{-\Delta\Delta Ct}$) method (Livak and Schmittgen, 2001). The Wilcoxon paired sample test was applied to estimate significance level. The oligonucleotide primers are presented in Supplementary Table S1.

Results

Differential expression analysis between mid-secretory and early secretory endometrium

In the present study, we used RNA-seq to analyse gene expression of endometrial samples; 70 individual cDNA libraries (35 ESE and 35 MSE samples) were sequenced and an average of 54 million read pairs were generated per library. The summary of RNA-seq results are provided in Supplementary Table SII. After filtering out low-quality reads from the datasets, an average of approximately 90% of the qualified read pairs across all groups were mapped to human genome version 19. A schematic overview of the analysis and results is presented in Fig. 2.

The total number of expressed transcripts ranged from 21,890 to 30,576 per studied sample (Supplementary Table SII). DEGs within three groups were detected using paired sample design, followed by meta-analysis of all expressed genes with METAL software. This resulted in total of 3591 genes (1800 up-regulated and 1791 down-regulated) that were identified as statistically significantly differentially expressed in MSE (Bonferroni p-value $<2.79 \times 10^{-6}$) and showed the same direction in expression change in all three groups, as presented in Supplementary Table SIII. The average logarithmic expression fold change (Log_2FC) of genes expressed in MSE ranged between 5.3 and 8.8 in both directions (Supplementary Table SIII). The genes with the highest expression rate change were *PAEP*, *C4BPA*, *C2CD4A*, *CXCL14*, *GPX3*, *SLC15A1*, *MT1G*, *HAP1*, *MFSD4A* and *IRX3*. The most significantly differentially expressed genes were *SLC15A1*, *SLC1A1*, *SNX10*, *DKK1*, *GADD45A*, *AOX1*, *GPX3*, *DPP4*, *HEY2* and *MYOCD*.

Expression profile deconvolution of endometrial tissue

To clarify whether the observed changes in gene expression levels might actually stem from fluctuations in biopsy cellular composition, we used gene expression data for isolated endometrial epithelial and stromal cells to perform endometrial whole-tissue gene expression profile deconvolution. As a first step of the analysis, gene expression profiles of homogenous cell populations derived from ESE and MSE were compared with whole-tissue transcriptomes to estimate the relative proportions of these cellular subpopulations in the whole-tissue biopsy. This analysis revealed that the stromal fraction was larger in ESE biopsies, while in MSE biopsies the epithelial fraction was slightly larger, with considerable inter-individual variation (Fig. 3A). On average, ESE samples consisted of 65% stromal cells, while the average stromal fraction in MSE samples was 46%. We observed a good correlation between gene-expression-estimated epithelial cell proportions and endometrial cycle date according to histological evaluation ($r = 0.88$ (95% CI 0.71–0.96), $P = 1.1 \times 10^{-6}$), confirming that epithelial fraction becomes more dominant as the cycle progresses (Fig. 3B).

As a second step, the epithelial and stromal cells' proportions estimated by computational analysis of gene expression profiles were further validated by direct histological evaluation using endometrial tissue samples (Fig. 3C). The histology results suggested that the proportion of epithelial cells in ESE is around 30%, increasing to approximately 50% by the time of WOI, and this complies with the proportions calculated based on cell-type specific gene expression (35% and 54% for epithelial cells, respectively) (Fig. 3D).

Transcripts differentially expressed in mid-secretory endometrium

The proportions obtained from deconvolution were then used as cofactors in differential expression analysis to account for differences in cellular composition. After meta-analysis of deconvoluted gene expression profiles from different groups, 1211 (679 up- and 532 down-regulated) transcripts showed significant expression change in
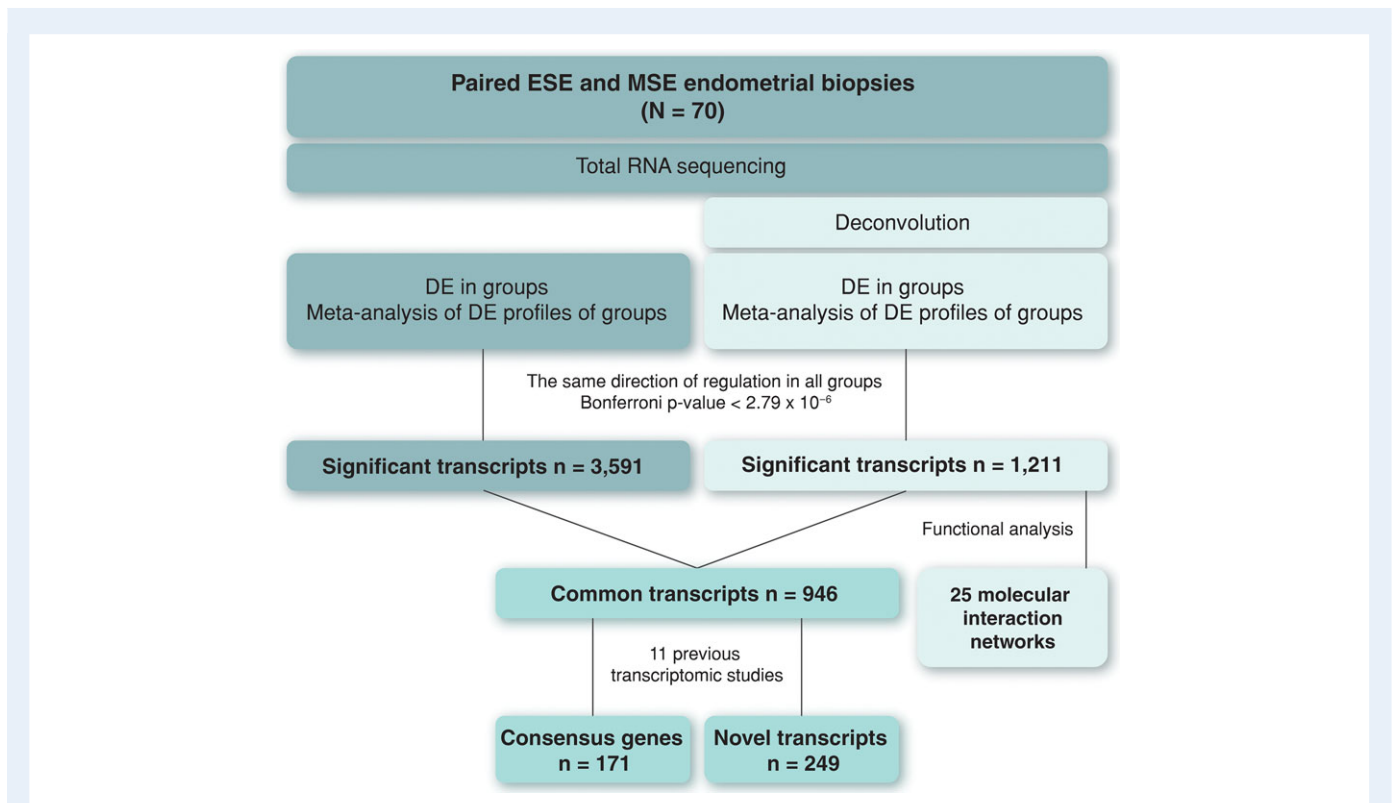
Figure 2 Overview of the endometrial RNA-seq data analysis pipeline and results. Sequenced endometrial total RNA samples were processed using two parallel approaches: (i) differential expression analysis followed by meta-analysis of results from different groups and (ii) deconvolution followed by DE and meta-analysis. Functional analysis was performed on all significant genes using the deconvolution-based approach. Significant transcripts identified by the two approaches were compared resulting in a set of common transcripts. ESE—early secretory endometrium; MSE—mid-secretory endometrium; *N*—number of analysed samples; *n*—numbers of DE transcripts identified during each step of the analysis; DE—differential expression; Consensus genes—genes reported in 11 selected receptivity-related microarray and RNA-seq studies (please, see main text for the references).

