Transcription profile analysis of the endometrium revealed molecular markers of the personalized ‘window of implantation’ during in vitro fertilization

O. V. Burmenskaya, V. K. Bozhenko, V. Yu. Smolnikova, E. A. Kalinina, I. E. Korneeva, A. E. Donnikov, E. P. Beyk, V. A. Naumov, N. V. Aleksandrova, P. I. Borovikov and D. Y. Trofimov

Federal State Budget Institution ‘Research Center for Obstetrics, Gynecology and Perinatology’ Ministry of Healthcare of the Russian Federation, Moscow, Russia

ABSTRACT
To determine the most informative markers for assessing the functional state of endometrium during the ‘window of implantation’ and creating a model for assessment of the readiness of endometrium for embryo implantation.

Forty-seven women with tubal infertility and a successful IVF pregnancy participated in the study. Pipelle endometrial sample was performed during the supposed ‘window of implantation’ in natural cycle with subsequent histological study, and transcriptional profile of genes GPX3, PAEP, DPP4, TAGLN, HABP2, IMPA2, AQP3, HLA-DOB, MSX1, POSTN determined by real-time quantitative polymerase chain reaction (qRT-PCR). Differences in the level of mRNA expression of all the studied genes in the receptive endometrium were found in comparison to the prereceptive one, which allowed us to classify two functional states of the endometrium. The results of histological examination responded to the stage of maturation of the endometrium in 78.7% of cases. Receptive endometrial status can be determined based on the integral evaluation of mRNA expression level of 4 PAEP, DPP4, MSX1, and HLA-DOB genes. The model for determining a personalized ‘window implantation’ is offered for practical application in ART.

Introduction
Human implantation is a complex and multifactorial process. The success of implantation depends on the quality of the embryos and the readiness of the endometrium for embryo nidation. Recent studies of global transcriptome gene profiles have been conducted under different conditions that affect the endometrium qualitative state. The microarray-based gene expression technology of endometrium identified that the ‘window of implantation’ not always correlates with the 19th–21th day of menstrual cycle. The ‘window of implantation’ is displaced for 2–3 days in every fourth patient with IVF failure and this has fatal consequences for implementation of the assisted reproductive technology (ART) [1]. The main result of the studies of global transcription profiles of genes was the identification of the most informative markers of the ‘window of implantation’, and some of them are already included in the ‘gold standard’ of the endometrium receptivity [2,3]. The method of quantitative reverse transcription and polymerase chain reaction (RT-PCR) can be an alternative to expensive, time-consuming, and redundant method of microarray analysis in clinical practice. To determine the personalized ‘window of implantation’ in women with unsuccessful attempts of IVF and ET program, it may be sufficient to study a small number of genes by real time RT-PCR – a method whose sensitivity and specificity is equal to microarray analysis, but the latter is more accessible to routine practice. We were the first to carry out different versions of multivariate analysis of expression data of endometrial tissue at different stages of menstrual cycle to identify a complex of potentially informative genes. Furthermore, we tested the markers for determining the receptive status of the endometrium using the real-time quantitative polymerase chain reaction (qRT-PCR). The aim of the study is to determine the most informative markers for assessing the functional state of endometrium during the ‘window of implantation’ and creating a model for assessing the readiness of endometrium for embryo implantation.

Materials and methods
Patients with tubal infertility factor underwent endometrial biopsy procedure using Pipelle endometrial sampler (Pipelle de Cornier®, Laboratorie C.C.D., France) on LH +7–8 days in cycle before stimulation superovulation (7–8 days after the peak of luteinizing hormone). Peak level of LH was diagnosed according to the US data and/or urine Clear Blue test (Unipath Ltd, UK). We enrolled in the study only women with successful pregnancy after IVF (n = 47).

The sample was divided into two parts. One part was fixed and stored in a special transport medium for RNA stabilization in IntactRNA bioassays in accordance with the Instructions for the Kit (Eurogen, Russia). The other part was prepared for histological examination.

