



OPEN ACCESS

EDITED BY

Zhiqiang Liu,
Shandong University, China

REVIEWED BY

Fatima Maqoud,
University of Bari Aldo Moro, Italy
Nicolae Bacalbasa,
Carol Davila University of Medicine and
Pharmacy, Romania

*CORRESPONDENCE

Lei Zhang

✉ leizhanggyn@outlook.com

Ruiheng Zhao

✉ kafuka1026@163.com

†These authors have contributed equally to
this work

RECEIVED 18 February 2024

ACCEPTED 28 October 2024

PUBLISHED 22 November 2024

CITATION

Ma R, Zheng Y, Wang J, Xu H, Zhang R,
Xie Z, Zhang L and Zhao R (2024)
Identification of key genes associated with
endometriosis and endometrial cancer by
bioinformatics analysis.
Front. Oncol. 14:1387860.
doi: 10.3389/fonc.2024.1387860

COPYRIGHT

© 2024 Ma, Zheng, Wang, Xu, Zhang, Xie,
Zhang and Zhao. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Identification of key genes associated with endometriosis and endometrial cancer by bioinformatics analysis

Ruyue Ma[†], Yu Zheng[†], Jianing Wang[†], Hong Xu, Ruirui Zhang, Zhijia Xie, Lei Zhang* and Ruiheng Zhao*

Department of Obstetrics and Gynecology, Suzhou Ninth People's Hospital, Suzhou Ninth Hospital Affiliated to Soochow University, Suzhou, China

Background: Endometriosis (EMS) is acknowledged as a risk factor for the development of endometrial cancer (EC), although the precise molecular mechanisms that underpin this association have yet to be fully elucidated. The primary objective of this investigation is to harness bioinformatics methodologies to identify pivotal genes and pathways that may be implicated in both EMS and EC, potentially offering novel therapeutic biomarkers for the management of endometriosis.

Methods: We acquired four datasets pertaining to EMS and one dataset concerning EC from the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) in EMS and EC cohorts, in comparison to controls, were ascertained utilizing the limma package. Subsequently, we conducted a series of bioinformatic analyses, including Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, and protein-protein interaction (PPI) analysis, to delineate pathways associated with the identified DEGs.

Results: Our bioinformatics analyses disclosed 141 shared DEGs between EMS and EC groups relative to the control cohort. GO analysis demonstrated that these genes are predominantly involved in the regulation of growth and development, as well as signal transduction pathways. KEGG analysis underscored the significance of these genes in relation to the JAK-STAT signaling pathway and leukocyte transendothelial migration. Furthermore, PPI analysis pinpointed ten central genes (APOE, FGF9, TIMP1, BGN, C1QB, MX1, SIGLEC1, BST2, ICAM1, MME) exhibiting high interconnectivity. Notably, the expression levels of APOE, BGN, C1QB, and BST2 were found to correlate with cancer genomic atlas data, and were implicated in tumor immune infiltration. Strikingly, only APOE and BGN demonstrated a significant correlation with patient prognosis.

Conclusion: This comprehensive bioinformatics analysis has successfully identified key genes that may serve as potential biomarkers for EC. These findings significantly enhance our comprehension of the molecular underpinnings of EC pathogenesis and prognosis, and hold promise for the identification of novel drug targets.

KEYWORDS

bioinformatics, biomarker, gene expression, endometriosis, endometrial cancer

Introduction

Endometriosis (EMS), characterized by the presence of endometrioid tissue outside the uterine cavity, affects approximately 5% to 10% of women of reproductive age worldwide and is a leading cause of dysmenorrhea, infertility, and chronic pelvic pain (1). EMS is a benign lesion that causes pelvic tissue adhesion, forming nodules and masses, and is characterized by its dependence on sex hormones. The malignant manifestations of EMS are similar to those of tumors, with lesions that are extensive, diverse, highly invasive, and recurrent.

Endometrial cancer (EC) is the seventh most common cancer among women worldwide and is one of the most prevalent gynecological tumors (2). With an average annual growth rate of 1.9%, EC significantly impacts the quality of life for patients and places a substantial burden on healthcare resources (3). Recent studies have indicated that EC is more prevalent in individuals with EMS than in the general population, and EMS is an independent risk factor for EC prognosis. This suggests that patients with EMS may have a higher risk of developing EC (4, 5). However, the molecular mechanisms underlying the association between EMS and EC are not well understood.

Patients with EMS and EC share similar pathological mechanisms, including genomic alterations, inflammation, stem cells, immunogenicity, and estrogen dependence (5, 6). An increasing number of studies suggest that endometriosis is a systemic inflammatory disease. The concentration of cytokines in endometriosis is abnormally elevated, including IL-1 β , IL-6, IL-28, tumor necrosis factor, CCL2, CCL5, and VEGF (7, 8), which are related to the progression of EMS and the development of EC (9–11). Research has found that immune cells are involved in the pathogenesis of EMS, with an increased number of lymphocytes in peritoneal fluid and a higher concentration of peritoneal macrophages. Additionally, an imbalance between Th1/Th2 cells contributes to the progression of the disease (12, 13). Therefore, we consider whether there are key genes and pathways that underlie common molecular mechanisms in the development of these two diseases. Screening differentially expressed genes (DEGs) between endometrial cancer and endometriosis may offer an alternative approach to understanding the mechanisms of carcinogenesis and recurrence in endometriosis, as well as constructing new potential serum biomarkers for the early diagnosis of EC patients.