MSE endometrium, with Log$_2$FC range between −5.3 and 7.9 (Supplementary Table S4). The genes with the highest expression rate change were *PAEP*, *HAP1*, *GPX3*, *CXCL14*, *C4BPA*, *SLC15A1*, *IRX3*, *MMP26*, *C2CD4A* and *SULT1C2P*. The most significantly differentially expressed genes were *SLC15A1*, *SLC1A1*, *SNX10*, *DKK1*, *GADD45A*, *AOX1*, *GPX3*, *DPP4*, *HEY2* and *MYOCD*.

As a result of RT-qPCR-based validation, *MYOCD*, *LINC01320*, *SLC8A1*, *TRPC4* and *GGATP1* showed significant expression rate change in MSE samples (Fig. 4).

Further, all 1211 transcripts that exhibited significant expression change in MSE after adjusting for differences in cellular composition were subject to functional annotation in order to investigate their role in endometrial maturation processes. DE genes with the highest significance level (*n* = 90) and expression fold change (*n* = 90) are depicted in Fig. 5 (Supplementary Table SIV).
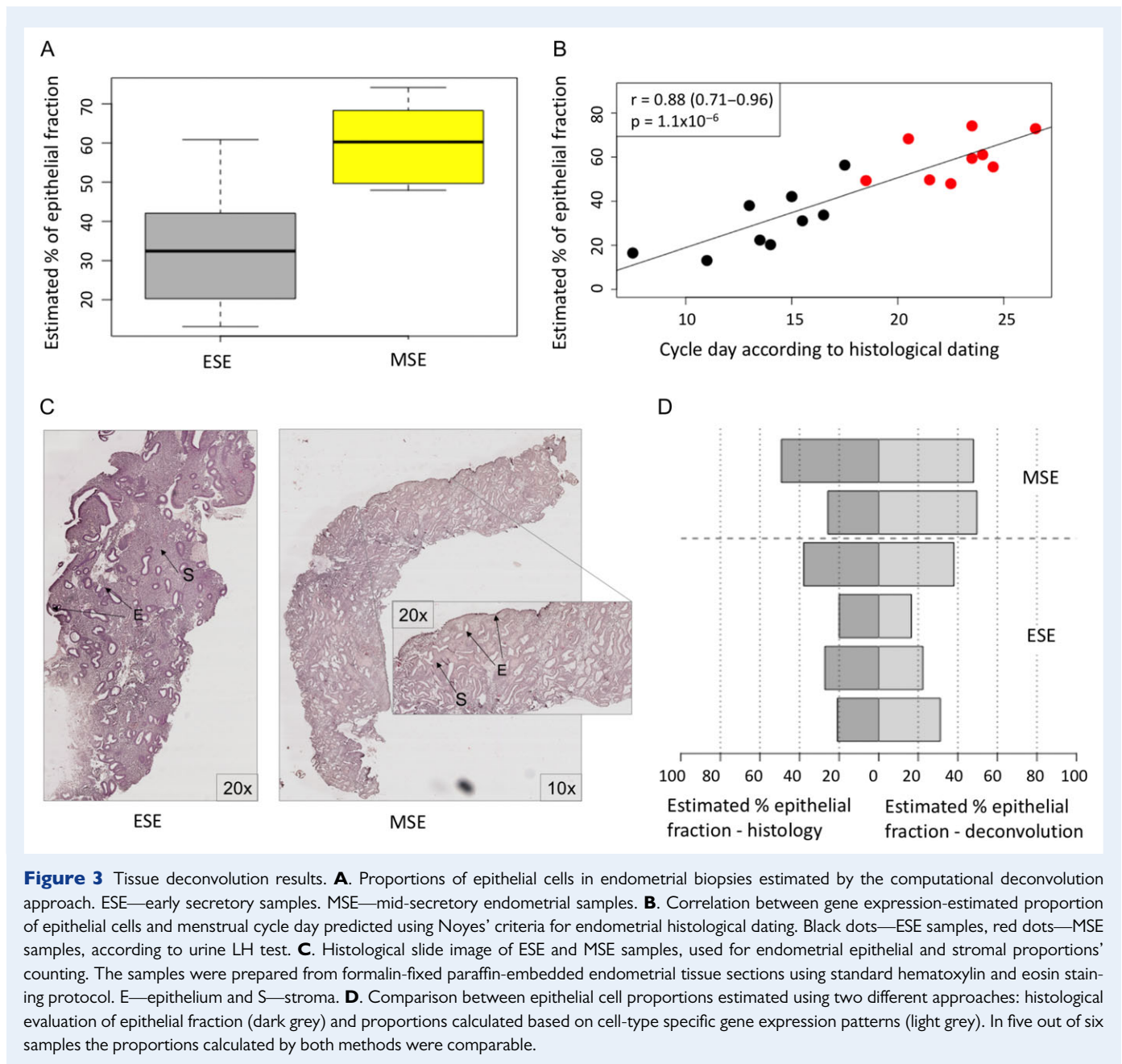
Pathway mapping with IPA indicated that WOI DEGs are involved in many basic cellular functions and processes and, in particular, cell fate, cellular movement, cell-to-cell signalling, cellular growth and development, inflammatory and vascular development processes (Supplementary Table SV). In total, the IPA tool identified 25 Ingenuity Knowledge Database pathways significantly enriched in the MSE transcriptome relative to ESE (Benjamini-Hochberg FDR < 0.05), highlighting the major regulative role of NFkB (Supplementary Figure S1A), interferon α (Supplementary Figure S1B), steroid hormone activity

(Supplementary Figure S1C) and MAPK/ERK signalling (Supplementary Figure S1D–F) in human endometrium. All molecular interaction networks with a score at least 21 (Krämer *et al.*, 2014) are presented in Supplementary Figure S1A–K.