After appropriate standard treatment, the samples were fixed in 10% neutral formalin for 24 h and then were encapsulated in paraffin. Sections 5 microns thick were made and stained with...
Bioinformatic analysis of open access data of normal endometrium expression performed in order to search for key genes and their expression changes in the processes of endometrial regeneration. The data available in free access were analyzed. The first stage of bioinformatic analysis is the search for data suitable for studying the selected object. In this case, it was interesting for us to study the expression of genes in the normal endometrium in different phases of the menstrual cycle. For this purpose, a search was made in the database of expressions ArrayExpress [https://www.ebi.ac.uk/arrayexpress/] [citations: https://www.ebi.ac.uk/arrayexpress/about.html]. As search terms, we used «endometrium» and «Homo sapiens». The following datasets were selected:

- E-GEOD-46735 Plasma microRNA profiles in women with and without endometriosis at different menstrual cycle phases
- E-GEOD-35287 Unique Transcriptome, Pathways, and Networks in the Human Endometrial Fibroblast Response to Pregestosterone in Endometriosis
- E-MTAB-694 Transcription profiling by array of human peritoneum, peritoneal endometriosis lesion and eutopic endometrium from endometriosis patients
- E-GEOD-6364 Transcription profiling of human endometrium from patients with a history of endometriosis vs. myometrial tissue
- E-GEOD-23339 Gene expression profiles of endometriosis
- E-GEOD-25628 Endometriosis transcription profiling
- E-GEOD-44207 CCAAT/enhancer-binding protein alpha is epigenetically silenced by histone acetylation in endometriosis and promotes the pathogenesis of endometriosis: a novel therapeutic target
- E-GEOD-40186 miRNA expression profiles in endometriotic cyst stalk cells (ECCks)
- E-MEXP-1251 Transcription profiling by array of vessels from endometriotic nodules and control endometrial and myometrial tissue
- E-GEO-40007 TNFalpha and IL1beta stimulate differential gene expression in endometrial stromal cells
- E-GEOD-31683 Knappel-like Factor 9 and Progesterone Receptor Co-Regulation of Decidualizing Endometrial Stromal Cells: Implications for its Loss in the Pathogenesis of Endometriosis
- E-GEOD-7846 Transcription profiling of human endometrial endothelial cells of eutopic endometrium of patients with endometriosis compared with control

Results

The results of PCA-analysis are shown in Figure 1. On the PCA plot, we see that the samples are divided according to the stage at which they were taken. The next step is to identify genes, the expression of which is the same at different stages of the cycle (expression increases/decreases). For this purpose, Ward’s hierarchical agglomerative clustering method was used [6]. Eight stable clusters of genes were identified (Figure 2). One cluster included 2148 genes, 2–1537, 3–2830, 4–3569, 5–3517, 6–2513, 7–3611, and 8–3795 genes.

Analyzing the obtained results and comparing them with the data of the ERA study [3], 10 genes were isolated, which would allow isolate the samples of the middle secretion stage in the cohort of endometrial specimens (Table 3).

Furthermore, we conducted a study of the transcriptional profiles of 10 genes in the endometrial specimens obtained at different stages of the menstrual cycle by RT-PCR. Cluster analysis helped to identify 2 groups of samples. The division into groups was well correlated, but not in all cases coincided with histological examination of endometrial biopsy specimens. Among 22 samples of the medium secretion stage, according to the histological study, only 15 (68.2%) were determined by RT-PCR as ‘receptive endometrium,’ other 7 samples as ‘pre-receptive endometrium.’
Among 25 samples of the early secretion stage, according to the histological examination, 22 (88%) samples were determined by RT-qPCR as ‘pre-receptive endometrium’, and 4 samples – as ‘receptive endometrium’.

Compared to the stage of early secretion in ‘window of implantation,’ the level of mRNA expression of 7 genes GPX3 increases 105 times (43–173, \( p = 1.5 \times 10^{-12} \)), PAEP – 104 times (29–602, \( p = 5.8 \times 10^{-14} \)), DPP4 – 29 times (10–73, \( p = 1.7 \times 10^{-12} \)), HABP2 – 15 times (9–32, \( p = 4.0 \times 10^{-11} \)), IMPA2 – 5 times (4–9, \( p = 2.5 \times 10^{-11} \)), AQP3 – 4 times (1.2–6, \( p = 1.0 \times 10^{-7} \)), TAGLN – 3 times (2.5–6, \( p = 1.5 \times 10^{-6} \)). The level of mRNA expression of 3 genes HLA-DOB decreases in 21 times (7–35, \( p = 8.2 \times 10^{-12} \)), MSX1 – 2.1 times (1.3–4, \( p = 3.4 \times 10^{-7} \)), POSTN – 1.5 times (1–2.3, \( p = .002 \)) (Figure 3).

During the ‘implantation window’ compared to the stage of early secretion, the level of mRNA expression of 7 genes increases.

A discriminant function has been chosen on the basis of multifactor analysis that allows to classify samples according to the level of expression of 4 genes: PAEP, DPP4, MSX1, and HLA-DOB. The optimum value of the cutoff is determined using the ROC-analysis. The area under the ROC curve was AUC =1.0 (\( p = 6.6 \times 10–13 \)). The cut off value of the IRE (index receptivity endometrium) was 62. The values of cut off above 62 classified the samples as ‘receptive endometrium.’ The sensitivity and specificity of the proposed method for evaluating the endometrial receptivity in the range of the threshold value was 100% compared to cluster analysis.

