Analysis of Proteomic Differences between Eutopic Endometrium and Ectopic Endometrium in Patients with Endometriosis

Yongchang Chen  
China Medical University

Junbo Zhao  
Shengjing Hospital of China Medical University

Yuanyuan An  
Shengjing Hospital of China Medical University

Fangfang Bi (bifangfang168@163.com)  
Shengjing Hospital of China Medical University

Keywords: Proteomic, Eutopic Endometrium, Ectopic Endometrium, Endometriosis

DOI: https://doi.org/10.21203/rs.3.rs-45924/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Objective

To analyze the difference in protein expression between eutopic endometrium and ectopic endometrium from a proteomic aspect, to investigate a preclinical research basis in the occurrence and development of endometriosis (EMT), and provide a theoretical basis for new diagnosis and treatment of EMT.

Methods

Eutopic and ectopic endometrium tissue were collected from 8 patients with an ovarian endometriotic cyst. The protein was extracted for the screening of differential proteins, database retrieval and bioinformatics. Functional classification of differential proteins was performed by gene ontology (GO), secondary annotation, and subcellular structure positioning. The functions of the proteins were analyzed by GO, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, and protein domain enrichments. Results were further validated by clustering analysis.

Results

Among a total of 6,233 identified proteins, 5,395 proteins had quantitative information, including 190 up-regulated proteins and 153 down-regulated proteins. Bioinformatics of differential proteins showed that the up-regulated proteins were primarily distributed in the nuclei and closely associated with conformation change of DNA and ribosome duplication, while the down-regulated proteins were predominantly distributed in the extracellular, cytoplasmic membrane and other positions and were related to the migration of endometrial tissues.

Conclusion

The differential expression of proteins in eutopic endometrium and ectopic endometrium plays an important role in the occurrence and development of EMT and its resultant infertility. The proteomic analysis provides a new basis for future research investigating the pathogenesis of EMT.

Introduction

Endometriosis (EMT) is defined as the endometrial tissues being present in positions other than the uterine cavity. The morbidity rate of EMT is up to 10% in women of childbearing age \cite{1}, and this is increasing by 1.6% annually \cite{2}. Furthermore, EMT seriously influences the physical and psychological health of women and pathogenesis of EMT is complex. Following on the implantation theory proposed by Sampson in 1921, investigators have continuously explored the pathogenesis of EMT from various
aspects, such as gene, molecule, cell and tissue. Currently, EMT is generally considered a hormone-dependent disease regulated by multiple genes\(^3\). Infertility, dysmenorrhea, and dyspareunia are the primary clinical manifestations of EMT. As reported in relevant literature, the infertility rate is approximately 30% ~ 50% in EMT patients and there are currently additional preclinical studies investigating EMT-induced infertility. It is commonly thought that EMT-induced infertility is caused by the participation of multiple genes and the regulation of multiple molecules. Many different methods are used to diagnose and treat EMT, however its relapse rate remains high after treatment\(^4\). Therefore, it is also vital to investigate methods to prevent EMT relapse. According to recent studies, there are several differential genes between eutopic endometrium and ectopic endometrium\(^5\), which may be responsible for the occurrence and development of EMT, and the resultant infertility. In this study, the difference between eutopic endometrium and ectopic endometrium was further analyzed from a proteomic aspect, to offer a preclinical research basis to the occurrence and development of EMT and provide a theoretical basis for new diagnosis and treatment of EMT.

**Materials And Methods**

General data

Endometrial tissue from 8 patients with an ovarian cyst was collected after patients were admitted into the Ward of Micro-invasive Gynecology (1) in Shengjing Hospital of China Medical University from March 2017 to January 2018 (Table 1). In all patients, an ovarian cyst resection was performed under laparoscope. Ovarian EMT was confirmed by postoperative paraffin pathological examination and diagnostic dilation and curettage completed during the operation. The study protocol and study contents were both approved by the Ethics Committee of the hospital. Prior to the operation, all patients signed the informed consent form. During the operation, the tissues of ectopic endometrium and eutopic endometrium were collected and then snap frozen at -80°C.

Inclusion and exclusion criteria

The inclusion criteria included: patients with regular menstruation before the operation; patients with anti-Müllerian hormone (AMH) > 0.01 mIU/ml; patients confirmed as having ovarian EMT foci by postoperative paraffin pathological examination, and patients diagnosed as having hyperplasic endometrium by diagnostic dilation and curettage. The exclusion criteria included: patients with other endometrial lesions; patients with submucous hysteromyoma, and patients with concomitant adenomyosis. The eutopic and ectopic endometrium tissues served as group A and B, respectively.