Marker candidate genes for endometrial receptivity

When significant DEG lists before (3591 transcripts) and after (1211) deconvolution analysis were compared, 946 transcripts came out significant in both approaches (Fig. 6, Supplementary Tables SIII and SIV) and therefore were considered as genes with the true expression change in MSE compared to ESE. There were 2645 transcripts, which are almost 74% from all DE transcripts identified without deconvolution, not considered significantly differentially expressed after adjustment for tissue cellular composition (Fig. 6; Supplementary Table SIII). Of these, the genes with the highest expression change between MSE and ESE in whole tissue (Log$_2$FC > 3), yet not differentially expressed after deconvolution, are presented in the Table I. Moreover, 265 transcripts came out significant only after deconvolution was applied (Fig. 6; Supplementary Table SIV).

Further, we compared our results with those from eleven publications that have reported genes differentially expressed during WOI in healthy women, including nine microarray (Carson *et al.*, 2002; Kao

Figure 3 Tissue deconvolution results. **A**. Proportions of epithelial cells in endometrial biopsies estimated by the computational deconvolution approach. ESE—early secretory samples. MSE—mid-secretory endometrial samples. **B**. Correlation between gene expression-estimated proportion of epithelial cells and menstrual cycle day predicted using Noyes' criteria for endometrial histological dating. Black dots—ESE samples, red dots—MSE samples, according to urine LH test. **C**. Histological slide image of ESE and MSE samples, used for endometrial epithelial and stromal proportions' counting. The samples were prepared from formalin-fixed paraffin-embedded endometrial tissue sections using standard hematoxylin and eosin staining protocol. E—epithelium and S—stroma. **D**. Comparison between epithelial cell proportions estimated using two different approaches: histological evaluation of epithelial fraction (dark grey) and proportions calculated based on cell-type specific gene expression patterns (light grey). In five out of six samples the proportions calculated by both methods were comparable.

et al., 2002; Borthwick *et al.*, 2003; Riesewijk *et al.*, 2003; Mirkin *et al.*, 2005; Talbi *et al.*, 2006; Altmäe *et al.*, 2010; Tseng *et al.*, 2010; Díaz-Gimeno *et al.*, 2011) and two RNA-seq studies (Hu *et al.*, 2014; Sigurgeirsson *et al.*, 2016). Publications were selected based on the availability and quality of gene expression datasets. As a result, 171 transcripts overlapped among both microarray and RNA-seq studies (Supplementary Table SVIA), 526 transcripts were found in either microarray or RNA-seq study, while 249 transcripts were recognized as novel or previously not reported in endometrial WOI-related transcriptome (Supplementary Table SVIB). The top significant novel transcripts are presented in the Table II.

In addition to protein coding genes, possible regulatory transcripts, e.g. non-coding RNAs, were identified. The list of significant DE long intergenic non-coding RNAs (lincRNAs) is presented in Table III.

Discussion

One of the main objectives of our study was to investigate how histological changes during the menstrual cycle affect the detection of gene expression profiles of heterogeneous endometrial tissue. While the histological changes that distinguish MSE from ESE are well described (Noyes *et al.*, 1975; Crum *et al.*, 2003), to the best of our knowledge, we are the first to show how epithelial and stromal fractions are actually altered between the pre-receptive and receptive state, by using a computational deconvolution approach.

The morphological criteria used for endometrial dating have major limitations for predicting endometrial receptivity (Coutifaris *et al.*, 2004; Murray *et al.*, 2004), therefore endometrial gene expression analysis has become the new golden standard to predict the WOI. Since
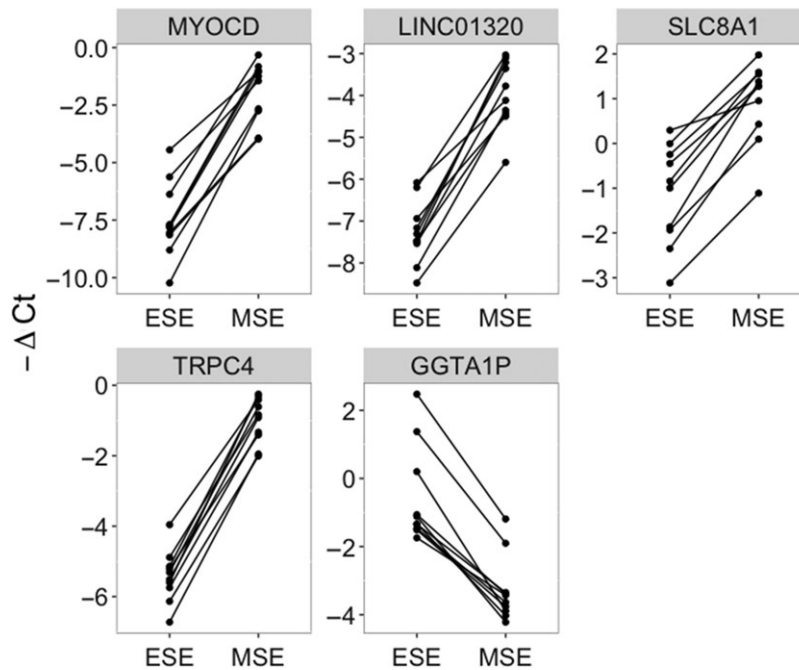
Figure 4 RT-qPCR validation of five marker candidates, using ten paired early secretory (ESE) and mid-secretory (MSE) endometrial samples. The dots connected with a solid line represent samples from the same patient. Significant: paired Wilcoxon test P-value <0.01.

endometrial biopsies contain a range of cell types in varying proportions, due to both biological and technical reasons, some of the observed changes in gene expression levels might actually stem from differences in biopsy cellular composition. Although this issue has been discussed before (Edgell et al., 2013; Altmäe et al., 2014), it has never been experimentally addressed in previous endometrial transcriptome studies. As a substantial improvement compared to previous endometrial gene expression studies, we applied gene expression profile deconvolution to account for biopsy cellular composition that could artificially alter the observed expression profile of a whole-tissue biopsy. Only 26% of the transcripts identified by meta-analysis without prior adjustment for cellular heterogeneity remained significant after deconvolution analysis was applied (Fig. 6, Supplementary Table SIII and SIV), while the remaining 74% rather reflect the variation in tissue cellular composition between the two compared time points. Among the latter category are much studied receptivity-associated genes: *CLDN4*, *FOS*, *CCND2*, *AR*, *VC*, *CXCL13*, *PENK*, *OPRK1*, *KIF4A*, *POSTN*, *OLFM4*, *PRC1*, *KMO*, *AMIGO2*, *ECM1*, *RASSF2*, *KIF20A* and *LRP4* (Díaz-Gimeno et al., 2011; Altmäe et al., 2017). For example, olfactomedin 4 (*OLFM4*; Table I) has been shown to be significantly down-regulated in MSE samples (Díaz-Gimeno et al., 2011) but shows no significant expression change between MSE and ESE in either epithelial or stromal cells when evaluated separately (Dassen et al., 2010; Kodithuwakku et al., 2011).

