Identification of key genes and pathways in endometriosis by integrated expression profiles analysis

Ding Cui[1], Yang Liu[1], Junyan Ma[1], Kaiqing Lin[2], Kaihong Xu[2] and Jun Lin[2]

[1] Department of Laboratory, Women's Hospital, Zhejiang University School of Medicine, Hangzhou, China
[2] Department of Gynecology and Obstetrics, Women's Hospital, Zhejiang University School of Medicine, Hangzhou, China

ABSTRACT

The purpose of this study was to integrate the existing expression profile data on endometriosis (EM)-related tissues in order to identify the differentially expressed genes. In this study, three series of raw expression data were downloaded from GEO database. Differentially expressed genes (DEGs) in three tissue types were screened. GO, KEGG pathway enrichment analysis, core differential genes (CDGs) protein–protein interaction (PPI) network and weighted gene co-expression network analysis (WGCNA) were performed, finally, the dysregulation of Hippo pathway in ectopic endometrium (EC) was detected by Western blotting. A total of 1,811 DEGs between eutopic (EU) and normal endometrium (NE), 5,947 DEGs between EC and EU, and 3,192 DEGs between EC and NE datasets were identified. After screening, 394 CDGs were obtained, and 5 hub genes identified in the PPI network. CDGs enrichment and WGCNA network analysis revealed cell proliferation, differentiation, migration and other biological processes, Hippo and Wnt signaling pathways, and a variety of tumor-related pathways. Western blotting results showed that YAP/TAZ was upregulated, and MOB1, pMOB1, SAV1, LATS1 and LATS2 were downregulated in EC. Moreover, CDGs, especially the hub genes, are potential biomarkers and therapeutic targets. Finally, the Hippo pathway might play a key role in the development of endometriosis.

INTRODUCTION

Endometriosis (EM) is a common gynecological disease and one of the leading causes of female infertility and chronic pelvic pain (*Giudice & Kao, 2004*). Although the prevalence of EM in fertile women is approximately 10% to 15% (*Olive & Pritts, 2001*), its exact pathogenesis is not yet clear. A classic hypothesis of EM etiology is the "retrograde menstruation theory" proposed by Sampson (*Sampson, 1927*) in 1927. According to this theory, endometrial debris enters the abdominal cavity with menstrual reflux, and then adheres in this ectopic position forming lesions. However, there are still many gaps in this theory, and it has not been proven conclusively. Subsequent studies found that EM is also associated with genetic, immunological, and inflammatory factors, along with

Corresponding author
Jun Lin, linjun@zju.edu.cn

endometrial stem/progenitor cells (*Burney & Giudice, 2012*; *Riccio et al., 2018*; *Cousins, Dorien & Gargett, 2018*). As a chronic disease, EM seriously affects the quality of life of patients and causes a heavy social and economic burden. Therefore, it is extremely important to explore its underlying molecular mechanisms, and discover novel markers to facilitate early diagnosis and individualized treatment.

Gene chip technology is a high-throughput detection method that can analyze the expression of all genes in a given tissue/organism at one time, making it an ideal technique for screening disease-related genes and biomarkers. Since the advent of gene-chip technology, a large amount of high-throughput data has been generated, uploaded and stored in public databases. Unfortunately, the use of these data is limited by the narrow objectives of most studies, and its full potential has not been tapped yet.

To understand the pathogenesis of EM, it is essential to compare the ectopic (EC), eutopic (EU) and normal (NE) endometrial tissues since all three are concurrently involved in EM development. Ectopic endometrium (EC) is defined as a condition that endometrium-like tissue grows outside the uterine cavity in patients with endometriosis, ectopic endometriotic nodules are formed due to the cyclical bleeding of EC and the fibrosis of the surrounding tissues, which could eventually lead to symptoms of dysmenorrhea, chronic pelvic pain, abnormal menstruation and infertility. Eutopic endometrium (EU) refers to the is the endometrial tissue grows in the uterine cavity in patients with endometriosis, while, the normal endometrium (NE) refers to the endometrial tissue grows in uterine cavity in patients with non-endometriosis. Some studies have reported that the EU in endometriosis patients has stronger proliferation, migration and invasion ability than the NE in non-endometriosis patients, thus, EU cells attach and colonize more easier outside of uterine cavity, thereby causing the occurrence of endometriosis. However, studies profiling the gene expression in EM have only focused on two out the three tissues, either due to the limitation of the research scale or the experimental design (*Hever et al., 2007*; *Burney et al., 2007*). In addition, traditional expression studies ultimately focus on the structural and/or functional changes in a single or a few genes. However, interactions between genes are complex, and any pathophysiological condition is a result of the synergistic effects of a large number of genes.

To overcome these limitations and gain new insights into EM, we first used a bioinformatics approach to integrate the existing expression profile data on all EM-related tissues in order to identify the differentially expressed genes. The interaction between these genes was then elucidated using biological networks.

MATERIALS & METHODS

Data mining

The gene expression profiles used in this study were downloaded from the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo) in three different series: GSE7305, GSE6364 and GSE4888 (Data S1). GSE7305 contains datasets of 10 paired EC and EU samples (*Hever et al., 2007*), GSE6364 contains datasets of 21 EU and 16 NE samples (*Burney et al., 2007*), and GSE4888 contains datasets of 27 NE samples (*Talbi et al., 2006*). After

Table 1 Source datasets overview.

Series	Samples			Platform
	NE	**EU**	**EC**	
GSE7305	–	10 / 10	10 / 10	HG-U133 PLUS 2.0
GSE6364	16 / 16	21 / 21	–	HG-U133 PLUS 2.0
GSE4888	27 / 12	–	–	HG-U133 PLUS 2.0

Notes.
Before "/" is the number of original samples contained in the data set, and after "/" is the actual number of samples used for analysis after excluding duplicate samples.

removing duplicate samples, 69 samples (EC = 10, EU = 31, NE = 28) were included for the samples (Table 1).

Data preprocessing and principal component analysis

The raw data was analyzed by the Affy package and batch effect removal was performed by sva package in R. After obtaining the normalized gene expression profiles, the principal component analysis (PCA) of 69 samples was performed based on the matrix.

Identification of differentially expressed genes

The annotation for the probes was downloaded from Affymetrix official website (http://www.affymetrix.com/support/technical/annotationfilesmain.affx). limma package was used to perform the pairwise comparisons of the three groups to screen for differentially expressed genes (DEGs) at P value < 0.01. Core differential genes (CDGs) shared by the three pairwise comparisons were thus obtained.

GO and KEGG pathway analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (*Kanehisa & Goto, 2000*) analysis of the candidate genes were performed using DAVID online tool (https://david.ncifcrf.gov/summary.jsp). The enrichment significance threshold was set to $P < 0.05$.