Screening of differential proteins and bioinformatics

1. Protein extraction

The samples were removed from the -80°C freezer, 40 mg of each sample was weighed and put into a mortar precooled with liquid nitrogen, liquid nitrogen added to the mortar and the tissue fully ground into
a powder. In each group, lysis buffer solution of a volume equivalent to 4 folds of ground powder (8 M urea + 1% proteinase inhibitor and 2 mM ethylenediaminetetraacetic acid [EDTA]) was added to the mortar, sample powder lysed ultrasonically, and centrifuged at 4°C at 12000 g for 10 minutes. The supernatant was transferred to a new centrifuge tube, and the protein concentration detected using a BCA protein detection kit.

2. Trypsin digestion and TMT labeling

Dithiothreitol was added to the protein solution to give a final concentration of 5 mM and reduced at 56°C for 30 min. Iodoacetamide was then added to give a final concentration of 11 mM and incubated at room temperature, in the dark, for 15 min and finally diluted to give a urea concentration of < 2 M in the sample. Subsequently, trypsin at a mass ratio of 1:50 (trypsin:protein) was added to the protein solution and enzymolysed overnight at 37°C. The following morning, trypsin at a mass ratio of 1:100 (trypsin:protein) was added and the solution continuously enzymolysed for 4 h. The enzymolytic peptide segments from the trypsin digestion were desalted with Strata X C18 (Phenomenex, USA), and lyophilized under vacuum. Following this, the peptide segments were dissolved with 0.5 M triethylammonium bicarbonate (TEAB) buffer, and labeled. Briefly, the thawed labeling reagent (TMT kit) was dissolved in acetonitrile, the peptide segments mixed in and incubated at room temperature for 2 h. The labeled peptide segments were mixed, desalted, and lyophilized under vacuum.

3. Grading by high performance liquid chromatography (HPLC)

The peptide segments were graded by high-pH reverse-phase high performance liquid chromatography (HPLC). Agilent 300 Extend C18 chromatographic column (particle size: 5 μm, internal diameter: 4.6 mm, length: 250 mm) was used. In total 60 components were separated from peptide segments with acetonitrile (pH 9) by a grading gradient of 8% - 32% for 60 minutes, and then combined into 18 components. The combined components were lyophilized under vacuum.

4. Liquid chromatography - mass spectroscopy (LC-MS) joint analysis

The peptide segments were dissolved with LC mobile phase A (0.1% (v/v) formic acid aqueous solution), and separated by the EASY-nLC 1000 ultra-high performance liquid chromatography system. Mobile phase A was the aqueous solution containing 0.1% formic acid and 2% acetonitrile; mobile phase B was the aqueous solution containing 0.1% formic acid and 90% acetonitrile. The LC gradient was set as follows: 0 - 40 minutes, 9% - 25% B; 40 - 52 minutes, 25% - 36% B; 52 - 56 minutes, 36% - 80% B; 56 - 60 minutes, 80% B. The flow rate was maintained at 700 nL/min.

The peptide segments were separated by the ultra-high performance liquid chromatography system, injected into a NSI ion source (voltage: 2.0 kV), ionized, and analyzed by Q Exactive™ Plus mass spectrometer. The parent ions and secondary fractions of peptide segments were detected and analyzed using a high-resolution Orbitrap. The following parameters were used: primary mass spectroscopy, scanning range 350 - 1800 m/z, scanning resolution 70,000; secondary mass spectroscopy, scanning
range starting from 100 m/z, scanning resolution 35,000. The data were acquired with the data dependent acquisition (DDA) program. In order to improve the effective utilization rate of mass spectroscopy, the parameters were set as follows: automatic gain control (AGC) 5E4; signal threshold, 10,000 ions; maximum injection time, 200 ms. For tandem mass spectroscopic scanning, the dynamic exclusion time was set as 30 s, to avoid repeated scanning of parent ions.

5. Database search

The secondary mass spectroscopic data were retrieved in the SwissProt Human database (including 20,130 sequences) using Maxquant (v1.5.2.8). The reverse library was added to calculate the false discovery rate (FDR) caused by the randomized matching. The common contamination library was also added into the database to eliminate the influence of contaminated proteins from the identification results. The following retrieval parameters were used: enzyme digestion mode, Trypsin/P; number of omitted digestion loci, 2; minimum length of peptide segment, 7 amino acid residues; maximum modification number of peptide segments, 5; mass error tolerance of primary parent ions, 20 ppm (First search), 5 ppm (Main search); mass error tolerance of secondary fractions, 0.02 Da; fixed modification for cysteine alkylation; variable modification for methionine oxidation and protein N-terminal acetylation; quantitative method, TMT-10plex (W8003TQ) and TMT-6plex (71170TQ); FDR, 1% (for both protein identification and peptide-spectrum match [PSM] identification).

