Identification of genes and pathways associated with endometriosis using bioinformatic analysis

Yaxiong Huang1,2, Yuanzhen Zhang1, Bin Li2, Lina Ke2

1Department of Gynecology and Obstetrics, Zhongnan Hospital, Wuhan University, Wuhan 430000, Hubei, P. R. China; 2Department of Gynecology and Obstetrics, Dongfeng Hospital, Hubei University of Medicine, Shiyan 442008, Hubei, P. R. China

Received November 10, 2019; Accepted April 23, 2020; Epub September 15, 2020; Published September 30, 2020

Abstract: Endometriosis can lead to infertility and the exact mechanism remains unclear. The aim of this study was to elucidate the essential genes and pathways linked to uterine receptivity in women with endometriosis and normal women by bioinformatic analyses. Methods: Gene expression profiles for GSE51981 and GSE6364 were downloaded from the GEO database. Differentially expressed genes (DEGs) between 72 endometriosis and 41 normal tissue samples were analyzed using GEO2R. Gene Ontology (GO) and pathway enrichment analyses were performed and Cytoscape software was used to visualize the protein-protein interaction (PPI) network. A total of 612 DEGs were identified in normal samples and 314 DEGs were identified in women with endometriosis between proliferative phase endometrium (PE) and mid-secretory phase endometrium (MSE). We also identified 1,692 DEGs in PE and 69 DEGs in MSE between endometriosis and normal samples. The DEGs of normal samples were enriched in GO terms and were correlated with ion channel activity and pathways related to Hedgehog signaling. Whereas the DEGs of the endometriosis samples were enriched in GO terms and were associated with receptor binding and pathways related to the p53 effectors pathway. The hub genes CDK1, CCNB1, KIF11 and BUB1 were identified from the PPI network in normal samples, whereas the hub genes CDC20, CCNB1 and CCNB2 were defined in endometriosis samples. The endometrium of women with endometriosis seems to differ from that of normal women. Our study identified hub genes and signaling pathways, which enhances our understanding of the potential molecular mechanisms of implantation failure in women with endometriosis.

Keywords: Endometriosis, implantation, infertility, uterine receptivity, enrichment analysis

Introduction

Endometriosis is defined by the growth of functional endometrium outside the uterus. The clinical manifestations of this disease range from no symptoms to chronic pelvic pain, dyspareunia and infertility. It affects approximately 10% of fertile women, and 30%-50% of women with endometriosis are infertile. Previous studies have demonstrated that endometriosis is associated with poor oocyte quality, abnormal embryonic development and defective endometrial receptivity, which contribute to infertility. Molecular markers of endometrial receptivity, including homeobox genes, leukemia inhibitory factor, integrin αvβ3, etc., have been reported to display aberrant expression patterns in both humans and mice with endometriosis.

John S Tamaresis et al [1] used the dataset GSE51981 to establish a prompt, low-risk, low-cost diagnostic method with high accuracy. However, Richard O Burney et al [2] used the dataset GSE6364 to identify progesterone resistance-related genes and candidate sensitivity genes in patients with endometriosis. While other research on implantation failure in endometriosis exists, these studies have certain limitations. First, the number of samples analyzed by microarrays was small, which may lead to a high false positive rate. Second, these studies were limited to the implantation period, whereas changes in the uterine endometrium are continuous; no detailed analysis in the proliferation period was included. Although plenty of studies provide insight into endometriosis-related infertility, the exact pathological mecha-
nism of endometrial receptivity deficiency in endometriosis remains largely unknown.

In the present study, we analyzed the gene expression data of both GSE51981 and GSE6364 databases to identify the differentially expressed genes (DEGs) between proliferative phase endometrium (PE) and mid-secretory phase endometrium (MSE) in the endometriosis group and normal group in order to explore the molecular mechanisms of endometrial receptivity of endometriosis. Simultaneously, a comparative study on microarray data between the endometriosis and normal groups was performed to show specific differences.

Material and methods

Microarray data search

Microarray data and corresponding clinical data were obtained from the Gene Expression Omnibus database (http://ncbi.nlm.nih.gov/geo/): GSE51981 and GSE6364. The GSE51981 dataset included 29 endometrium biopsy samples obtained from patients with endometriosis in PE; 28 from patients with endometriosis in MSE, 20 from normal women in PE and 8 from normal women in MSE. The GSE6364 dataset included endometrium samples from women with endometriosis and normal women. A total of 28 endometrium samples, including 13 control samples and 15 endometriosis samples, were eligible for further analysis. The control samples consisted of 5 samples in PE and 8 samples in MSE; the endometriosis samples included 6 samples in PE and 9 samples in MSE. A total of 72 samples from women with endometriosis (PE=35 and MSE=37) and 41 normal samples (PE=25 and MSE=16) were included in this study. These datasets were both based on the GPL570 Affymetrix platform and the Human Genome U133 Plus 2.0 Array.