Protein–protein interaction network analysis

A CDGs-based protein–protein interaction (PPI) network was constructed using STRING (https://string-db.org) with high confidence >0.7) (*Franceschini et al., 2013*). Cytoscape (*Shannon et al., 2003*) was used to visualize the PPI network and analyze the connectivity degree. A node with a degree of connectivity greater than 10 was defined as a hub gene and is indicated in the figure.

Weighted gene co-expression network analysis

To further study the synergy between genes, a weighted gene co-expression network was constructed using WGCNA package (*Langfelder & Horvath, 2008*). The probe with coefficient of variation (CV) ≥ 0.2 was selected for co-expression network construction. The module was then used to correlate the source trait (EC/ EU/ NE) of the endometrial

tissues with the expression value of eigengene. Finally, candidate gene/probe expression values were aggregated into 28 different co-expression modules.

Clinical sample collection

In order to verify the association of the Hippo pathway with EM, we analyzed protein expression in EC, EU and NE endometrium by Western blotting. All subjects gave their written informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the Obstetrics and Gynecology Hospital of Zhejiang University School of Medicine (IRB approval number: 20160054). The patients' age was between 25–40 years, their menstrual cycles were regular, and all tissue samples were collected in the proliferative phase. No steroids were given to the patients at least 6 months prior to study inclusion.

Three paired EC and EU tissue samples were collected from patients at stage III/IV ovarian EM (diagnosed under a laparoscope according to the revised AFS classification), and three NE tissue samples were obtained from women without visible endometriosis.

Protein extraction and western blot

Total protein was extracted from the tissues using an immunoprecipitation lysate containing protease and phosphatase inhibitors, and measured using a BCA protein quantification kit (Thermo Fisher Scientific). Equal amounts (30 µg) of protein samples were separated in a 10% SDS polyacrylamide gel and transferred to polyvinylidene fluoride (PVDF) membranes (Thermo Fisher Scientific). The membranes were incubated overnight with primary rabbit antibodies against YAP/TAZ (1:1000, #8579; Cell Signaling Technology), SAV1 (1:1000, #8579; Cell Signaling Technology), MOB1 (1:1000, #8579; Cell Signaling Technology), phospho-MOB1 (1:1000, #8579; Cell Signaling Technology), LATS1 (1:1000, #8579; Cell Signaling Technology), LATS2 (1:1000, DF7516; Affinity Biosciences) and GAPDH (1:1000, ab9485; Abcam) at 4 °C, followed by HRP-conjugated goat anti-rabbit secondary antibody (1:1000; Abcam) for 1 h at room temperature. The signals were detected with an enhanced chemiluminescence detection reagent (Thermo Fisher Scientific), and the band densities were quantified and analyzed with Quantity One software (Bio-Rad Laboratories, Hercules, CA, USA). Relative protein levels were normalized to GAPDH.

RESULTS

Principal component analysis

The downloaded raw data was standardized to eliminate inter-assay differences. PCA was performed based on the expression profile of the 69 samples. The results showed that sample clustering was not related to the batch (Fig. 1A, Data S2), which eliminated the batch effect. The EC samples were significantly different from the other two groups, while the latter did not differ significantly amongst themselves (Fig. 1B, Data S2).
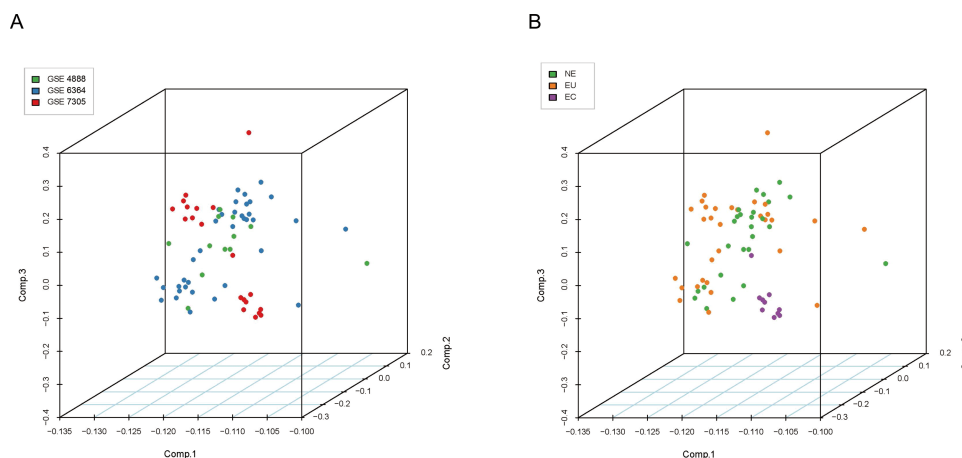
Figure 1 Principal component analysis (PCA) of gene expression profiles. Samples are marked with different colors based on the batch or the histologic origin. (A) Green for GSE4888, blue for GSE7305, and red for GSE6364; (B) Green for NE, orange for EU, and purple for EC.

Identification of DEGs between EC, EU and NE

EC, EU and NE samples were compared in pairs to identify DEGs, and Venn diagrams were drawn. The threshold was set at $P < 0.01$. A total of 1,811 DEGs (10,78 upregulated and 733 downregulated) between EU and NE, 5947 DEGs (3014 upregulated and 2,933 downregulated) between EC and EU, and 3,192 DEGs (1,754 upregulated and 1,438 downregulated) between EC and NE were observed (Fig. 2A, Data S3, Data S4). The top 5 up- and down-regulated genes in the three pairwise comparisons based on log2FC are listed in Table 2. To summarize, the least number of DEGs was seen between EU and NE ($n = 1811$) and the highest between EC and EU ($n = 5,947$). This was contrary to our expectation that greater differences ought to be seen between EC and NE rather than EC and EU. A total of 394 CDGs were shared across all three pairs (Fig. 2B, Data S5).

GO and KEGG pathway analysis of CDGs

GO analysis was performed based on the 394 CDGs (Data S6), and Fig. 3A shows the top 20 entries of BP terms. The enriched CDGs included not only those involved in important cellular activities like cytokinesis, cell differentiation, and cell migration but also the regulation of the key biological processes such as the EGFR and Wnt signaling pathways, inflammatory response etc.

KEGG pathway analysis was also performed on the CDGs (Data S7) and the top 20 entries with significant enrichment are shown in Fig. 3B. The enriched CDGs in pathways relevant to endometrial cancer and colorectal cancer may be related to the malignant features of EM. Furthermore, enriched CDGs in the Hippo signaling pathway suggests its involvement in the pathophysiological process of EM.