6. Quality control test for mass spectroscopy

The results of the quality control test for mass spectroscopic data are shown in Fig. 1. The majority of peptide segments were at a length of 8 - 20 amino acid residues, which conformed to the standards concerning trypsin digestion of peptide segments, indicating that the preparation of samples met the relevant standards.

7. Bioinformatics analysis

The proteins with quantitative information were determined with the following parameters: change threshold; fold change value of 2 fold and \( p < 0.05 \). The degree of linear correlation between the data in two groups was measured using Pearson correlation coefficient. Using three factors (i.e., biological process, cellular component and molecular function), the function of proteins was classified with Gene Ontology (GO) to further determine the biological role of the proteins. The prediction of the subcellular structure and classification statistics of differential proteins were determined using Wolfpsort software. In this study, the enrichment analysis of GO, Kyoto Encyclopedia of Genes and Genomes (KEGG) and protein domains were conducted to determine if the proteins with differential expression tended to be significantly enriched in some functional types. Using Fisher's exact test, the negative logarithmic transformation was made on the obtained \( p \) value.

Results
1. Identification of proteins with differential expression in eutopic and ectopic endometrium

Among a total of 6,233 identified proteins, 5,395 proteins had quantitative information, including 190 up-regulated proteins and 153 down-regulated proteins, as identified according to the change criteria (i.e., a fold change of > 2 fold or < ½ fold, indicated significance for up-regulation or down-regulation, respectively). A reproducibility test was performed on the samples using the Pearson correlation coefficient (Fig. 2), which showed that the samples had good reproducibility.

2. Functional classification of differential proteins

2.1 GO secondary annotation classification

Using the GO secondary annotation, the up-regulated proteins (Table 2) and down-regulated proteins (Table 3) were classified by function. The biological role of differential proteins was determined using 3 factors (i.e., biological process, cellular component and molecular function). The number of differential proteins was the largest in the secondary categories of cellular process, cell component and binding.

2.2 Subcellular structure localization of differential proteins

To further determine the position of differential proteins in the cellular structure, the structural localization was performed. The up-regulated proteins were distributed in various positions of the cell: nucleus (74 proteins); cytoplasm (32 proteins); extracellular (31 proteins), and mitochondrion (23 proteins) (Fig. 3). However, the distribution of down-regulated proteins was quite different from that of the up-regulated proteins: extracellular (43 proteins); cytoplasm (37 proteins), and nucleus (31 proteins) (Fig. 4).

3. Analysis on the enrichment function of differential proteins

3.1 GO enrichment analysis of differential proteins

In the following aspects, the up-regulated proteins were noticeably enriched: ribosome subunit; ribosome; chromosome; chromosome parts; organelle ribosome; mitochondrion ribosome; nucleic acid binding; structural components of ribosome; DNA binding; chromosome binding; RNA binding; activity of DNA topoisomerase; activity of structural molecules; activity of DNA topoisomerase II; conformation change of DNA; high molecular biosynthetic process of cells; biosynthetic process of macromolecules in cells; chromosome tissue; DNA packaging; protein-DNA composite subunit tissue; organic biosynthetic process; biosynthetic process; peptide metabolism process; biosynthetic process of cells; biosynthetic process of peptide; translation termination of mitochondrion; translation extension of mitochondrion, and translation (Fig. 5).

3.2 Enrichment of differential proteins in the KEGG pathway

The visualization of the KEGG pathway obtained by the enrichment analysis of differential proteins is shown in Fig. 6. Further analysis demonstrated that among the up-regulated proteins, HSA03010
ribosome was significantly enriched, followed by HSA03030 DNA duplication (Fig. 7). Among the down-regulated proteins, the digestion and absorption of HSA04974 protein was noticeably enriched (Fig. 8).

3.3 Domain enrichment of differential proteins

In the up-regulated proteins the following domains were enriched: minichromosome maintenance protein (MCM) and its N-terminal; OB domain; fibrinogen; α/β/γ chain and spiral coil domain (Fig. 9). In the down-regulated proteins, the calcium binding domain of epidermal growth factor and the cysteine-rich domain of growth factor receptor were enriched (Fig. 10).

4. Clustering analysis

According to the expression fold, the differential proteins were divided into four parts: Q1 (0 < Ratio ≤ 1/2); Q2 (1/2 < Ratio ≤ 1.5); Q3 (1.5 < Ratio ≤ 2), and Q4 (Ratio > 2) (Fig. 11). To further determine the functional correlation among proteins of different expression folds, the GO classification (Fig. 12), enrichment analysis of the KEGG pathway and protein domains (Fig. 13), and clustering analysis were performed for each part.