### Discussion

The study revealed changes in the expression of all the studied genes during ‘implantation window.’ The results are completely identical to the literature data [3]. The clustering method allowed us to distinguish two clusters of samples discriminating the receptive and pre-receptive endometrium. The results of this classification in 78.7% of cases corresponded to the histological conclusion, which corresponds to the literature data.

Each fifth female in the general population and every fourth woman with repeated unsuccessful IVF attempts is believed to have a histological conclusion that does not correspond to the receptor status of the endometrium [7].

For practical application, a classification, determined by the multivariate analysis method, can be proposed for the level of mRNA expression of 4 genes: PAEP, DPP4, MSX1, and HLA-DOB. The expression level of two of them is significantly increased (PAEP, DPP4), and the expression level of the other two (MSX1 and HLA-DOB) is lower in the receptive endometrium compared with the pre-receptive endometrium.

The gene PAEP (progestagen associated endometrial protein), also known as a pregnancy-associated endometrial alpha-2-globulin (PAEG) or placental protein 14 (PP14), encodes a glycoprotein of the lipocalin superfamily, the genes of most of which are grouped on the long arm of the chromosome 9. PAEP has three isoforms that play an important role in maintaining an intrauterine environment for providing, initiating, and elongating of pregnancy. In addition to the endometrium and amniotic fluid, high expression of this glycoprotein is found in the follicular fluid and

### Table 2. Data for processing from E-GEOD-6364.

| Source Name | Scan Name | Factor Value [disease] | Factor Value [clinical information] |
|-------------|-----------|------------------------|------------------------------------|
| GSM150214   | MSE_D_543 | endometriosis           | mid secretion phase                |
| GSM150221   | MSE_N_617 | normal                 | mid secretion phase                |
| GSM150216   | MSE_D_678 | endometriosis           | mid secretion phase                |
| GSM150195   | PE_D_651  | endometriosis           | proliferative phase                |
| GSM150210   | ESE_N_664 | normal                 | early secretion phase              |
| GSM150193   | PE_D_594  | endometriosis           | proliferative phase                |
| GSM150211   | MSE_D_33A | endometriosis           | mid secretion phase                |
| GSM150222   | MSE_N_626 | normal                 | mid secretion phase                |
| GSM150199   | MSE_D_73A | endometriosis           | mid secretion phase                |
| GSM150207   | ESE_D_599 | endometriosis           | early secretion phase              |
| GSM150205   | ESE_D_517 | endometriosis           | early secretion phase              |
| GSM150225   | MSE_N_153 | normal                 | mid secretion phase                |
| GSM150191   | PE_D_508  | endometriosis           | proliferative phase                |
| GSM150192   | PE_D_587  | endometriosis           | proliferative phase                |
| GSM150198   | PE_N_M16S | normal                 | proliferative phase                |
| GSM150220   | MSE_N_610 | normal                 | mid secretion phase                |
| GSM150208   | ESE_N_629 | normal                 | early secretion phase              |
| GSM150227   | MSE_N_M163 | normal                | mid secretion phase                |
| GSM150194   | PE_D_647  | endometriosis           | proliferative phase                |
| GSM150213   | MSE_D_540 | endometriosis           | mid secretion phase                |
| GSM150206   | ESE_D_575 | endometriosis           | early secretion phase              |
| GSM150197   | PE_N_598  | normal                 | proliferative phase                |
| GSM150201   | PE_N_M182 | normal                 | proliferative phase                |
| GSM150223   | MSE_N_659 | normal                 | mid secretion phase                |
| GSM150217   | MSE_D_97A | endometriosis           | mid secretion phase                |
| GSM150218   | MSE_D_72A | endometriosis           | mid secretion phase                |
| GSM150196   | PE_N_562  | normal                 | proliferative phase                |
| GSM150203   | ESE_D_489 | endometriosis           | early secretion phase              |
| GSM150215   | MSE_D_645 | endometriosis           | mid secretion phase                |
| GSM150199   | PE_N_M169 | normal                 | proliferative phase                |
| GSM150209   | ESE_N_650 | normal                 | early secretion phase              |
| GSM150212   | MSE_D_516 | endometriosis           | mid secretion phase                |
| GSM150224   | MSE_N_G98A| normal                 | mid secretion phase                |
| GSM150190   | PE_D_26A  | endometriosis           | proliferative phase                |
| GSM150204   | ESE_D_496 | endometriosis           | early secretion phase              |
| GSM150226   | MSE_N_M158 | normal               | mid secretion phase                |
| GSM150202   | ESE_D_27A | endometriosis           | early secretion phase              |
sperm plasma, where its functions are associated with providing a sequence of events necessary for fertilization [8].