Protein–protein interaction network of CDGs

Using STRING database (*Franceschini et al., 2013*) and Cytoscape software (*Shannon et al., 2003*), CDGs were identified for protein-protien interactions and a PPI network was
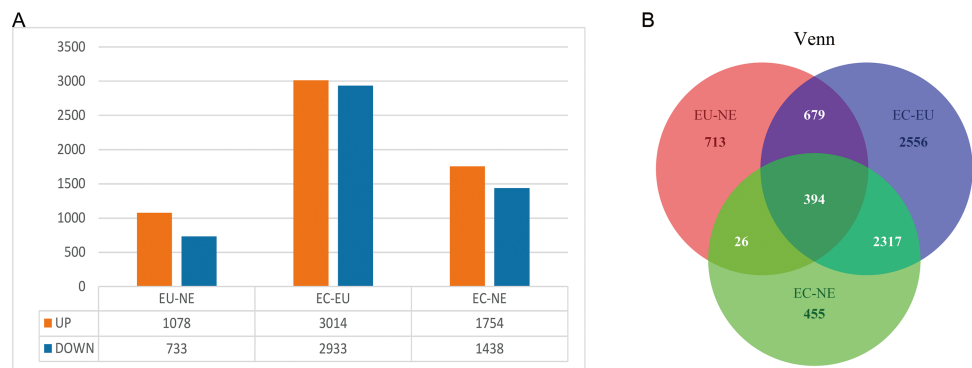
Figure 2 DEGs between EC, EU and NE. (A) Statistical analysis of DEGs between the three pairwise comparisons. Orange represents the number of up-regulated DEGs, and blue represents the number of down-regulated DEGs; (B) Venn diagrams of DEGs. Red represents DEGs between EU and NE, blue represents DEGs between EC and EU, and green represents DEGs between EC and NE. Crossing areas represent shared DEGs.

Table 2 Differentially expressed genes between EC, EU, and NE (top five).

Up-regulated			Down-regulated		
EU-NE	**EC-EU**	**EC-NE**	**EU-NE**	**EC-EU**	**EC-NE**
RRM2	DLK1	CLDN11	ABCA8	SCGB1D2	MMP26
CDCA7	CLDN11	DLK1	CEMIP	LINC01541	SCGB1D2
NDC80	GATA6	CHL1	GATA6	SCGB2A1	SCGB2A1
GDF7	SCN7A	SCN7A	CHI3L1	GABRP	LINC01541
HOXA10	COL8A1	GATA6	LOC102723694	SLC26A7	KIAA1324

constructed. After screening, 134 CDGs formed a PPI network consisting of 134 nodes and 227 edges (Fig. 4A). The size of the node is proportional to the degree of connectivity and a node with a connectivity degree > 10 is defined as a hub gene. A total of five hub genes were obtained: PIK3CA, CDK1, BIRC5, KIF2C, ITGA1 (Table 3), and marked in orange in the Fig. 4B. Using PIK3CA, the most connected hub gene, as the core, the nodes were directly connected to it to form the subnetwork containing 20 nodes and 33 edges (Fig. 4B).

Weighted gene co-expression network analysis

To further understand the synergy between genes, we performed WGCNA analysis on the gene expression profiles. The gene co-expression network was constructed on the basis of the correlation between the expression values of all the sample genes. Highly related genes formed a module, and 28 co-expression modules were finally identified (Fig. 5A; marked with different colors). The association of different modules with traits (i.e., different sources of endometrium: EC/ EU/ NE) showed EC correlated most with the "indianred4" module ($R = 0.32$, $P = 0.007$), followed by the "mediumpurple2" module ($R = 0.26$, $P = 0.03$) (Fig. 5B). The high correlation of the genes in these two modules with EC suggests their involvement in EM.

Figure 3 GO and KEGG analysis of CDGs (top 20). (A) GO (BP) enrichment analysis for 394 CDGs. (B) KEGG enrichment analysis for 394 CDGs.

Function annotation of EC-related modules

We conducted further analysis on the two modules showing highest correlation with EC. After annotating, the "indianred4" module contained 421 genes, and the "mediumpurple2" module contained 121 genes (Table 4, Data S8). KEGG pathway enrichment analysis revealed that the genes of the "indianred4" module were enriched in the Tight junction, RNA transport, Lysine degradation and Hippo signaling pathway entries (Fig. 6A, Table 4), while those of the "mediumpurple2" module were enriched mainly in Natural killer cell mediated cytotoxicity, Antigen processing and presentation, T cell and B cell receptor signaling pathway and other immune-related pathways, and VEGF Signaling pathway and
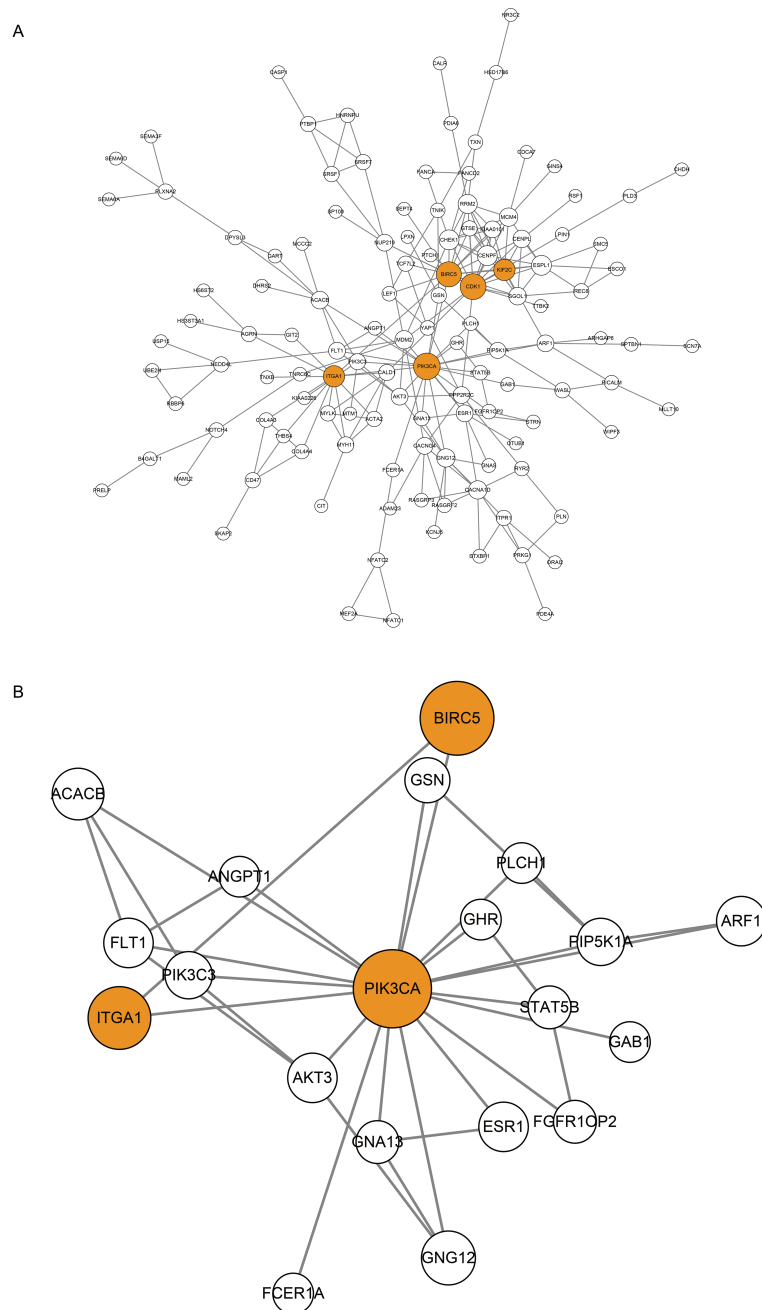
Figure 4 Protein-Protein Interaction (PPI) Network of CDGs. (A) Nodes represent CDGs, and edges represent interactions. Hub genes with connectivity degree >10 are marked in orange. (B) Subnetwork constructed with PIK3CA as the core.