Discussion

It is well known that EMT has a high morbidity rate, a high infertility rate, and a high relapse rate, posing a difficult challenge to clinicians in both preclinical and clinical medicine. Dorien et al observed by comparison that anti-epithelial cell adhesion molecules (endometrial epithelial cells), CD10 (endometrial interstitial cells) and keratin in the abdominal uid were not statistically significantly different between EMT patients and non-EMT patients, and therefore suspected the association of endometrial cells in the abdominal uid with the occurrence of EMT [6]. In the present study, the difference in proteins between eutopic and ectopic endometrium was analyzed to provide a new theoretic basis for the pathogenesis of EMT and to provide a new direction for treating EMT, reducing its relapse and decreasing the infertility rate of EMT patients.

In the present study, the protein solution was degraded by trypsin enzymolysis, the obtained peptide segments labeled with TMT and then graded by HPLC to ensure number and purity. As shown by LC-MS joint analysis and the quality control test of mass spectroscopy, the majority of peptide segments were 8–20 amino acids long, which conformed to the standards for trypsin digestion of peptides, indicating that the preparation of samples met the relevant standards, proving that the experimental methods were effective. A total of 6,233 proteins were identified, of which 5,395 proteins had quantitative information, including 190 up-regulated proteins and 153 down-regulated proteins. The biological analysis was investigated using several methods (e.g., functional classification of differential proteins, localization of subcellular structures, the KEGG pathway and protein domains), to provide a basis for additional investigations including the screening of functional proteins and the functions of target proteins.
Using the GO secondary annotation classification, it was determined that the molecular function of differential proteins primarily included binding function, catalytic activity and structural molecule activity. Furthermore, by conducting additional GO enrichment analysis in the up-regulated proteins, it was found that the nucleic acid binding, structural components of ribosome, DNA binding, chromatin binding and RNA binding were all statistically significant. It is well known that, ribosome, DNA and chromatin are the core of cell proliferation; by up-regulating the proteins that bind with DNA and RNA, the speed of both cell duplication and transcription is increased in the tissues of ectopic endometrium, suggesting a basis for the malignant, fast growth of EMT. In the down-regulated proteins it was found that the actin binding, cell skeletal protein binding, components of muscular structure and calcium ion binding were statistically significant. By decreasing the binding of ectopic cells with actin, the expression of actin is increased, allowing improved migration capability of these cells. This may enable the migration of ectopic tissues and the reduction of binding with calcium ions. As a result, an increase in free calcium ions may allow the binding of more calcium ions with muscular tissues, and the motion capability of myoprotein. Similar results were reported by Zhang C et al [7]. By improving the motion capability of myoprotein, the migration of tissues of ectopic endometrium is made considerably easier. Following this, the migrated ectopic cells accelerate the duplication, transcription and translation through binding with DNA and RNA, thus promoting the growth of ectopic tissues. Although the motion and duplication capabilities of cells in ectopic tissues are enhanced, an adhesion effect between normal cells can inhibit the migration of cells. Moreover, there are many monitoring systems in the human body to phagocytize the cells of abnormal migration, indicating the need for further investigations.

The biological progress analysis and enrichment analysis performed in this study, and the additional observations found that: the expression of differential proteins was increased specifically in the change of DNA conformation, the biosynthetic process of macromolecules in cells, the biosynthetic process of macromolecules, chromosome tissues and DNA packaging. In contrast the expression of differential proteins was decreased specifically in cell adhesion, biological adhesion, development of muscular structure, development of extracellular structure, and extracellular matrix (ECM) tissues. These observations further indicate that ectopic endometrial cells have higher duplication and translation capabilities, and that their decreased adhesion is the foundation for their migration. It is also clear that ectopic cells decrease the recognition capability of immune cells by changing the extracellular structure and ECM tissue thereby making it easier to migrate. To further verify these findings, the cellular composition was analyzed. It was found that both the up-regulated and down-regulated proteins were predominantly located in the: cells; organelles; cell membrane; membrane-enclosed lumens; extracellular areas, and cell interfaces. As shown by the additional GO enrichment analysis, the ribosome subunit, the ribosome, the chromosome and partial proteins of the chromosome in the ectopic endometrial tissues were up-regulated. However, the protein ECM, contractile fibers, and ECM were down-regulated. The intracellular positions of both up-regulated and down-regulated proteins provide a structural foundation for the above biological processes and molecular functions.
In this study, the subcellular structure was further localized, and the results showed that the up-regulated proteins were primarily located in the nucleus and the cytoplasm while the down-regulated proteins were predominantly located in the extracellular sites and cytoplasm. It also confirmed the functions of the above proteins. In addition, several of the differentially expressed proteins were present in the cytoplasmic membrane, which is one of the reasons for the adherence, adhesion, and invasion capabilities of ectopic cells [8].