Data processing and identification of differential genes

DEGs between PE and MSE in the endometriosis group and normal group were analyzed by GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r). We also conducted a comparative study on gene expression data from the endometrium between PE and MSE in both the endometriosis and normal groups (Table 1). The log2-fold change (log2 FC) was calculated. The cut-off criterion for DEG screening was set as |log2 FC| ≥2, and P≤0.05 was considered significant.

Gene Ontology (GO) and enrichment analysis of functional categories

The functional enrichment analysis tool FunRich was used to carry out functional enrichment analysis for DEGs. The following functional categories were analyzed: Gene Ontology (GO) items (cellular components, molecular function and biological processes) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. P<0.01 was considered the threshold value.

Protein-protein interaction (PPI) network integration and analysis

For the DEGs identified in these datasets, the Search Tool for the Retrieval of Interacting Genes (STRING) database was used to visualize protein-protein networks and to evaluate the validated interactions; a confidence score >0.9 was defined as significant. Cytoscape 3.5.1 was used to visualize the PPI networks, and the nodes that had a high degree of connectivity were defined as key genes with critical biological functions.

Results

Identification of DEGs in endometriosis and normal samples

A total of 612 genes were identified in normal samples and 314 genes were differentially expressed in endometriosis samples between PE and MSE in the GSE51981 and GSE6364 datasets using the cut-off criteria of an adjusted significance value of (P≤0.05) and |log2 FC| ≥2 as thresholds. We also identified 1,692 DEGs in PE and 69 DEGs in MSE between endometriosis and normal samples.

GO and pathway enrichment analysis

Following GO and pathway analyses of the DEGs, the results were collected and shown in Figures 1 and 2.

The DEGs in normal samples were enriched in molecular functions, including extracellular matrix structural constituent, ion channel activity, peptide hormone, guanylate cyclase ac-
Analyses of implantation failure in endometriosis

Table 1. Differentially expressed genes in the proliferative phase endometrium (PE) and mid-secretory phase endometrium (MSE) between women with endometriosis and normal women

<table>
<thead>
<tr>
<th></th>
<th>logFC</th>
<th>symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Down-regulated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.27</td>
<td>DIO2</td>
<td>deiodinase, iodothyronine, type II</td>
</tr>
<tr>
<td></td>
<td>-3.89</td>
<td>MYH10</td>
<td>myosin, heavy chain 10, non-muscle</td>
</tr>
<tr>
<td></td>
<td>-3.86</td>
<td>SMC3</td>
<td>structural maintenance of chromosomes 3</td>
</tr>
<tr>
<td></td>
<td>-3.78</td>
<td>PKP4</td>
<td>plakophilin 4</td>
</tr>
<tr>
<td></td>
<td>-3.69</td>
<td>HSP90AB1</td>
<td>heat shock protein 90 alpha family class B member 1</td>
</tr>
<tr>
<td>Up-regulated</td>
<td>5.75</td>
<td>FOSB</td>
<td>FosB proto-oncogene, AP-1 transcription factor subunit</td>
</tr>
<tr>
<td></td>
<td>4.71</td>
<td>FOS</td>
<td>Fos proto-oncogene, AP-1 transcription factor subunit</td>
</tr>
<tr>
<td></td>
<td>3.82</td>
<td>EGR1</td>
<td>early growth response 1</td>
</tr>
<tr>
<td></td>
<td>3.62</td>
<td>SOX13</td>
<td>SRY-box 13</td>
</tr>
<tr>
<td></td>
<td>3.6</td>
<td>JUNB</td>
<td>JunB proto-oncogene, AP-1 transcription factor subunit</td>
</tr>
<tr>
<td>MSE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Down-regulated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.98</td>
<td>DIO2</td>
<td>deiodinase, iodothyronine, type II</td>
</tr>
<tr>
<td></td>
<td>-2.74</td>
<td>PKP4</td>
<td>plakophilin 4</td>
</tr>
<tr>
<td></td>
<td>-2.69</td>
<td>CWH43</td>
<td>cell wall biogenesis 43 C-terminal homolog</td>
</tr>
<tr>
<td></td>
<td>-2.68</td>
<td>MUC15</td>
<td>mucin 15, cell surface associated</td>
</tr>
<tr>
<td></td>
<td>-2.61</td>
<td>NORAD</td>
<td>non-coding RNA activated by DNA damage</td>
</tr>
<tr>
<td>Up-regulated</td>
<td>5.53</td>
<td>FOS</td>
<td>Fos proto-oncogene, AP-1 transcription factor subunit</td>
</tr>
<tr>
<td></td>
<td>3.94</td>
<td>EGR1</td>
<td>early growth response 1</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>JUNB</td>
<td>JunB proto-oncogene, AP-1 transcription factor subunit</td>
</tr>
<tr>
<td></td>
<td>2.27</td>
<td>ZFP36</td>
<td>ZFP36 ring finger protein</td>
</tr>
<tr>
<td></td>
<td>2.09</td>
<td>IER3</td>
<td>immediate early response 3</td>
</tr>
</tbody>
</table>