Neurotrophin signaling pathway entries (Fig. 6B, Table 4). Therefore, the "indianred4" module may be related to the adhesion and growth of ectopic endometrium while the "mediumpurple2" module to EMs-related immune responses, and the formation of new blood vessels and nerves.

Table 3 Hub genes in PPI networks of CDGs (degree of connectivity >10).

Gene symbol	Connectivity degree
PIK3CA	19
CDK1	18
BIRC5	17
KIF2C	12
ITGA1	12

Dysregulation of Hippo signaling pathway in EMs

KEGG results indicated a key role of the Hippo signaling pathway in EM pathogenesis. In order to verify the association between the Hippo signaling pathway and EM, we confirmed the expression of the proteins encoded by the key genes (indicated on the map) of this pathway by Western blotting (Fig. 7A). Compared to the EU and NE samples, YAP/TAZ expression was increased, and SAV1, MOB1, p-MOB1, LATS1, and LATS2 were decreased in EC (Fig. 7B). These result indicate that the Hippo signaling pathway is abnormally activated in EC compared to EU and NE.

DISCUSSION

In order to compare the overall molecular phenotypes of ectopic (EC) and eutopic (EU) endometrial and non-endometrial (NE) tissues, and to identify key genes and pathways associated with EM, we performed bioinformatics analysis on gene-chip expression profiles from GEO database. We screened DEGs, constructed PPIs and weighted gene co-expression networks, and identified hub nodes and important co-expression modules in the networks. Finally, Western blotting was performed to validate the dysregulation of Hippo signaling pathway in EC tissues. These results provide new insights into EM.

Due to the different screening thresholds set in various studies, and inconsistencies in the batches, platforms, and observation priorities, it was difficult to directly compare those studies. According to the classical implantation theory (*Sampson, 1927*), ectopic endometrial lesions may originate from the recurrent endometrial tissue that flows back into the abdominal cavity with menstrual blood. Subsequent studies also reported that compared to normal endometrium, ectopic and eutopic endometria show stronger proliferation, adhesion, migration and invasiveness (*Lin et al., 2018*; *Yotova et al., 2011*; *Wingfield et al., 1995*; *Sundqvist et al., 2012*), which is important for the ectopic implantation and growth of the endometrial tissues. Ideally, the EC tissues ought to be closer to EU in terms of biological behavior and overall molecular phenotype. However, our results showed the highest number of DEGs between EC and EU, which prompts us to rethink the histologic origin of the ectopic endometrium.

The implantation hypothesis proposed by *Sampson (1927)* is accepted by most researchers. However, this theory cannot explain why only 10% of menstruating women are effected by EM when most women have retrograde menstruation (*Olive & Pritts, 2001*), or why extra-pelvic EM and pre-menarche EM occur (*Jubanyik & Comite, 1997*; *Schifrin, Erez & Moore, 1973*). With the discovery of endometrial stem cells, Sampson's theory has been
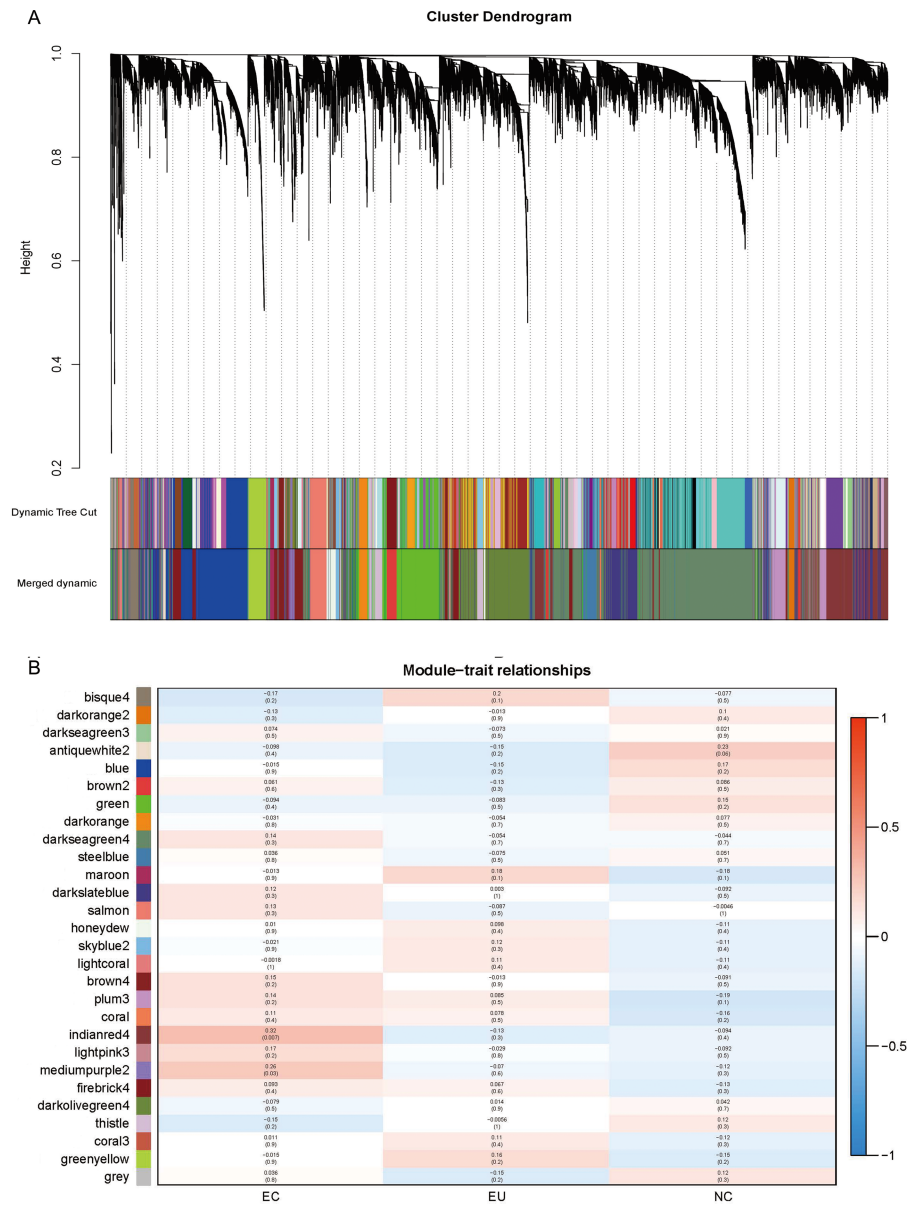
Figure 5 **Weighted gene co-expression network analysis (WGCNA) and Module trait relationship.** (A) The co-expression modules for each gene below the hierarchical clustering tree are marked with different colors. Dynamic Tree Cut corresponds to the original module; Merged dynamic corresponds to the merged modules finally obtained. (B) Red indicates positive correlation and blue indicates negative correlation. The number above each cell indicates the correlation coefficient between the module and the corresponding endometrial tissue, and the number in parentheses indicates the corresponding *P* value.