The KEGG pathway is an information network connecting the interactions between known molecules. To determine the link between the upstream genes of differential proteins, the enrichment analysis of the KEGG pathway was performed. The results showed that in differential proteins, the pathway-associated proteins relevant to DNA duplication were up-regulated, while those relevant to protein digestion and absorption were down-regulated, thus promoting the duplication of ectopic endometrium and decreasing the digestion and absorption of proteins. The ribosome pathway is a pathway to synthesize various ribosomal proteins (RPs) and mitoribosomal proteins. It is known that RPs participate in several processes including DNA repair and cell development, proliferation, differentiation, apoptosis and cancertation. The RPs of abnormal expression influences various processes of tumor cells (e.g., apoptosis, aging, growth and invasion) via various mechanisms [9]. In differential proteins, the expression of the ribosome pathway was statistically significantly different, which is one of the contributing factors in the occurrence and development of EMT, and its high relapse rate after treatment. In the DNA duplication pathway, DNA duplication is a complicated network involving interactions among various proteins and biological enzymes. An increase in DNA duplication is seen in many malignant tumors. For example, DNA duplication is up-regulated in ectopic endometrium, one of the reasons for the rapid growth of EMT. In the down-regulated pathways, there was a significant down-regulation of a pathway relevant to protein digestion and absorption. The upstream genes of this pathway are involved in the synthesis of collagen. Collagen is an important component of the ECM [10], and plays an important role in the occurrence and development of tumors [11, 12]. It has also be shown that collagen acts as a barrier against the invasion and metastasis of tumor cells [13]. According to the findings of some studies, collagen can change the adhesion and promote the movement of tumor cells by self-degradation and remodeling, thus facilitating the metastasis of these cells [14, 15]. In the present study, the down-regulation of pathways relevant to protein digestion and absorption was statistically significant, which was one of the factors contributing to the invasion and migration of EMT.

Protein domain refers to some components repeatedly appearing in different protein molecules and with a similar sequence, structure, and function. In differential proteins, MCM and its N-terminal were up-regulated. As the starting point of DNA duplication, MCM ensures that DNA is duplicated only once in a cell cycle. Overexpression of MCM has been observed in several malignant tumors [16]. The N-terminal of MCM, acts as the starting point of protein synthesis, and plays an important role in the stability and activity of proteins, intracellular localization, and intracellular signal conduction. Due to the up-regulation of MCM and its N-terminal, DNA duplication in ectopic endometrial tissues is accelerated, another factor in the malignant biological behaviors of EMT. Among the differential proteins, the cadherin domain of...
epidermal growth factor and the cysteine-rich domain of growth factor receptor were down-regulated.

Cadherin is a transmembrane protein, that mediates the mucosa of non-calcium ions (Ca\textsuperscript{2+})-dependent passing cells in epithelial cells. In the study by Zhu X et al, after the knockout of the cadherin gene (CDH1), the migration and invasion of in vitro endometrial epithelial cells were increased. It has also been shown that cadherin plays an important role in the development of cancer\textsuperscript{[17]}. However, Matsuzaki S et al concluded in their study, that in EMT patients, the down-regulation of cadherin in the middle secretory phase of the menstrual cycle was one of the factors associated with infertility in these patients\textsuperscript{[18]}. Results of more recent study showed that oxidative stress was a contributor to infertility in EMT patients, whereby cysteine is involved by producing glutathione\textsuperscript{[19]}. In the present study, the cysteine-rich domain of growth factor was down-regulated, however, its upstream genes and downstream pathways require further investigation.

In this study, the differential proteins were grouped according to their fold difference. The GO enrichment classification and the enrichment of KEGG and protein domain were performed in each group, further validating the results of the initial investigations. In the up-regulated proteins, the changes of DNA conformation and DNA topoisomerase were the primary contributors to abnormal DNA duplication\textsuperscript{[20, 21]}. In the down-regulated proteins, there was more significant enrichment of myoprotein, differentiation of striated muscle cells, differentiation of muscle cells, actin skeleton and band I. These findings are in support of the findings of the initial results.

This study demonstrated that ectopic cells offer a foundation for the growth of ectopic endometrium through the following factors: changing the structure of DNA; up-regulating the endonuclear proteins associated with DNA duplication; promoting the cell duplication pathways, and increasing the structural domains relevant to cell duplication. The ectopic cells promoted their own migration by down-regulating the extracellular and intracytoplasmic adhesion-associated proteins, changing the skeletal structure of actin, and evading the immune surveillance. In addition, the ectopic cells induced infertility in EMT patients by down-regulating the proteins that participate in cell oxidation (e.g., calcium-binding protein of epidermal growth factor and cysteine-rich domain receptors) thereby promoting oxidative stress. In this study, we preliminarily investigated the contributors to the invasion and malignant growth of EMT and its resultant infertility at a proteomic level, which provided a basis for the subsequent research on the functions and pathways of the differentially expressed proteins.