Further, we described DE transcripts that were significantly differentially expressed only after deconvolution analysis was applied, but were non-significant before deconvolution (n = 265, Fig. 6). The expression change of these genes in whole-tissue biopsy may be eclipsed by fluctuations in cell-type proportions in endometrial biopsies, potentially leading to underestimation of processes and pathways involved in endometrial maturation. Among these transcripts we found

genes that have not been previously reported to be related to WOI. For example, cell migration-associated protein GLIPR2 and fibroblast growth factor 2 (FGF2) are believed to be involved in extracellular signal-regulated kinase (ERK) signalling (Paiva et al., 2011; Huang et al., 2013), one of the key pathways in endometrial gene expression regulation (Supplementary Table SV; Figure S1D) (Fluhr et al., 2013). Inflammatory response-associated gap-junction alpha-1 (GJA1) (Supplementary Figure S1D), also found significantly differentially expressed only after deconvolution analysis, is differentially expressed in stromal cells during MSE, and promotes stromal cell differentiation and maintenance of the decidualized endometrial phenotype (Yu et al., 2011, 2014). Although, to the best of our knowledge, these genes have not yet been reported in human MSE transcriptome studies previously, their aberrant expression has been shown in endometrial studies focusing on implantation failure and insufficient endometrial growth (Tapia et al., 2008; Maekawa et al., 2017), supporting their involvement in endometrial processes. It is important to outline that in our study these genes did not show significant expression change in whole-tissue samples, suggesting that dynamic epithelial and stromal cell proportions may also be the reason they have not been reported before in whole-tissue transcriptomic studies.

Our information about endometrial receptivity-associated gene expression keeps improving. Thousands of genes have been suggested to take part in the processes of endometrial maturation, with respect to their WOI-marker potential. Previous findings summarized in two recent publications have outlined that the WOI could be characterized by a relatively small number of differentially expressed genes (Altmäe et al., 2017; Enciso et al., 2018). In the present study, we also report a set of 1211 transcripts differentially expressed between MSE and ESE after deconvolution (Supplementary Table SIV). Figure 5 summarizes
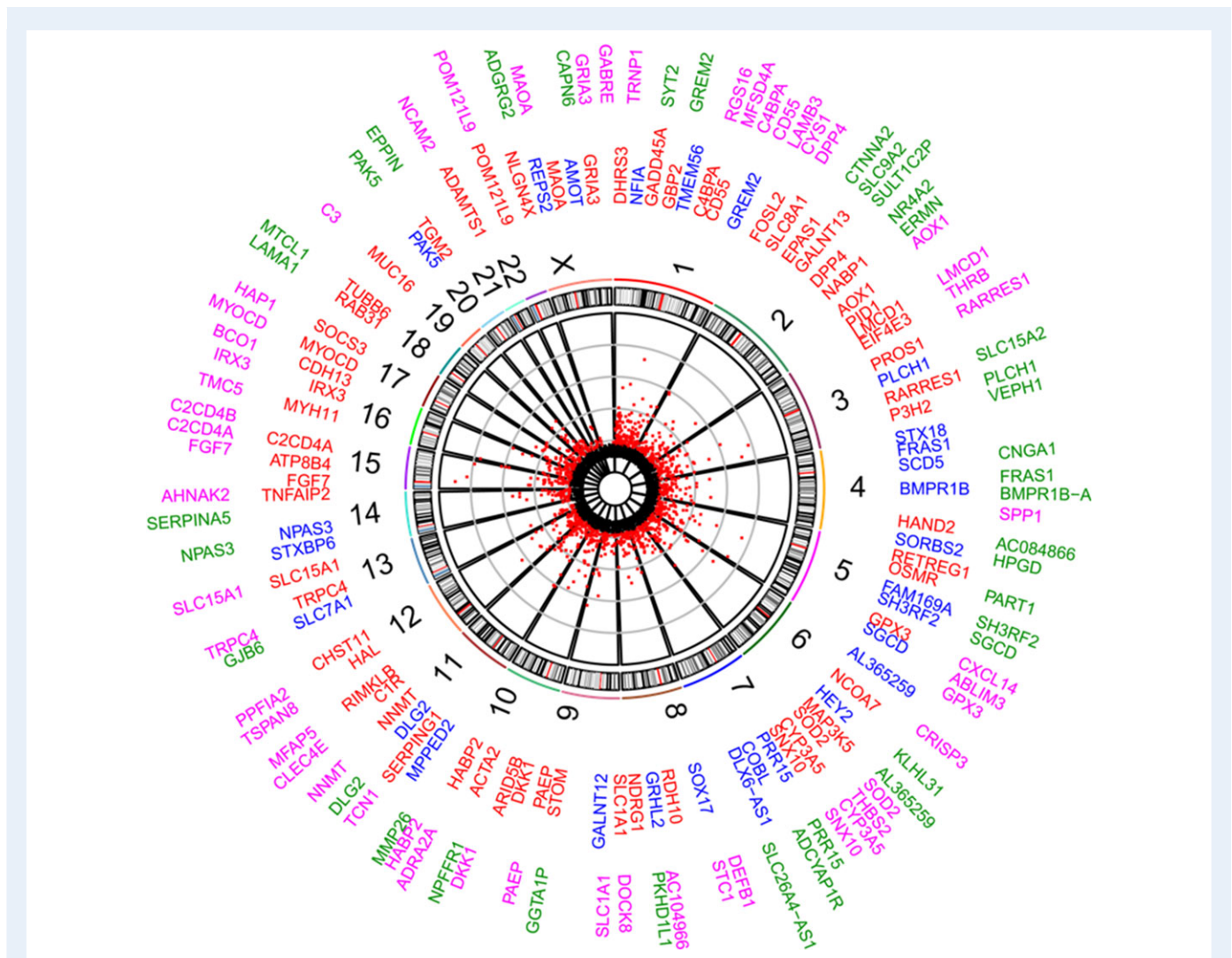
Figure 5 Differentially expressed genes (DEGs) between early secretory (ESE) and mid-secretory (MSE) endometrial samples. Scatterplot represents DEGs significant after deconvolution (Bonferroni *P*-value <2.79 × 10^{-6}). Numbers represent chromosomes. The inner circle gene names contain top significant DEGs (red—up-regulated and blue—down-regulated in MSE, respectively), while the outer circle contains DEGs with the highest expression rate fold change between ESE and MSE (pink—up-regulated and green—down-regulated in MSE, respectively).