tivity, metallopeptidase activity, lipid transporter activity and the pathways related to mesenchymal-to-epithelial transition, Hedgehog, transmembrane transport of small molecules, FoxO family signaling, SLC-mediated transmembrane transport and integrin family cell surface interactions.

The DEGs in endometriosis samples were mainly enriched in molecular functions associated with extracellular matrix structural constituents, receptor binding, nucleic acid binding, calcium ion binding, intracellular ligand-gated ion channel activity and metallopeptidase activity as well as pathways associated with mesenchymal-to-epithelial transition, FoxO family signaling, direct p53 effectors, transmembrane transport of small molecules, glypican pathway and melatonin degradation II.

**PPI network construction**

The PPI network of DEGs was constructed in the STRING database (Figures 3 and 4). In the PPI network, the hub genes cyclin dependent kinase 1 (CDK1), cyclin B1 (CCNB1), kinesin family member 11 (KIF11) and budding uninhibited by benzimidazoles 1 (BUB1) had higher connectivity degrees in normal samples; whereas the hub genes CDC20, CCNB1 and cyclin B2 (CCNB2) were defined in endometriosis.

**Discussion**

**Normal samples**

Embryo implantation is a critical step in the reproductive process, but successful implantation requires a receptive endometrium. The endometrium undergoes significant changes from proliferative phases to secretory phases in the menstrual cycle. As this transformation is accompanied by a continuous, well-planned expression of a large number of specific genes, gene expression profiles between proliferative PE and MSE in normal women can provide new insight.

In our study, 612 DEGs were identified in PE samples compared with MSE samples throu-
Analyses of implantation failure in endometriosis

Figure 1. Gene Ontology (GO) and pathway enrichment analysis of differentially expressed genes (DEGs) in normal women. X-axis is the-log10 of P and the percentage of genes. We considered P<0.01 to be statistically significant. Y-axis is the names of the biological processes.

Through gene expression profiling in the GSE51981 and GSE6364 datasets. In the PPI network, the hub genes CDK1, CCNB1, KIF11 and BUB1 had higher connectivity degrees than normal.

It is recognized that cell proliferation occurs through the G1, S, G2 and M phases of the cell cycle and is controlled by cyclins and Cdkks [3]. CDK1 plays a vital role in regulating the eukaryotic cell cycle by interacting with multiple cyclic proteins [4]. Studies have shown that CDK1 expression is inhibited in stromal cell decidualization [5], which is consistent with the results of our research in which CDK1 was significantly decreased in the secretory phase. Inhibition of CDK1 activity may be responsible for the conversion of endoreduplication in decidual cells, which induces the formation of polyploid stromal cells. The polyploid state may increase cell resistance to survival by inhibiting the apoptosis of decidual cells [5].

BUB1, a protein kinase located in the centromere in early mitosis, is required for centromere localization of other SAC proteins [6] and for monitoring microtubules connected to centromeres [7]. Depletion of BUB1 in mammalian cells can lead to nonlinear chromosome segregation during mitosis [8] and chromatid cohesion loss [9]. Studies have shown that BUB1 has a specific role in female fertility in meiosis [10].

CCNB1 is a regulatory protein participating in mitosis that is mainly expressed in the G2/M phase. Cell cycle control is crucial in the process of the G2/M (mitosis) transition [11]. Temporary supervision of CCNB1 during the cell cycle can lead to uncontrolled cell growth and abnormal cell function [12]. Experimental data confirm that the expression of CCNB1 in normal endometrial cells is inhibited in the secretory phase [13], which is in accordance with our results. It has also been shown that
Analyses of implantation failure in endometriosis

CCNB1 plays an essential role in the rhythmic proliferation and differentiation of normal human endometrial cells under steroids [13].