**Conclusion**

The differential expression of proteins in eutopic endometrium and ectopic endometrium plays an important role in the occurrence and development of EMT and its resultant infertility. The proteomic analysis provides a new basis for future research investigating the pathogenesis of EMT.

**Declarations**
Ethics approval and consent to participate:

NA

Consent for publication:

All the authors are consent for the publication.

Availability of data and material:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Competing interests:

There is no conflict of interests in this work.

Funding:

This work was supported by the 2020 innovation and entrepreneurship training plan for college students of China Medical University, National Natural Science Foundation of China (No. 81872125) and Liaoning Provincial doctoral research initial fund (No. 201601128)

Authors' contributions:

Fangfang Bi designed the article. Yongchang Chen and Junbo Zhao did the experiments. Yuanyuan An and Fangfang Bi wrote the manuscript. All the figures were prepared by Yongchang Chen and revised by Fangfang Bi.

Acknowledgements:

This work was supported by the 2020 innovation and entrepreneurship training plan for college students of China Medical University, National Natural Science Foundation of China (No. 81872125) and Liaoning Provincial doctoral research initial fund (No. 201601128)

References

1. Shafrir AL, Farland LV, Shah DK, Harris HR, Kvaskoff M, Zondervan K, et al. Risk for and consequences of endometriosis: A critical epidemiologic review. Best practice & research Clinical obstetrics & gynaecology. 2018; 51: 1–15.
2. Eisenberg VH, Weil C, Chodick G, Shalev V. Epidemiology of endometriosis: a large population-based database study from a healthcare provider with 2 million members. BJOG: an international journal of obstetrics gynaecology. 2018;125:55–62.

3. Rolla E. Endometriosis: advances and controversies in classification, pathogenesis, diagnosis, and treatment. F1000Research. 2019; 8.

4. Lee SY, Kim ML, Seong SJ, Bae JW, Cho YJ. Recurrence of Ovarian Endometrioma in Adolescents after Conservative, Laparoscopic Cyst Enucleation. J Pediatr Adolesc Gynecol. 2017;30:228–33.

5. Golestan Jahromi M, Aflatoonian R, Afsharian P, Aghajanpour S, Shahhoseini M, Aflatoonian A. Altered expression of 3 paralogous HOX A-D clusters in endometriosis disease: A case-control study. International journal of reproductive biomedicine (Yazd Iran). 2018;16:549–56.

6. O DF, Roskams T, Van den Eynde K, Vanhie A, Peterse DP, Meuleman C, et al. The Presence of Endometrial Cells in Peritoneal Fluid of Women With and Without Endometriosis. Reproductive sciences (Thousand Oaks. Calif). 2017;24:242–51.

7. Zhang C, Gao L, Yi Y, Han H, Cheng H, Ye X, et al. Adenosine Triphosphate Regresses Endometrial Explants in a Rat Model of Endometriosis. Reproductive sciences (Thousand Oaks, Calif). 2016; 23: 924 – 30.

8. Sundqvist J, Andersson KL, Scarselli G, Gemzell-Danielsson K, Lalitkumar PG. Expression of adhesion, attachment and invasion markers in eutopic and ectopic endometrium: a link to the aetiology of endometriosis. Human reproduction (Oxford England). 2012;27:2737–46.

9. Guo X, Shi Y, Gou Y, Li J, Han S, Zhang Y, et al. Human ribosomal protein S13 promotes gastric cancer growth through down-regulating p27(Kip1). J Cell Mol Med. 2011;15:296–306.

10. Hynes RO, Naba A. Overview of the matrisome—an inventory of extracellular matrix constituents and functions. Cold Spring Harb Perspect Biol. 2012;4:a004903.

11. Hamada S, Masamune A. Elucidating the link between collagen and pancreatic cancer: what's next? Expert Rev Gastroenterol Hepatol. 2018;12:315–7.

12. Xie SA, Zhang T, Wang J, Zhao F, Zhang YP, Yao WJ, et al. Matrix stiffness determines the phenotype of vascular smooth muscle cell in vitro and in vivo: Role of DNA methyltransferase 1. Biomaterials. 2018;155:203–16.

13. Fang M, Yuan J, Peng C, Li Y. Collagen as a double-edged sword in tumor progression. Tumour biology: the journal of the International Society for Oncodevelopmental Biology Medicine. 2014;35:2871–82.

14. Ray A, Lee O, Win Z, Edwards RM, Alford PW, Kim DH, et al. Anisotropic forces from spatially constrained focal adhesions mediate contact guidance directed cell migration. Nature communications. 2017;8:14923.