the set of transcripts, significantly differentially expressed in MSE after endometrial gene expression profile was adjusted for epithelial and stromal fractions, by extracting the transcripts with the highest expression rate change (outer gene circle, *n* = 90) and significance level (inner gene circle, *n* = 90). These top genes were also among the 946 potential endometrial receptivity markers that exhibited significant expression change in whole-tissue transcriptome analysis. Moreover, 31 genes (*C4BPA, CD55, DPP4, AOX1, LMCD1, PLCH1, FRAS1, BMPR1B, SH3RF2, GPX3, SOD2, CYP2A5, SNX10, PRR15, SLC1A1, PAEP, DKK1, HABP2, DLG2, NNMT, TRPC4, SLC15A1, NPAS3, FGF7, C2CD4A, IRX3, MYOCD, PAK5, POM121L9, MAOA* and *GRIA3*) were located in both circles, which suggests their great biomarker potential. Twelve of these genes, *PAEP, GPX3, C4BPA, DPP4, DKK1, CD55, NNMT, MAOA, HABP2, SOD2, AOX1* and *SLC1A1*, are already applied endometrial receptivity markers (Díaz-Gimeno *et al.*, 2011; Altmäe *et al.*, 2017; Enciso *et al.*, 2018), while others have been previously reported in endometrial receptivity-related transcriptome studies (Altmäe *et al.*,

2012; Hu *et al.*, 2014; Sigurgeirsson *et al.*, 2016; Chen *et al.*, 2017a). Still, many of the genes identified in our study (Supplementary Table SIV) have not been reported in MSE transcriptome studies before. For example, solute carrier family member, *SLC8A1*, one of the top differentially expressed genes, is recruited in ERK1/2 activation and can modulate immunoglobulin-mediated immune response through ERK1/2 or through interaction with major histocompatibility complex (MHC) class I proteins (Supplementary Table SV and SVIB, Supplementary Figure SI D).

Our results also support the previously discussed important role of immune responses in the pre- and peri-implantation period (Altmäe *et al.*, 2010; Gnainsky *et al.*, 2014; Haller-Kikkatalo *et al.*, 2014; Altmäe *et al.*, 2017). For example, type I interferon-activated proteins caspase 1 (CASP1), tumour suppressor protein SAMD9L and phospholipid scramblase 1 (PLSCR1) induce the innate immune response and promote endometrial antiviral status (Catalano *et al.*, 2007; Boon *et al.*, 2014; Wang *et al.*, 2016) (Supplementary Figure SVI B). We also noted

that the expression change of these three genes in MSE before deconvolution was under 2-fold (Supplementary Table SIII), which may be the reason why these genes have never been reported in studies using at least a 2-fold expression change threshold. The 171 DE genes, overlapping between our results and previously published RNA-seq and

Figure 6 Numbers of significantly differentially expressed transcripts identified from meta-analysis of transcriptomic profiles of endometrial tissue samples from different groups. A. Transcripts identified in whole endometrial tissue (before deconvolution, BD; n = 3591). B. Transcripts identified in cellular heterogeneity-adjusted expression profiles (after deconvolution, AD; n = 1211). C. Transcripts significant only in whole-tissue transcriptome and non-significant after deconvolution (blue, n = 2645) were considered genes whose expression change in whole-tissue may be derived from variation in cell types' proportions. D. Differentially expressed transcripts significant only after tissue deconvolution (pink, n = 265) were considered genes whose expression change may be eclipsed in mid-secretory endometrium due to fluctuations in cell-type proportions, leading to underestimation of molecular processes and pathways involved in endometrial maturation. E. Transcripts significant in both approaches (purple, n = 946) are considered potential WOI markers.

array studies, can be considered as genes with the highest endometrial receptivity marker potential, while novel marker candidates reported in our study need further investigation in order to understand how they are involved in endometrial maturation (Supplementary Table SVI A and B). It is also interesting that 126 of 238 initial endometrial receptivity array (ERA) genes (Díaz-Gimeno et al., 2011) exhibited significant expression change after adjustment for biopsy cellular composition, while for the remaining 112 ERA genes, the expression changes may arise from uneven cellular content in biopsies or from the dramatic transcriptional changes in minor cell types not analysed in our study (e.g. immune cells).

With the implementation of transcriptomic tools in endometrial analysis, it has become clear that not only protein coding genes are important for endometrial receptivity. LincRNAs have significant regulative potential on transcriptional and translational levels through direct interactions with different types of RNAs and proteins (Kung et al., 2013; Groen et al., 2014). Recently, Sigurgeirsson and colleagues showed the potential role of lincRNAs in endometrial receptivity (Sigurgeirsson et al., 2016). We report eight lincRNAs significantly up- or down-regulated in MSE (Table III). Our results demonstrate that LINC01320 in particular has a strong effect in all studied groups (Table II, Supplementary Table SIII and SIV), confirmed by RT-qPCR (Fig. 4). Previously, differential expression of LINC01320 was reported in endometrial and ovarian cancer cells (Chen et al., 2017a), although, to the best of our knowledge, it has never been associated with endometrial receptivity before.

Certain limitations of our study also need to be acknowledged. For gene expression profile deconvolution, only stromal and epithelial cell fractions were considered, whereas other cells, such as endothelial and immune cells, or even cell types we may not be aware of, also could contribute to the gene expression profile. Although there is no available data on gene expression profiles of minor endometrial cell types, the failure to consider their expression changes may provide slightly biased results. Second, using actual counts for cellular subpopulations obtained by FACS as covariates in differential expression analysis is likely to provide more accurate expression change estimates than the computational deconvolution applied in this study.
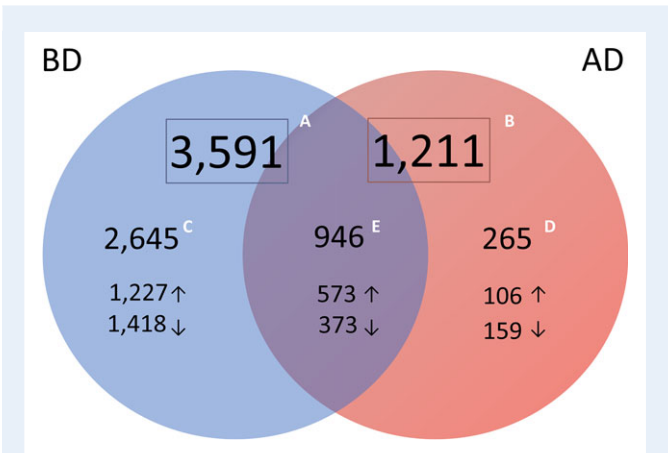
Table I Top significantly differentially expressed genes in endometrial whole-tissue samples, that remained sub-significant after deconvolution analysis. FC - average expression fold change between mid-secretory and early secretory samples. P-value—Bonferroni P-value. BD—before deconvolution.