KIF11, which is distributed throughout the cytoplasm [14], is a mitotic kinesin that plays an essential role in the foundation of the mitotic spindle [15]. It is speculated that KIF11 is involved in the progression of many diseases.

From our study, we found that hub genes CDK1, CCNB1 and BUB1 were enriched in Polo-like kinase signaling events in the cell cycle pathway. Serine/threonine protein kinase 1 (PLK1) belongs to the threonine kinase family and is highly conserved from yeast to humans. It has been confirmed that PLK1 is involved in the regulation of checkpoints in different cell cycles, ensuring that cell cycle events occur in a strict order. PLK1 expression was highest in the late stage of proliferation and in early stages of secretion, while it was significantly reduced in the late stage of secretion [16]. This expression pattern may be related to apoptosis of the endometrium or endometrial glandular cells. PLK1 has been previously found in decidual cells and apoptotic trophoblast cells during implantation in mice [17]. The mechanism of the PLK1 pathway is unknown, although studies have shown that PLK1 can bind to BUB1 [18].

In the present study, Polo-like kinase signaling events in the cell cycle pathway were enriched in the hub genes (CDK1, CCNB1 and BUB1). Down-regulation of CDK1, BUB1 and CCNB1 at protein and mRNA levels can decrease cell proliferation. Our results further confirm that cell proliferation was inhibited from the proliferative phase to the mid-secretory phase.

Endometriosis samples

Endometriosis is marked by the existence of an extra-uterine functional endometrium, which affects 6-10% women of childbearing age in a variety of ways, including changes in follicular development, fertilization disorders, implant defects and poor oocyte quality with decreased implantation ability [19]. Studies have shown inconsistent results regarding whether
Analyses of implantation failure in endometriosis

Figure 3. Protein-protein interaction network of differentially expressed genes (DEGs) in normal women. Edges indicate connections between proteins and their thicknesses represent the combined score. Blue nodes represent down-regulated DEGs; red nodes represent up-regulated DEGs. The size of the node indicates P, the higher the value, the smaller the node. Hub genes are represented with quadrilateral and red font.

Figure 4. Protein-protein interaction network of differentially expressed genes (DEGs) in endometriosis. Edges indicate connections between proteins and their thicknesses represent the combined score. Blue nodes represent down-regulated DEGs; red nodes represent up-regulated DEGs. The size of the node indicates P, the higher the value, the smaller the node. Hub genes are represented with quadrilateral and red font.

the endometrium, oocyte or both affect conception rates in endometriosis. Therefore, we compared the gene expression profiles between proliferative PE and MSE in women with endometriosis to
Analyses of implantation failure in endometriosis

evaluate whether women with endometriosis have uterine receptivity defects.

In our study, we obtained data from the GSE-51981 and GSE6364 datasets and identified 314 DEGs between the PE and MSE tissues. The hub genes CDC20, CCNB1 and CCNB2 were defined in endometriosis.

CDC20 is one of the controllers of the spindle checkpoint. In the cell cycle, CDC20 seems to interact with several other proteins [20]. In mammals, CDC20 can participate in delayed and late mitosis and its expression is necessary for cell division [21]. In human malignant tumors, overexpression of CDC20 is related to chromosomal instability and aneuploidy [22].

The B-type cyclins B1 and B2 are important components of the cell cycle regulation system. CCNB1 is one of the key components controlling the G2/M phase transition. Studies have shown that ovarian sex hormones may regulate the expression of CCNB1 not only in other cells (endothelial cells and cancer) but also in human endometrial cells, including ectopic endometrium cells [23]. CCNB2 is also a member of the B cell cyclin family and plays a vital role in regulating the cell cycle mediated by transforming growth factor beta [24]. Both CCNB1 and CCNB2 are expressed in uterine stromal cells during pregnancy, but the level of CCNB2 is significantly higher than that of CCNB1 [25]. Previous studies have shown that knockout of the Ccnb1 gene can cause embryonic lethality in mice [26]. In contrast, mutant mice lacking Ccnb2 were fertile, although the number of offspring produced in mice with homozygous Ccnb2 deletion was significantly lower; the reasons for this difference remain unclear.