15. Vellinga TT, den Uil S, Rinkes IH, Marvin D, Ponsoen B, Alvarez-Varela A, et al. Collagen-rich stroma in aggressive colon tumors induces mesenchymal gene expression and tumor cell invasion. Oncogene. 2016;35:5263–71.
16. Nowinska K, Chmielewska M, Piotrowska A, Pula B, Pastuszewski W, Krecicki T, et al. Correlation between levels of expression of minichromosome maintenance proteins, Ki-67 proliferation antigen and metallothionein I/II in laryngeal squamous cell cancer. Int J Oncol. 2016;48:635–45.

17. Zhu X, Li Y, Zhou R, Wang N, Kang S. Knockdown of E-cadherin expression of endometrial epithelial cells may activate Wnt/beta-catenin pathway in vitro. 2018; 297: 117–23.

18. Matsuzaki S, Darcha C, Maleysson E, Canis M, Mage G. Impaired down-regulation of E-cadherin and beta-catenin protein expression in endometrial epithelial cells in the mid-secretory endometrium of infertile patients with endometriosis. J Clin Endocrinol Metab. 2010;95:3437–45.

19. Ferreira EM, Giorgi VSI, Rodrigues JK, de Andrade AZ, Junior AAJ, Navarro PA. Systemic oxidative stress as a possible mechanism underlying the pathogenesis of mild endometriosis-related infertility. Reprod Biomed Online. 2019;39:785–94.

20. Atkin ND, Raimer HM, Wang YH. Broken by the Cut: A Journey into the Role of Topoisomerase II in DNA Fragility. 2019; 10.

21. Cai J, Luo S, Lv X, Deng Y, Huang H, Zhao B, et al. Formulation of injectable glycyrrhizic acid-hydroxycamptothecin micelles as new generation of DNA topoisomerase I inhibitor for enhanced antitumor activity. International journal of pharmaceutics. 2019;571:118693.

Tables

Table 1: Basic information of EMT patients

| Num | Age | ASRM stage | CA125 U/ml | AMH miu/ml | Symptom | Type               |
|-----|-----|------------|------------|-----------|---------|--------------------|
| 1   | 38  | IV         | 36         | 0.39      | dysmenorrhea | -                  |
| 2   | 28  | III        | 48         | 0.96      | NA      | -                  |
| 3   | 28  | IV         | 53         | 2.08      | Sexual intercourse pain | - |
| 4   | 29  | IV         | 37         | 2.65      | dysmenorrhea | -               |
| 5   | 31  | III        | 36         | 1.53      | dysmenorrhea | Primary infertility |
| 6   | 38  | III        | 37         | 0.59      | dysmenorrhea, Sexual intercourse pain | - |
| 7   | 37  | IV         | 126        | 1.84      | dysmenorrhea, Infertility | Primary infertility |
| 8   | 38  | IV         | 282        | 0.84      | dysmenorrhea, Sexual intercourse pain, Infertility | Secondary infertility |

Note: -: Not applicable.
Table 2 Distribution of up-regulated proteins in GO secondary annotation

| B/A |
| GO Terms Level 1 | GO Terms Level 2 | No. of Protein |
|-----------------|-----------------|---------------|
| **Biological Process** | cellular process | 172 |
| | metabolic process | 143 |
| | single-organism process | 131 |
| | biological regulation | 114 |
| | cellular component organization or biogenesis | 102 |
| | response to stimulus | 75 |
| | multicellular organismal process | 68 |
| | localization | 66 |
| | developmental process | 64 |
| | signaling | 35 |
| | multi-organism process | 35 |
| | immune system process | 22 |
| | reproduction | 16 |
| | locomotion | 14 |
| | biological adhesion | 13 |
| | other | 12 |
| **Cellular Component** | cell | 183 |
| | organelle | 176 |
| | membrane-enclosed lumen | 103 |
| | membrane | 93 |
| | macromolecular complex | 81 |
| | extracellular region | 68 |
| | cell junction | 18 |
| | other | 9 |
| **Molecular Function** | binding | 171 |
| | catalytic activity | 66 |
| | structural molecule activity | 23 |
| | transcription factor activity, protein binding | 11 |
Table 3 Distribution of down-regulated proteins in GO secondary annotation