Gene symbol	Gene name	FC BD	P-value BD
OLFM4	Olfactomedin 4	−15.8	9.02E−33
LEFTY1	Left-right determination factor 1	13.35	6.02E−35
LHFPL3	LHFPL tetraspan subfamily member 3	14.04	1.04E−30
BRINP1	BMP/retinoic acid inducible neural specific 1	12.58	4.87E−23
POSTN	Periostin	−12.48	5.11E−39
SLC47A1	Solute carrier family 47 member 1	−11.87	7.01E−44
FGF1	Fibroblast growth factor 1	−10.97	2.13E−49
GABRA3	Gamma-aminobutyric acid type A receptor alpha3 subunit	−9.52	1.83E−52
COL17A1	Collagen type XVII alpha-1 chain	8.9	4.96E−30
TMEM154	Transmembrane protein 154	8.68	2.42E−40

Table II Top significant novel genes. Meta-analysis p-value using whole-tissue expression profiling (Before Deconvolution, BD) and after adjusting for cellular composition (After Deconvolution, AD). Log2FC—average logarithmic expression fold change between mid-secretory and early secretory samples. *—RT-qPCR validated.

Gene symbol	Gene name	P-value BD	Log$_2$FC BD	P-value AD	Log$_2$FC AD
SLC26A4-AS1	SLC26A4 antisense RNA 1	5.75E−91	−5.07	6.68E−14	−4.17
LINC01320*	Long intergenic non-coding RNA 1320	2.77E−80	3.67	7.67E−11	2.03
PPFIA2	PTPRF interacting protein alpha 2	5.35E−71	2.45	1.17E−21	3
C12orf75	Chromosome 12 open reading frame 75	1.44E−66	1.89	2.95E−17	2.06
VEPH1	Ventricular zone expressed PH domain containing 1	2.38E−61	−2.06	5.48E−19	−2.54
SLC8A1*	Solute carrier family 8 member A1	3.18E−61	1.93	2.18E−18	2.28
TSPAN15	Tetraspanin 15	5.51E−60	2.32	4.86E−10	1.97
KLKP1	Kallikrein pseudogene 1	6.80E−59	−3.74	9.26E−07	−2.52
MTMR7	Myotubularin related protein 7	5.16E−55	2.59	1.80E−08	1.98
GLYATL2	Glycine-N-acyltransferase like 2	8.92E−55	−2.28	5.20E−14	−2.7

Table III Long intergenic non-coding RNAs, significantly differentially expressed in human endometrium during WOI. Log2FC—average logarithmic expression fold change between mid-secretory and early secretory samples. P-value—Bonferroni p-value. BD—before deconvolution. AD—after deconvolution. *—RT-qPCR validated.

Gene symbol	P-value BD	Log$_2$FC BD	P-value AD	Log$_2$FC AD
LINC01320*	2.77E−80	3.67	7.67E−11	2.02
LINC00665	6.84E−32	−1.04	2.18E−09	−1.18
LINC00886	3.53E−26	−1.22	6.84E−09	−1.53
LINC00964	2.84E−33	−1.96	2.23E−08	−2.13
LINC01341	2.01E−26	1.42	1.26E−07	1.66
LINC00939	1.35E−10	−0.76	1.77E−07	−1.41
LINC00930	5.92E−18	−1.36	6.60E−07	−1.77
LINC00958	8.52E−07	−0.48	7.43E−07	−1.23

Nonetheless, computational deconvolution is becoming increasingly popular (Reinartz *et al.*, 2016; Scott *et al.*, 2016; Yu and He, 2017), and has been shown to have accuracy similar to methods using actual cell counts (Newman *et al.*, 2015). As a marked methodological improvement, we addressed the transcriptomic changes in human endometrium using meta-analysis of samples from different populations, which not only helps to further eliminate population-specific and batch effects, but also to gain the statistical power required. As a result, we suggest a set of robust endometrial receptivity markers that reflect actual expression changes that take place in endometrial tissue, irrespective of its cellular composition. However, novel potential endometrial transcriptomic markers identified in our study, such as LINC01320, need further investigation to elucidate their role in endometrial functioning.

In conclusion, our results show how the proportions of endometrial epithelial and stromal cells change in endometrial biopsies during the transition from the pre-receptive to receptive state and demonstrate how these changes affect the gene expression profile of biopsied tissue. Overall, our results highlight that the cellular composition of biopsies must be taken into consideration in endometrial 'omics' studies where whole-tissue samples are involved.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Acknowledgements

Authors' roles

M.S. and V.K. wrote the manuscript. M.S., A.V.M., C.S., A.S. and T.L. participated in designing the study; M.S., A.V.M., J.F.M.B., F.M.C. and F.V. performed the experiments. Data analysis was coordinated by T.L., performed by M.S. and V.K., and contributions were made by

References

Altmäe S, Esteban FJ, Stavreus-Evers A, Simón C, Giudice L, Lessey BA, Horcajadas JA, Macklon NS, D'Hooghe T, Campoy C et al. Guidelines for the design, analysis and interpretation of 'omics' data: focus on human endometrium. *Hum Reprod Update* 2014;**20**:12–28.

Altmäe S, Koel M, Võsa U, Adler P, Suhorutšenko M, Laisk-Podar T, Kukushkina V, Saare M, Velthut-Meikas A, Krjutškov K et al. Meta-signature of human endometrial receptivity: a meta-analysis and validation study of transcriptomic biomarkers. *Sci Rep* 2017;**7**:10077.

Altmäe S, Martinez-Conejero JA, Salumets A, Simon C, Horcajadas JA, Stavreus-Evers A. Endometrial gene expression analysis at the time of embryo implantation in women with unexplained infertility. *Mol Hum Reprod* 2010;**16**:178–187.

Altmäe S, Reimand J, Hovatta O, Zhang P, Kere J, Laisk T, Saare M, Peters M, Vilo J, Stavreus-Evers A et al. Research resource: interactome of human embryo implantation: identification of gene expression pathways, regulation, and integrated regulatory networks. *Mol Endocrinol* 2012;**26**: 203–217.

Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;**31**:166–169.

Boon ACM, Williams RW, Sinasac DS, Webby RJ. A novel genetic locus linked to pro-inflammatory cytokines after virulent H5N1 virus infection in mice. *BMC Genomics* 2014;**15**:1017.