The gene *DPP4* (dipeptidyl peptidase 4), also known as the CD26 differentiation cluster, encodes dipeptidyl peptidase-4, which is a membrane enzyme that hydrolyzes the peptide linkage of proline at the N-terminal position. The gene product is the surface antigen of activation of CD26 T cells, is expressed on the surface of most cells of the body, participates in immune regulation, signal transduction, glucose homeostasis and in apoptosis [9].

The gene *MSX1* (*msh homeobox 1*) or homeobox gene 7 *HOX7* encodes a protein of the homeobox family. As a result of interaction with components of transcription complexes, this protein provides repression of differentiation genes during embryogenesis, plays a role in suppression of tissue growth [10].

The gene *HLA-DOB* (major histocompatibility complex, class II, DO beta) encodes a beta chain of a protein that performs an auxiliary function in antigens presentation. The HLA-DO protein interacting with HLA-DM suppresses the peptide loading of the molecules of the main histocompatibility complex of class II, preventing the presentation of antigens [11].

We believe that the proposed model can be used to classify the pre-receptive and receptive endometrium in the natural cycles of IVF (LH +7) or in cycles using hormone replacement therapy (on the fifth day of progesterone administration), but it requires further validation for more quantities of samples. In addition, the issue of determining the postreceptive status of the endometrium remains urgent.

There were no samples of the ‘post-receptive’ endometrium in our study. Molecular classification of the three functional states of the endometrium (pre-receptive, receptive, and postreceptive)
would significantly optimize preimplantation diagnostics of the endometrium and fasten the process of determining the personalized ‘window of implantation’.

**Conclusion**

The study of the transcriptional profiles of the GPX3, PAEP, DPP4, TAGLN, HABP2, IMPA2, AQP3, HLA-DOB, MSX1, and POSTN genes, revealed significant differences in the level of mRNA expression of these genes in the pre-receptive and receptive endometrium and it allowed to be classified the samples according to their receptive status.

A proposed model for determining a personalized ‘window of implantation’ can be tested in future and used in ART programs.

**Disclosure statement**

The authors report no conflicts of interest.
Funding

The study was performed with the support from grant of Russian Foundation for Basic Research (RFBR). Code of grant: 16-29-07434.

References

1. Ruiz-Alonso M, Galindo N, Pellicer A, Simon C. What a difference two days make: “personalized” embryo transfer (pET) paradigm: a case report and pilot study. Hum Reprod 2014;29:1244–7.
2. Horcajadas JA, Pellicer A, Simón C. Wide genomic analysis of human endometrial receptivity: new times, new opportunities. Hum Reprod Update 2007;13:77–86.
3. Díaz-Gimeno P, Horcajadas JA, Martínez-Conejero JA, et al. A genomic diagnostic tool for human endometrial receptivity based on the transcriptomic signature. Fertil Steril 2011;95:50–60.
4. Ritchie ME, Phipson B, Wu D, et al. “limma powers differential expression analyses for RNA-sequencing and microarray studies.”. Nucleic Acids Study 2015;43:e47.
5. Venables WN, Ripley BD. Modern applied statistics with S. New York: Springer-Verlag; 2002.
6. Murtagh F, Legendre P. Ward’s hierarchical agglomerative clustering method: which algorithms implement Ward’s criterion? J Classif 2014;31:274.
7. Garrido-Gómez T, Ruiz-Alonso M, Blesa D, et al. Profiling the gene signature of endometrial receptivity: clinical results. Fertil Steril 2013;99:1078–85.
8. Seppäälä M, Taylor RN, Koistinen H, et al. Glycodelin: a major lipocalin protein of the reproductive axis with diverse actions in cell recognition and differentiation. Endocr Rev 2002;23:401–30.
9. Zhao S, Song M, Fan Y, et al. Elevation of plasma soluble CD26 levels during pregnancy. J Obstet Gynaecol Res 2012;38:272–9.
10. Zhang H, Catron KM, Abate-Shen C. A role for the Msx-1 homeodomain in transcriptional regulation: residues in the N-terminal arm mediate TATA binding protein interaction and transcriptional repression. Proc Natl Acad Sci USA 1996;93:1764–9.
11. Guce AI, Mortimer SE, Yoon T, et al. HLA-DO acts as a substrate mimic to inhibit HLA-DM by a competitive mechanism. Nat Struct Mol Biol 2013;20:90–8.