expanded; according to the new stem cell theory, endometrial stem cells are abnormally shed during menstruation and enter the abdominal cavity with retrograde menstruation and are ectopically implanted, eventually leading to EM (*Starzinski-Powitz et al., 2001*; *Leyendecker et al., 2002*; *Gargett, 2007*). This theory enhances the interpretive capacity of the morbidity model of EM, and at the same time explains our results; it is likely that the

Table 4 Genes within EC-related modules and KEGG pathway analysis (top 10).

Module	Gene	KEGG Term	Count	P-value
indianred4	421	Tight junction	8	2.68×10^{-3}
		RNA transport	7	3.07×10^{-2}
		Lysine degradation	4	3.43×10^{-2}
		Hippo signaling pathway	6	5.64×10^{-2}
mediumpurple2	121	Natural killer cell mediated cytotoxicity	19	5.47×10^{-20}
		Antigen processing and presentation	10	2.09×10^{-9}
		T cell receptor signaling pathway	7	7.88×10^{-5}
		Fc epsilon RI signaling pathway	6	1.08×10^{-4}
		Leukocyte transendothelial migration	7	1.68×10^{-4}
		Fc gamma R-mediated phagocytosis	6	2.95×10^{-4}
		Chemokine signaling pathway	8	3.02×10^{-4}
		VEGF signaling pathway	5	8.66×10^{-4}
		Neurotrophin signaling pathway	5	1.00×10^{-2}
		B cell receptor signaling pathway	4	1.27×10^{-2}
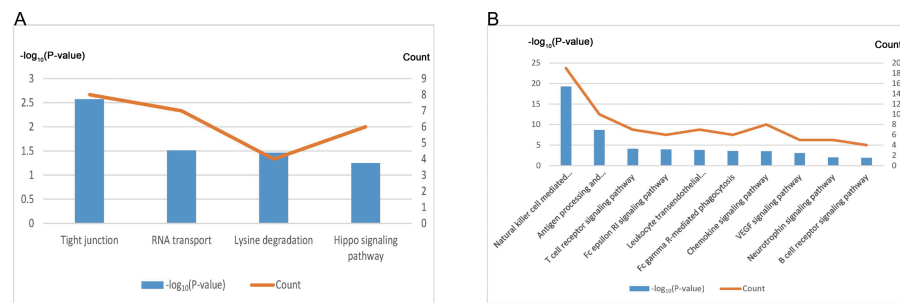
Figure 6 Functional annotation of EC-related modules. (A) KEGG pathway analysis of genes within the "indianred4" module. (B) KEGG pathway analysis of genes within the "mediumpurple2" module.

EC tissue does not originate from completely differentiated EU cells, but may be formed as a separate lineage during the initial differentiation.

The PPI network analysis revealed 5 hub genes: PIK3CA, CDK1, BIRC5, KIF2C and ITGA1. *Anglesio et al. (2017)* reported several cancer driver mutations, including PIK3CA, in 5 out of 27 non-cancer EM. Other studies have reported aberrant expression of CDK1 in gynecologic tumors such as ovarian cancer, cervical cancer and endometrial cancer (*Bae et al., 2015*; *Luo et al., 2016*; *Li, Qu & Li, 2017*), but its association with EM has not been verified. BIRC5, also known as Survivin, is an anti-apoptotic and pro-proliferative factor associated with various tumors (*Ghaffari et al., 2016*; *Wang et al., 2014*; *Wang et al., 2016*), and has recently been confirmed as a crucial factor in EM development as well as a valid diagnostic biomarker (*Uegaki et al., 2015*; *Acimovic et al., 2016*). KIF2C, the most representative member of the mitotic and cell cycle regulating Kinesin-13 family, is closely related to hepatocellular carcinoma, breast cancer and other tumors (*Chen et al., 2017*; *Shimo et al., 2008*), but its role in EMs is unclear. ITGA1 indirectly influences the
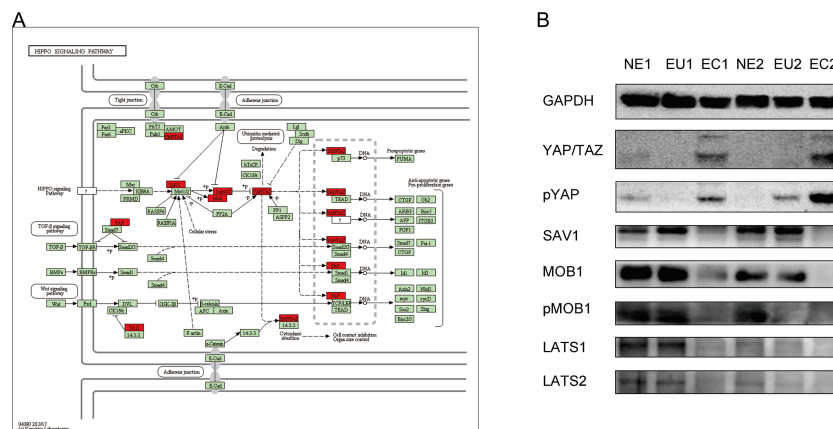
Figure 7 Dysregulation of Hippo signaling pathway in EC. (A) Schematic diagram of the Hippo signaling pathway. The key genes of interest are marked in red. (Image credit: the KEGG archive at https://www.kegg.jp). (B) Western blotting was performed to detect the expression and phosphorylation of Hippo signaling pathway proteins in EC, EU and NE.

proliferation, differentiation, and apoptosis of tumor cells by mediating cell adhesion processes (*Yim et al., 2013*; *Boudjadi et al., 2016*), but its association with EMs has not been reported.