Borthwick JM, Charnock-Jones DS, Tom BD, Hull ML, Teirney R, Phillips SC, Smith SK. Determination of the transcript profile of human endometrium. *Mol Hum Reprod* 2003;**9**:19–33.

Carson DD, Lagow E, Thathiah A, Al-Shami R, Farach-Carson MC, Vernon M, Yuan L, Fritz MA, Lessey B. Changes in gene expression during the early to mid-luteal (receptive phase) transition in human endometrium detected by high-density microarray screening. *Mol Hum Reprod* 2002;**8**:871–879.

Catalano RD, Critchley HO, Heikinheimo O, Baird DT, Hapangama D, Sherwin JRA, Charnock-Jones DS, Smith SK, Sharkey AM. Mifepristone induced progesterone withdrawal reveals novel regulatory pathways in human endometrium. *Mol Hum Reprod* 2007;**13**:641–654.

Chen BJ, Byrne FL, Takenaka K, Modesitt SC, Olzomer EM, Mills JD, Farrell R, Hoehn KL, Janitz M. Transcriptome landscape of long intergenic non-coding RNAs in endometrial cancer. *Gynecol Oncol* 2017a; **147**:654–662.

Chen Z, Huang A, Sun J, Jiang T, Qin FX-F, Wu A. Inference of immune cell composition on the expression profiles of mouse tissue. *Sci Rep* 2017b;**7**:40508.

Coutifaris C, Myers ER, Guzick DS, Diamond MP, Carson SA, Legro RS, McGovern PG, Schlaff WD, Carr BR, Steinkampf MP et al. Histological dating of timed endometrial biopsy tissue is not related to fertility status. *Fertil Steril* 2004;**82**:1264–1272.

Crum CP, Hornstein MD, Nucci MR, Mutter GL. Hertig and beyond: a systematic and practical approach to the endometrial biopsy. *Adv Anat Pathol* 2003;**10**:301–318.

Dassen H, Punyadeera C, Delvoux B, Schulkens I, Marchetti C, Kamps R, Klomp J, Dijcks F, de Goeij A, D'Hooghe T et al. Olfactomedin-4 regulation by estrogen in the human endometrium requires epidermal growth factor signaling. *Am J Pathol* 2010;**177**:2495–2508.

Díaz-Gimeno P, Horcajadas JA, Martínez-Conejero JA, Esteban FJ, Alamá P, Pellicer A, Simón C. A genomic diagnostic tool for human endometrial receptivity based on the transcriptomic signature. *Fertil Steril* 2011;**95**: 50–60.

Edgell TA, Rombauts LJF, Salamonsen LA. Assessing receptivity in the endometrium: the need for a rapid, non-invasive test. *Reprod Biomed Online* 2013;**27**:486–496.

Enciso M, Carrascosa JP, Sarasa J, Martínez-Ortiz PA, Munné S, Horcajadas JA, Aizpurua J. Development of a new comprehensive and reliable endometrial receptivity map (ER Map/ER Grade) based on RT-qPCR gene expression analysis. *Hum Reprod* 2018;**33**:220–228.

Evans GE, Martínez-Conejero JA, Phillipson GTM, Simón C, McNoe LA, Sykes PH, Horcajadas JA, Lam EYN, Print CG, Sin IL et al. Gene and protein expression signature of endometrial glandular and stromal compartments during the window of implantation. *Fertil Steril* 2012;**97**:1365–1373.

Fischer AH, Jacobson KA, Rose J, Zeller R. Hematoxylin and eosin staining of tissue and cell sections. *CSH Protoc* 2008;**2008**:pdb. prot4986.

Fluhr H, Spratte J, Bredow M, Heidrich S, Zygmunt M. Constitutive activity of Erk1/2 and NF-κB protects human endometrial stromal cells from death receptor-mediated apoptosis. *Reprod Biol* 2013;**13**:113–121.

Gnainsky Y, Granot I, Aldo P, Barash A, Or Y, Mor G, Dekel N. Biopsy-induced inflammatory conditions improve endometrial receptivity: the mechanism of action. *Reproduction* 2014;**149**:75–85.

Gong T, Szustakowski JD. DeconRNASeq: a statistical framework for deconvolution of heterogeneous tissue samples based on mRNA-Seq data. *Bioinformatics* 2013;**29**:1083–1085.

Groen JN, Capraro D, Morris KV. The emerging role of pseudogene expressed non-coding RNAs in cellular functions. *Int J Biochem Cell Biol* 2014;**54**:350–355.

Gómez E, Ruíz-Alonso M, Miravet J, Simón C. Human endometrial transcriptomics: implications for embryonic implantation. *Cold Spring Harb Perspect Med* 2015;**5**:a022996.

Haller-Kikkatalo K, Altmäe S, Tagoma A, Uibo R, Salumets A. Autoimmune activation toward embryo implantation is rare in immune-privileged human endometrium. *Semin Reprod Med* 2014;**32**:376–384.

Harper MJ. The implantation window. *Baillieres Clin Obstet Gynaecol* 1992; **6**:351–371.

Hayman DM, Blumberg TJ, Scott CC, Athanasiou KA. The effects of isolation on chondrocyte gene expression. *Tissue Eng* 2006;**12**:2573–2581.

Hu S, Yao G, Wang Y, Xu H, Ji X, He Y, Zhu Q, Chen Z, Sun Y. Transcriptomic changes during the pre-receptive to receptive transition in human endometrium detected by RNA-Seq. *J Clin Endocrinol Metab* 2014;**99**:E2744–E2753.

Huang S, Liu F, Niu Q, Li Y, Liu C, Zhang L, Ni D, Pu X. GLIPR-2 overexpression in HK-2 cells promotes cell EMT and migration through ERK1/2 activation. *PLoS One* 2013;**8**:e58574.

Kao LC, Tulac S, Lobo S, Imani B, Yang JP, Germeyer A, Osteen K, Taylor RN, Lessey BA, Giudice LC. Global gene profiling in human endometrium during the window of implantation. *Endocrinology* 2002;**143**:2119–2138.

Kodithuwakku SP, Ng P-Y, Liu Y, Ng EHY, Yeung WSB, Ho P-C, Lee K-F. Hormonal regulation of endometrial olfactomedin expression and its suppressive effect on spheroid attachment onto endometrial epithelial cells. *Hum Reprod* 2011;**26**:167–175.

Krjutškov K, Katayama S, Saare M, Vera-Rodriguez M, Lubenets D, Samuel K, Laisk-Podar T, Teder H, Einarsdottir E, Salumets A *et al*. Single-cell transcriptome analysis of endometrial tissue. *Hum Reprod* 2016;**31**:844–853.

Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* 2014;**30**:523–530.