| Activity                                           | Count |
|----------------------------------------------------|-------|
| nucleic acid binding transcription factor activity | 10    |
| transporter activity                               | 10    |
| molecular function regulator                       | 7     |
| other                                              | 6     |
| GO Terms Level 1            | GO Terms Level 2                                      | No. of Protein |
|-----------------------------|------------------------------------------------------|---------------|
| **Biological Process**      |                                                      |               |
| cellular process            |                                                      | 126           |
| single-organism process     |                                                      | 123           |
| biological regulation       |                                                      | 100           |
| multicellular organismal process |                                              | 85            |
| response to stimulus        |                                                      | 80            |
| cellular component organization or biogenesis |                                          | 71            |
| developmental process       |                                                      | 68            |
| metabolic process           |                                                      | 65            |
| localization                |                                                      | 45            |
| signaling                   |                                                      | 33            |
| biological adhesion         |                                                      | 33            |
| immune system process       |                                                      | 22            |
| locomotion                  |                                                      | 14            |
| multi-organism process      |                                                      | 13            |
| other                       |                                                      | 14            |
| **Cellular Component**      |                                                      |               |
| cell                        |                                                      | 132           |
| organelle                   |                                                      | 129           |
| extracellular region        |                                                      | 95            |
| membrane                    |                                                      | 74            |
| membrane-enclosed lumen     |                                                      | 42            |
| macromolecular complex      |                                                      | 39            |
| supramolecular complex      |                                                      | 32            |
| cell junction               |                                                      | 27            |
| synapse                     |                                                      | 11            |
| **Molecular Function**       |                                                      |               |
| binding                     |                                                      | 134           |
| catalytic activity          |                                                      | 46            |
| structural molecule activity|                                                      | 19            |
| molecular function regulator|                                                      | 12            |
molecular transducer activity 10
signal transducer activity 8
transcription factor activity, protein binding 6
transporter activity 5
other 5

Figures

Figure 1
Figure 2

Pearson’s Correlation

-1  -0.5  0   0.5  1
Figure 3
Figure 4

Figure 5
Figure 7

-\text{log10}(\text{Fisher' exact test p value})

| Pathway                                      | -\text{log10} | p value |
|----------------------------------------------|---------------|---------|
| hsa03010 Ribosome                            | 2.19          | 6.82    |
| hsa05322 Systemic lupus erythematosus        | 1.55          |         |
| hsa00040 Pentose and glucuronate...         | 1.41          |         |
| hsa00140 Steroid hormone biosynthesis       | 1.39          |         |

Figure 8

-\text{log10}(\text{Fisher' exact test p value})

| Pathway                                      | -\text{log10} | p value |
|----------------------------------------------|---------------|---------|
| hsa04974 Protein digestion and absorption   | 5.53          |         |
| hsa05414 Dilated cardiomyopathy (DCM)        | 5.47          |         |
| hsa04510 Focal adhesion                      | 4.64          |         |
| hsa04512 ECM-receptor interaction            | 3.59          |         |
| hsa03320 PPAR signaling pathway              | 2.8           |         |
| hsa04530 Tight junction                      | 2.09          |         |
| hsa00982 Drug metabolism - cytochrome P450   | 1.96          |         |
| hsa04610 Complement and coagulation cascades | 1.62          |         |
| hsa00520 Amino sugar and nucleotide sugar... | 1.36          |         |
| hsa04260 Cardiac muscle contraction          | 1.34          |         |
Figure 9
Figure 10

-Log_{10}(Fisher's exact test p value)

| Domain                                           | Value |
|--------------------------------------------------|-------|
| EGF-like calcium-binding domain                  | 7.72  |
| Growth factor receptor cysteine-rich domain      | 7.43  |
| Complement C1r-like EGF domain                  | 7.1   |
| EGF-like domain                                  | 6.98  |
| Zinc finger, LIM-type                            | 5.7   |
| Fibronectin type III                             | 5.08  |
| Immunoglobulin I-set                             | 4.48  |
| von Willebrand factor, type A                    | 4.04  |
| Alpha crystallin/Hsp20 domain                    | 3.84  |
| Calponin homology domain                         | 3.74  |
| TB domain                                        | 3.54  |
| Immunoglobulin subtype 2                         | 3.47  |
| Immunoglobulin subtype                           | 3.27  |
| Olfactomedin-like domain                         | 3.21  |
| Myosin S1 fragment, N-terminal                   | 3.21  |
| Alpha-crystallin, N-terminal                     | 3.21  |
| Peptidase M14, carboxypeptidase A                | 2.95  |
| Coagulation factor 5/8 C-terminal domain         | 2.95  |
| Carboxypeptidase, regulatory domain              | 2.8   |
| Green fluorescent protein                        | 2.74  |
| SoHo domain                                      | 2.74  |
| G2 nidogen/fibulin G2F                           | 2.74  |
| AEBP1/CPX, carboxypeptidase domain               | 2.74  |
| Carboxypeptidase-like, regulatory domain         | 2.67  |
| Immunoglobulin-like fold                         | 2.5   |
Figure 11
Figure 13