The Hippo signaling pathway was first discovered in 1995 (*Justice et al., 1995*; *Xu et al., 1995*). It is highly conserved in mammals and regulates organ size and homeostasis by maintaining the balance between cell proliferation and apoptosis. The key component of the human Hippo signaling pathway is the core kinase chain consisting of MST1/2, SAV1, LATS1/2, and MOB1, which inhibit cell proliferation and promote apoptosis by inhibiting the downstream YAP/TAZ transcriptional coactivator (*Yu, Zhao & Guan, 2015*). YAP is the main effector molecule of the Hippo signaling pathway, which can promote proliferation, inhibit apoptosis, and ultimately lead to loss of cellular contact inhibition and malignant transformation. Dysregulation of the Hippo signaling pathway in tumors has been confirmed by a large number of studies in recent years (*Janse van Rensburg & Yang, 2016*; *Cao & Huang, 2017*). Interestingly, the Hippo signaling pathway also regulates the proliferation and differentiation of stem cells in various tissues (*Li et al., 2015a*; *Yin & Zhang, 2015*; *Hiemer & Varelas, 2013*). However, its association with EMs has only been reported for the downstream effector molecule YAP (*Song et al., 2016*). Our results suggests that the Hippo signaling pathway is abnormally activated in EC and may play an important role in the occurrence of EM.

Interestingly, three of the five identified hub genes are closely related to the Hippo pathway and its core factors. Studies have shown that the overexpression of PIK3CA is associated with the poor clinical prognosis in squamous cell carcinoma of the head and neck (HNSCC) (*Garcia-Escudero et al., 2018*). The tissue array showed that the expression of PIK3CA gene protein product (PI3Kp110α) was related to the YAP nuclear localization in HNSCC. *Zhao et al. (2018)* found that PIK3CA can enhance the activity of YAP and TAZ in breast tumors in vitro via multiple pathways. This indicated that YAP and TAZ

were the key mediators of PIK3CA-induced breast tumorigenesis, and had a synergistic effect with PIK3CA in transforming the breast cells. Another study showed that the simultaneous activation of the PI3K and YAP pathways triggered the rapid development of liver tumors in mice, and a synergistic effect was observed between PI3K and YAP (*Li et al., 2015b*). The CDK1 phosphorylation by YAP can promote the mitotic defects, migration and invasion of tumor cells, which is essential for malignant transformation (*Yang et al., 2013*). On the other hand, CDK1 phosphorylation of TAZ may inhibit its carcinogenic activity (*Zhang et al., 2015*). BIRC5 is an important downstream effectors of the Hippo pathway, and its expression is regulated by the phosphorylation of YAP / TAZ, which in turn affects the cell proliferation, apoptosis, epithelial-mesenchymal transition, invasion and metastasis (*Moriyama & Hori, 2019*).

CONCLUSIONS

Through comprehensive bioinformatics analysis of common gene-chip data, we compared gene expression between EC, EU, and NE tissues. Surprisingly, the number of DEGs between EC and EU was far greater than that between EC and NE. This may have implications for our understating of the underlying mechanism of EM. A total of 394 CDGs were shared by the three pair-wise comparisons, and five hub genes in the PPI network to which the CDGs belonged were finally screened. Through weighted gene co-expression network analysis, EM-related gene modules were identified. The CDGs enrichment analysis and EM-related modules functional annotation suggest that the Hippo signaling pathway may play an important role in EM. Western blotting results confirmed that compared to EU and NE, the expression and phosphorylation of the Hippo signaling pathway proteins were abnormal in EC, suggesting that the dysregulation of this pathway is of great significance in endometrial development. These above findings will help us better understand the pathological basis of EM and provide new directions for discovering diagnostic biomarkers and therapeutic targets.

Abbreviations

EM	endometriosis
GEO	Gene Expression Omnibus database
DEGs	differentially expressed genes
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genome
CDGs	core differential genes
PPI	protein–protein internaction
WGCNA	weighted gene co-expression network analysis
EC	ectopic endometrium
EU	eutopic endometrium
NE	non-endometriosis endometrium
PPI	protein–protein internaction
YAP	Yes Associated Protein
TAZ	Tafazzin

MOB1	Mps One Binder Homolog A
SAV1	Salvador Family WW Domain Containing Protein 1
LATS1	Large Tumor Suppressor Kinase 1
LATS2	Large Tumor Suppressor Kinase 2
PCA	principal component analysis
GAPDH	glyceraldehyde phosphate dehydrogenase
BP	biological process
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
CDK1	Cyclin Dependent Kinases 1
BIRC5	Baculoviral Inhibitor of Apoptosis Repeat Containing 5
KIF2C	Kinesin Family Member 2C
ITGA1	Integrin Subunit Alpha 1

Funding

This work was supported by the National Natural Science Foundation of China (No. 81370694 and No. 81571413). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Natural Science Foundation of China: 81370694, 81571413.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Ding Cui conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Yang Liu performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Junyan Ma performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Kaiqing Lin, Kaihong Xu and Jun Lin conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The study was approved by the Ethics Committee of the Obstetrics and Gynecology Hospital of Zhejiang University School of Medicine (Ethical Application Ref: 20160054).

Data Deposition

The raw data are available in the Supplemental Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.10171#supplemental-information.

REFERENCES

Acimovic M, Vidakovic S, Milic N, Jeremic K, Markovic M, Milosevic-Djeric A, Lazovic-Radonjic G. 2016. Survivin and VEGF as novel biomarkers in diagnosis of endometriosis. *Journal of Medical Biochemistry* **35**:63–68 DOI 10.1515/jomb-2015-0005.

Anglesio MS, Papadopoulos N, Ayhan A, Nazeran TM, Noe M, Horlings HM, Lum A, Jones S, Senz J, Seckin T, Ho J, Wu RC, Lac V, Ogawa H, Tessier-Cloutier B, Alhassan R, Wang A, Wang Y, Cohen JD, Wong F, Hasanovic A, Orr N, Zhang M, Popoli M, McMahon W, Wood LD, Mattox A, Allaire C, Segars J, Williams C, Tomasetti C, Boyd N, Kinzler KW, Gilks CB, Diaz L, Wang TL, Vogelstein B, Yong PJ, Huntsman DG, Shih IM. 2017. Cancer-associated mutations in endometriosis without cancer. *The New England Journal of Medicine* **376**:1835–1848 DOI 10.1056/NEJMoa1614814.