Kung JTY, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics* 2013;**193**:651–669.

Li B, Liu JS, Liu XS. Revisit linear regression-based deconvolution methods for tumor gene expression data. *Genome Biol* 2017;**18**:127.

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;**25**:402–408.

Maekawa R, Taketani T, Mihara Y, Sato S, Okada M, Tamura I, Jozaki K, Kajimura T, Asada H, Tamura H *et al*. Thin endometrium transcriptome analysis reveals a potential mechanism of implantation failure. *Reprod Med Biol* 2017;**16**:206–227.

Miravet J, Ruiz-Alonso M, Taguchi S, Clemente M, Funabiki M, Gomez E, Marin C, Nakamura Y, Simon C. Endometrial injury does not affect the endometrial transcriptomic profile during the window of implantation. *Hum Reprod* 2017;**32**:i352.

Mirkin S, Arslan M, Churikov D, Corica A, Diaz JI, Williams S, Bocca S, Oehninger S. In search of candidate genes critically expressed in the human endometrium during the window of implantation. *Hum Reprod* 2005;**20**:2104–2117.

Murray MJ, Meyer WR, Zaino RJ, Lessey BA, Novotny DB, Ireland K, Zeng D, Fritz MA. A critical analysis of the accuracy, reproducibility, and clinical utility of histologic endometrial dating in fertile women. *Fertil Steril* 2004;**81**:1333–1343.

Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015;**12**:453–457.

Noyes RW, Hertig AT, Rock J. Dating the endometrial biopsy. *Am J Obstet Gynecol* 1975;**122**:262–263.

Paiva P, Hannan NJ, Hincks C, Meehan KL, Pruysers E, Dimitriadis E, Salamonsen LA. Human chorionic gonadotrophin regulates FGF2 and other cytokines produced by human endometrial epithelial cells, providing a mechanism for enhancing endometrial receptivity. *Hum Reprod* 2011;**26**:1153–1162.

Pan Y, Bush EC, Toonen JA, Ma Y, Solga AC, Sims PA, Gutmann DH, Pan Y, Bush EC, Toonen JA *et al*. Whole tumor RNA-sequencing and deconvolution reveal a clinically-prognostic PTEN/PI3K-regulated glioma transcriptional signature. *Oncotarget* 2017;**8**:52474–52487.

Reinartz S, Finkernagel F, Adhikary T, Rohnalter V, Schumann T, Schober Y, Nockher WA, Nist A, Stiewe T, Jansen JM *et al*. A transcriptome-based global map of signaling pathways in the ovarian cancer microenvironment associated with clinical outcome. *Genome Biol* 2016;**17**:108.

Riesewijk A, Martín J, Os R, van, Horcajadas JA, Polman J, Pellicer A, Mosselman S, Simón C. Gene expression profiling of human endometrial receptivity on days LH+2 versus LH+7 by microarray technology. *Mol Hum Reprod* 2003;**9**:253–264.

Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;**26**:139–140.

Ruiz-Alonso M, Blesa D, Díaz-Gimeno P, Gómez E, Fernández-Sánchez M, Carranza F, Carrera J, Vilella F, Pellicer A, Simón C. The endometrial receptivity array for diagnosis and personalized embryo transfer as a treatment for patients with repeated implantation failure. *Fertil Steril* 2013;**100**:818–824.

Schelker M, Feau S, Du J, Ranu N, Klipp E, MacBeath G, Schoeberl B, Raue A. Estimation of immune cell content in tumour tissue using single-cell RNA-seq data. *Nat Commun* 2017;**8**:2032.

Scott LJ, Erdos MR, Huyghe JR, Welch RP, Beck AT, Wolford BN, Chines PS, Didion JP, Narisu N, Stringham HM *et al*. The genetic regulatory signature of type 2 diabetes in human skeletal muscle. *Nat Commun* 2016;**7**:11764.

Shen-Orr SS, Gaujoux R. Computational deconvolution: extracting cell type-specific information from heterogeneous samples. *Curr Opin Immunol* 2013;**25**:571–578.

Sigurgeirsson B, Åmark H, Jemt A, Ujvari D, Westgren M, Lundeberg J, Gidlöf S, Comprehensive RNA. sequencing of healthy human endometrium at two time points of the menstrual cycle. *Biol Reprod* 2016;**96**:24–33.

Strell C, Hilscher MM, Laxman N, Svedlund J, Wu C, Yokota C, Nilsson M. Placing RNA in context and space - methods for spatially resolved transcriptomics. *FEBS J* 2018. Available from: http://doi.wiley.com/10.1111/febs.14435.

Talbi S, Hamilton AE, Vo KC, Tulac S, Overgaard MT, Dosiou C, Le Shay N, Nezhat CN, Kempson R, Lessey BA *et al*. Molecular phenotyping of human endometrium distinguishes menstrual cycle phases and underlying biological processes in normo-ovulatory women. *Endocrinology* 2006;**147**:1097–1121.

Tapia A, Gangi LM, Zegers-Hochschild F, Balmaceda J, Pommer R, Trejo L, Pacheco IM, Salvatierra AM, Henríquez S, Quezada M *et al*. Differences in the endometrial transcript profile during the receptive period between women who were refractory to implantation and those who achieved pregnancy. *Hum Reprod* 2008;**23**:340–351.

Tseng L-H, Chen I, Chen M-Y, Yan H, Wang C-N, Lee C-L. Genome-based expression profiling as a single standardized microarray platform for the diagnosis of endometrial disorder: an array of 126-gene model. *Fertil Steril* 2010;**94**:114–119.

Wang L, Fu H, Nanayakkara G, Li Y, Shao Y, Johnson C, Cheng J, Yang WY, Yang F, Lavallee M *et al*. Novel extracellular and nuclear caspase-1 and inflammasomes propagate inflammation and regulate gene expression: a comprehensive database mining study. *J Hematol Oncol* 2016;**9**:122.

Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;**26**:2190–2191.

Yu J, Berga SL, Zou W, Sun H-Y, Johnston-MacAnanny E, Yalcinkaya T, Sidell N, Bagchi IC, Bagchi MK, Taylor RN. Gap junction blockade induces apoptosis in human endometrial stromal cells. *Mol Reprod Dev* 2014;**81**:666–675.

Yu Q, He Z. Comprehensive investigation of temporal and autism-associated cell type composition-dependent and independent gene expression changes in human brains. *Sci Rep* 2017;**7**:4121.

Yu J, Wu J, Bagchi IC, Bagchi MK, Sidell N, Taylor RN. Disruption of gap junctions reduces biomarkers of decidualization and angiogenesis and increases inflammatory mediators in human endometrial stromal cell cultures. *Mol Cell Endocrinol* 2011;**344**:25–34.