In our study, we observed some pathways associated with the FOXM1 transcription factor network, which was also found to be enriched in the hub genes (CCNB1 and CCNB2). FOXM1, a member of the forked box transcription factor family, is highly expressed in proliferating cells and plays a decisive role in the replication of DNA and mitosis by modulating genes that regulate transformation between the G1-S and G2-M phases. FOXM1 is strongly expressed in almost all malignant tumors and cancer cell lines and in a wide range of tissues during embryonic development [27]. However, it is rarely found in normal adult tissues. The FOXM1 gene is expressed and regulated at the early stage of the post decidual implantation, and conditional knockout of the FOXM1 gene showed a regional uterine decidualization defect at the site of implantation [28]. Down-regulation of FOXM1 may lead to implant failure in endometriosis.

The results of our bioinformatics analysis suggest that hub genes CCNB1 and CCNB2 affect the FOXM1 transcription factor network and regulate implantation in endometriosis patients.

Comparison between endometriosis and normal samples

Many studies have compared differentially regulated genes in the endometrium between endometriosis and normal samples only during the implantation window. However, we know that the window of implantation is controlled by a series of proliferative and differentiation of the uterine stroma and epithelium in humans and that these events lead to an epithelium that is receptive to blastocyst attachment. One of essential conditions for a successful pregnancy is hormonal regulation of endometrial cell proliferation, but the associated molecular mechanisms are still unclear.

In our study, we identified 1,692 DEGs in PE and 69 DEGs in MSE between endometriosis and normal samples. From this result, we can see that the main difference occurs in the proliferative phase. At the same time, we can see that the most highly regulated genes (FOS and DOI2) are the same.

We observed an up-regulation of FOS, a transcription factor induced by cytokines, growth factors, and estradiol that modulates the expression of genes which are involved in regulating cell proliferation, differentiation, survival, and angiogenesis [29]. FOS is inhibited by P4 in the rat uterus [30] and highly up-regulated in the endometrium in a baboon model of endometriosis [31]. It needs to be determined why FOS is up-regulated in PE in women with endometriosis, though P4 resistance in the endometrium of women with endometriosis may be conducive to the apparent cycle-independent FOS up-regulation observed herein in the secretory phase.
DIO2 is the most down-regulated gene in endometriosis. The endometrium has also been shown to be an additional producer of thyroid T3 and T4. DIO2, which has been observed in the endometrium of different mammals, is the richest enzyme that transforms T4 into more potent T3 [32]. Regulation of the endometrial thyroid signaling pathway, including synthesis and activation-deactivation processes, is carried out by pituitary TSH [33]. Thyroid-derived T4 and T3 are up-regulated by E2, which also increases generation of thyroid-binding globulin (TBG) and increases hypothalamic TRH by hCG through LH receptors. This endocrine pathway in early pregnancy has been described, although the effects of the increased thyroid hormones have not been clarified.

The results of the present study suggested that irrespective of the proliferative phase or secretory phase, the biggest change is the same in terms of FOS and DIO2, and we can see that the greatest difference between healthy people and patients with endometriosis occurs in the proliferative phase of the endometrium. These observations are consistent with the hypothesis that the endometrium in endometriosis may also be abnormal with regard to embryonic receptivity, especially due to the strong relationship with implantation failure that occurs in patients with endometriosis [34].

Conclusions

In conclusion, the endometrium undergoes significant changes from the proliferating stage of the menstrual cycle to the secretory phase. Our study provides an integrated bioinformatics analysis of DEGs and pathways that may occur in normal tissue and women with endometriosis. In the normal implantation, Polo-like kinase signaling events in the cell cycle pathway and hub genes, including CDK1, CCNB1 and BUB1, play essential roles, whereas in endometriosis, the FOXM1 transcription factor network and DEGs, including CDC20, CCNB1 and CCNB2, changes dramatically. The results of this study may help us better understand the potential molecular mechanisms of implantation failure in normal women and patients with endometriosis and may provide a framework for further investigation of the underlying mechanism(s) in the eutopic endometrium of women with endometriosis. There were some limitations of the present study. The lack of experimental verification in this study is one limitation. Further large-scale genetic and experimental studies are required to verify these findings in the future.

Acknowledgements

This work was supported by the Science and Technology Research Program of Shiyan (No. 17Y42).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Yuanzhen Zhang, Department of Gynecology and Obstetrics, Zhongnan Hospital, Wuhan University, No. 169, Donghu Road, Wuhan 430000, Hubei, P. R. China. Tel: +86-027-67812888; Fax: +86-027-67812892; E-mail: chenghao9511356@126.com

References


Analyses of implantation failure in endometriosis