Bae T, Weon KY, Lee JW, Eum KH, Kim S, Choi JW. 2015. Restoration of paclitaxel resistance by CDK1 intervention in drug-resistant ovarian cancer. *Carcinogenesis* **36**:1561–1571 DOI 10.1093/carcin/bgv140.

Boudjadi S, Carrier JC, Groulx JF, Beaulieu JF. 2016. Integrin alpha1beta1 expression is controlled by c-MYC in colorectal cancer cells. *Oncogene* **35**:1671–1678 DOI 10.1038/onc.2015.231.

Burney RO, Giudice LC. 2012. Pathogenesis and pathophysiology of endometriosis. *Fertility and Sterility* **98**:511–519 DOI 10.1016/j.fertnstert.2012.06.029.

Burney RO, Talbi S, Hamilton AE, Vo KC, Nyegaard M, Nezhat CR, Lessey BA, Giudice LC. 2007. Gene expression analysis of endometrium reveals progesterone resistance and candidate susceptibility genes in women with endometriosis. *Endocrinology* **148**:3814–3826 DOI 10.1210/en.2006-1692.

Cao J, Huang W. 2017. Two faces of Hippo: activate or suppress the Hippo pathway in cancer. *Anti-Cancer Drugs* **28**:1079–1085 DOI 10.1097/CAD.0000000000000559.

Chen J, Li S, Zhou S, Cao S, Lou Y, Shen H, Yin J, Li G. 2017. Kinesin superfamily protein expression and its association with progression and prognosis in hepatocellular carcinoma. *Journal of Cancer Research and Therapeutics* **13**:651–659 DOI 10.4103/jcrt.JCRT_491_17.

Cousins FL, Dorien FO, Gargett CE. 2018. Endometrial stem/progenitor cells and their role in the pathogenesis of endometriosis. *Best Practice & Research Clinical Obstetrics & Gynaecology* **50**:27–38 DOI 10.1016/j.bpobgyn.2018.01.011.

Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, Von Mering C, Jensen LJ. 2013. STRING v9.1: protein–protein interaction networks, with increased coverage and integration. *Nucleic Acids Research* **41**:D808–D815 DOI 10.1093/nar/gks1094.

Garcia-Escudero R, Segrelles C, Duenas M, Pombo M, Ballestin C, Alonso-Riano M, Nenclares P, Alvarez-Rodriguez R, Sanchez-Aniceto G, Ruiz-Alonso A, Lopez-Cedrun JL, Paramio JM, Lorz C. 2018. Overexpression of PIK3CA in head and neck squamous cell carcinoma is associated with poor outcome and activation of the YAP pathway. *Oral Oncology* **79**:55–63 DOI 10.1016/j.oraloncology.2018.02.014.

Gargett CE. 2007. Uterine stem cells: what is the evidence? *Human Reproduction Update* **13**:87–101 DOI 10.1093/humupd/dml045.

Ghaffari K, Hashemi M, Ebrahimi E, Shirkoohi R. 2016. BIRC5 genomic copy number variation in early-onset breast cancer. *Iranian Biomedical Journal* **20**:241–245.

Giudice LC, Kao LC. 2004. Endometriosis. *Lancet* **364**:1789–1799 DOI 10.1016/S0140-6736(04)17403-5.

Hever A, Roth RB, Hevezi P, Marin ME, Acosta JA, Acosta H, Rojas J, Herrera R, Grigoriadis D, White E, Conlon PJ, Maki RA, Zlotnik A. 2007. Human endometriosis is associated with plasma cells and overexpression of B lymphocyte stimulator. *Proceedings of the National Academy of Sciences of the United States of America* **104**:12451–12456 DOI 10.1073/pnas.0703451104.

Hiemer SE, Varelas X. 2013. Stem cell regulation by the Hippo pathway. *Biochimica et Biophysica Acta* **1830**:2323–2334 DOI 10.1016/j.bbagen.2012.07.005.

Janse van Rensburg HJ, Yang X. 2016. The roles of the Hippo pathway in cancer metastasis. *Cellular Signalling* **28**:1761–1772 DOI 10.1016/j.cellsig.2016.08.004.

Jubanyik KJ, Comite F. 1997. Extrapelvic endometriosis. *Obstetrics and Gynecology Clinics of North America* **24**:411–440 DOI 10.1016/S0889-8545(05)70311-9.

Justice RW, Zilian O, Woods DF, Noll M, Bryant PJ. 1995. The Drosophila tumor suppressor gene warts encodes a homolog of human myotonic dystrophy kinase and is required for the control of cell shape and proliferation. *Genes & Development* **9**:534–546 DOI 10.1101/gad.9.5.534.

Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* **28**:27–30 DOI 10.1093/nar/28.1.27.

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**:559 DOI 10.1186/1471-2105-9-559.

Leyendecker G, Herbertz M, Kunz G, Mall G. 2002. Endometriosis results from the dislocation of basal endometrium. *Human Reproduction* **17**:2725–2736 DOI 10.1093/humrep/17.10.2725.

Li C, Kan L, Chen Y, Zheng X, Li W, Zhang W, Cao L, Lin X, Ji S, Huang S, Zhang G, Liu X, Tao Y, Wu S, Chen D. 2015a. Ci antagonizes Hippo signaling in the somatic cells of the ovary to drive germline stem cell differentiation. *Cell Research* **25**:1152–1170 DOI 10.1038/cr.2015.114.

Li L, Qu YW, Li YP. 2017. Over-expression of miR-1271 inhibits endometrial cancer cells proliferation and induces cell apoptosis by targeting CDK1. *European Review for Medical and Pharmacological Sciences* **21**:2816–2822.

Li X, Tao J, Cigliano A, Sini M, Calderaro J, Azoulay D, Wang C, Liu Y, Jiang L, Evert K, Demartis MI, Ribback S, Utpatel K, Dombrowski F, Evert M, Calvisi DF, Chen X. 2015b. Co-activation of PIK3CA and Yap promotes development of hepatocellular and cholangiocellular tumors in mouse and human liver. *Oncotarget* **6**:10102–10115 DOI 10.18632/oncotarget.3546.

Lin A, Yin J, Cheng C, Yang Z, Yang H. 2018. Decreased expression of FOXA2 promotes eutopic endometrial cell proliferation and migration in patients with endometriosis. *Reproductive Biomedicine Online* **36**:181–187 DOI 10.1016/j.rbmo.2017.11.001.

Luo Y, Wu Y, Peng Y, Liu X, Bie J, Li S. 2016. Systematic analysis to identify a key role of CDK1 in mediating gene interaction networks in cervical cancer development. *Irish Journal of Medical Science* **185**:231–239 DOI 10.1007/s11845-015-1283-8.

Moriyama K, Hori T. 2019. BCR-ABL induces tyrosine phosphorylation of YAP leading to expression of Survivin and Cyclin D1 in chronic myeloid leukemia cells. *International Journal of Hematology* **110**:591–598 DOI 10.1007/s12185-019-02726-7.

Olive DL, Pritts EA. 2001. Treatment of endometriosis. *The New England Journal of Medicine* **345**:266–275 DOI 10.1056/NEJM200107263450407.

Riccio L, Santulli P, Marcellin L, Abrao MS, Batteux F, Chapron C. 2018. Immunology of endometriosis. *Best Practice & Research Clinical Obstetrics & Gynaecology* **50**:39–49 DOI 10.1016/j.bpobgyn.2018.01.010.

Sampson JA. 1927. Metastatic or embolic endometriosis, due to the menstrual dissemination of endometrial tissue into the venous circulation. *The American Journal of Pathology* **3**:93–110.

Schifrin BS, Erez S, Moore JG. 1973. Teen-age endometriosis. *American Journal of Obstetrics and Gynecology* **116**:973–980 DOI 10.1016/S0002-9378(16)33845-5.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**:2498–2504 DOI 10.1101/gr.1239303.

Shimo A, Tanikawa C, Nishidate T, Lin ML, Matsuda K, Park JH, Ueki T, Ohta T, Hirata K, Fukuda M, Nakamura Y, Katagiri T. 2008. Involvement of kinesin family member 2C/mitotic centromere-associated kinesin overexpression in mammary carcinogenesis. *Cancer Science* **99**:62–70 DOI 10.1111/j.1349-7006.2007.00635.x.

Song Y, Fu J, Zhou M, Xiao L, Feng X, Chen H, Huang W. 2016. Activated hippo/yes-associated protein pathway promotes cell proliferation and anti-apoptosis in endometrial stromal cells of endometriosis. *The Journal of Clinical Endocrinology and Metabolism* **101**:1552–1561 DOI 10.1210/jc.2016-1120.

Starzinski-Powitz A, Zeitvogel A, Schreiner A, Baumann R. 2001. In search of pathogenic mechanisms in endometriosis: the challenge for molecular cell biology. *Current Molecular Medicine* **1**:655–664 DOI 10.2174/1566524013363168.

Sundqvist J, Andersson KL, Scarselli G, Gemzell-Danielsson K, Lalitkumar PG. 2012. Expression of adhesion, attachment and invasion markers in eutopic and ectopic endometrium: a link to the aetiology of endometriosis. *Human Reproduction* **27**:2737–2746 DOI 10.1093/humrep/des220.

Talbi S, Hamilton AE, Vo KC, Tulac S, Overgaard MT, Dosiou C, Le Shay N, Nezhat CN, Kempson R, Lessey BA, Nayak NR, Giudice LC. 2006. Molecular phenotyping of human endometrium distinguishes menstrual cycle phases and underlying biological processes in normo-ovulatory women. *Endocrinology* **147**:1097–1121 DOI 10.1210/en.2005-1076.

Uegaki T, Taniguchi F, Nakamura K, Osaki M, Okada F, Yamamoto O, Harada T. 2015. Inhibitor of apoptosis proteins (IAPs) may be effective therapeutic targets for treating endometriosis. *Human Reproduction* **30**:149–158 DOI 10.1093/humrep/deu288.

Wang R, Kang Y, Lohr CV, Fischer KA, Bradford CS, Johnson G, Dashwood WM, Williams DE, Ho E, Dashwood RH. 2016. Reciprocal regulation of BMF and BIRC5 (Survivin) linked to Eomes overexpression in colorectal cancer. *Cancer Letters* **381**:341–348 DOI 10.1016/j.canlet.2016.08.008.

Wang H, Zhang X, Wang L, Zheng G, Du L, Yang Y, Dong Z, Liu Y, Qu A, Wang C. 2014. Investigation of cell free BIRC5 mRNA as a serum diagnostic and prognostic biomarker for colorectal cancer. *Journal of Surgical Oncology* **109**:574–579 DOI 10.1002/jso.23526.

Wingfield M, Macpherson A, Healy DL, Rogers PA. 1995. Cell proliferation is increased in the endometrium of women with endometriosis. *Fertility and Sterility* **64**:340–346 DOI 10.1016/S0015-0282(16)57733-4.

Xu T, Wang W, Zhang S, Stewart RA, Yu W. 1995. Identifying tumor suppressors in genetic mosaics: the Drosophila lats gene encodes a putative protein kinase. *Development* **121**:1053–1063.

Yang S, Zhang L, Liu M, Chong R, Ding SJ, Chen Y, Dong J. 2013. CDK1 phosphorylation of YAP promotes mitotic defects and cell motility and is essential for neoplastic transformation. *Cancer Research* **73**:6722–6733 DOI 10.1158/0008-5472.CAN-13-2049.

Yim DH, Zhang YW, Eom SY, Moon SI, Yun HY, Song YJ, Youn SJ, Hyun T, Park JS, Kim BS, Lee JY, Kim YD, Kim H. 2013. ITGA1 polymorphisms and haplotypes are associated with gastric cancer risk in a Korean population. *World Journal of Gastroenterology* **19**:5870–5876 DOI 10.3748/wjg.v19.i35.5870.

Yin MX, Zhang L. 2015. Hippo signaling in epithelial stem cells. *Acta Biochimica et Biophysica Sinica* **47**:39–45 DOI 10.1093/abbs/gmu111.

Yotova IY, Quan P, Leditznig N, Beer U, Wenzl R, Tschugguel W. 2011. Abnormal activation of Ras/Raf/MAPK and RhoA/ROCKII signalling pathways in eutopic endometrial stromal cells of patients with endometriosis. *Human Reproduction* **26**:885–897 DOI 10.1093/humrep/der010.

Yu FX, Zhao B, Guan KL. 2015. Hippo pathway in organ size control, tissue homeostasis, and cancer. *Cell* **163**:811–828 DOI 10.1016/j.cell.2015.10.044.

Zhang L, Chen X, Stauffer S, Yang S, Chen Y, Dong J. 2015. CDK1 phosphorylation of TAZ in mitosis inhibits its oncogenic activity. *Oncotarget* **6**:31399–31412 DOI 10.18632/oncotarget.5189.

Zhao Y, Montminy T, Azad T, Lightbody E, Hao Y, SenGupta S, Asselin E, Nicol C, Yang X. 2018. PI3K positively regulates YAP and TAZ in mammary tumorigenesis through multiple signaling pathways. *Molecular Cancer Research* **16**:1046–1058 DOI 10.1158/1541-7786.MCR-17-0593.