Bioinformatics analysis integrates and analyzes biological data through various bioinformatics tools, which is one of the crucial means of life science research. To discover the pathogenesis of EMS related to EC, we collected three EMS datasets and one EC dataset from the Gene Expression Omnibus (GEO) database using various comprehensive bioinformatics tools. We identified hub genes associated with EMS in EC, predicted common molecular pathogenesis, and potential therapeutic drugs. We validated the expression patterns of hub genes in EC. Finally, we explored the immune cell characteristics of EC to reveal the association of hub genes with the immune system, which may provide a new direction for treatment.

Materials and methods

Data collection

In this study, data sets are from the National Center for Biotechnology Information (NCBI) pool of gene expression (GEO, <https://www.ncbi.nlm.nih.gov/geo/>). The data in the GEO database we used contained preprocessed and normalized data, which is suitable for direct use in downstream analysis. Three endometriosis datasets, GSE7305, GSE23339, and GSE25628, and one endometrial cancer dataset, GSE17025, were used. The GSE7305 dataset contains 10 ectopic endometrium and 10 samples of normal endometrium. GSE23339 contains 10 cases of ectopic endometrium and 9 cases of normal endometrium. GSE25628 contains 7 cases of ectopic endometrium, 9 cases of eutopic endometrium, and 6 cases of normal endometrium. GSE17025 contains 12 cases of normal endometrium and 91 cases of tumor. The downloaded data were processed using the R package limma followed by calibration, standardization, and log₂ transformation, and details about the dataset are summarized in Table 1.

Analysis of differentially expressed genes

R Studio software was utilized to process and standardize the files. Four GEO datasets were preprocessed. For the genes corresponding to multiple probes, the average value was taken as the expression level of the gene, and the expression levels of all genes were converted into log₂ (expression level). The “limma” R package was applied to analyze the differences between the experimental and control groups. The threshold of difference analysis was set as p-value <0.05 and FC >1.5. Overlapping DEGs from the four database screenings were utilized in subsequent GO enrichment, KEGG pathway analysis, and protein-protein interaction (PPI) analyses. We then used Venn software online (<https://jvenn.toulouse.inrae.fr/app/example.html>) to visualize the common DEGs.

Gene ontology and pathway discovery based on gene enrichment analysis

Enrichment analysis of gene sets provides unique biological properties of genes, including biological processes, cellular

TABLE 1 Details of GEO data used in this study.

GEO numbers	Experimental group	Control group	PMID
GSE7305	10	10	17640886
GSE17025	91	12	21619611
GSE23339	10	9	21436257
GSE25628	16	6	23460397

components, and molecular functions. KEGG pathway enrichment analysis explores the critical pathways involved in disease occurrence and progression. KEGG and GO enrichment analysis of differential genes was conducted with the R package clusterProfiler, with a p-value threshold of less than 0.05. We considered the input dataset to enrich the pathway or term if it met this criterion.

PPI network and Hub gene acquisition

The activity of protein-protein interactions (PPIs) is regarded as a significant target for cell biology research and is vital for elucidating key genes and necessary gene modules during cancer development. Common DEGs (version 11.0, <https://string-db.org/>) were inputted into STRING, a search tool for retrieving interacting genes, to generate PPI networks with a medium confidence score of 0.400. The STRING database comprises electronically inferred interaction data to ensure the most comprehensive coverage of potential protein interactions possible. The PPIs obtained through the analysis with Cytoscape software (<https://cytoscape.org/>) were visualized and analyzed using this powerful bioinformatics software. Additionally, the Cytoscape plugin CytoHubba (<https://apps.cytoscape.org/apps/cytohubba>) was employed to extract Hub genes, and the MCC function of CytoHubba was used to identify the top 10 Hub genes from the PPI network.

Expression verification of Hub genes

To further select precise biomarkers, we performed Gene Expression Profiling (GEPIA, <http://gepia.cancer-pku.cn/>). This web-based tool provides fast and customizable functionality to display gene expression profiles from the TCGA database while performing survival analysis of genes. To explore the survival differences between different groups. Immunohistochemical (IHC) data for common genes were retrieved in the Human Protein Atlas (HPA) project, which contains detailed information about samples.

TF-gene interactions

NetworkAnalyst (<https://www.networkanalyst.ca/>) was employed to identify genes, and TF has identified common genetic interactions, the research of transcription factors, and regulation networks between the target genes. NetworkAnalyst is a comprehensive network platform used to perform gene expression on many species, enabling them to be subjected to meta-analysis. TF-gene interaction network of network platform from NetworkAnalyst ENCODE (<https://www.encodeproject.org/>) contained in the database. The ENCODE database provides abundant information on transcription factor binding sites, which can provide strong support for the study of TF gene interactions.

TF miRNA co-regulation network

TF miRNA co-regulation interaction refers to the complex interaction among Transcription factors, miRNAs, and their target genes. The interaction of TF miRNA co-regulation was collected from the RegNetwork reservoir. Transcription factors are proteins that can bind to specific sequences of DNA, and miRNAs are small RNA molecules that regulate the transcription and translation process of genes by binding to the miRNA of target genes, thereby affecting the expression level of genes. NetworkAnalyst was used to visualize the TF miRNA co-regulation network.

Immune infiltration analysis of hub genes

Tumor immunity infiltration online database (TIMER, <https://cistrome.shinyapps.io/timer/>) is applied to analyze different types of cancer immunity infiltration. TIMER can analyze the relationship between the cancer immune cell infiltration level and hub genes. The relationship between somatic copy number alterations (SCNA) of potential prognostic hub genes and infiltrating immune cells was explored through related modules. A cutoff value of $P < 0.05$ was applied as the threshold.

Identification of candidate drugs

Drug molecular identification is a key component of research, and the drug-Gene Interaction Database (DGIdb, <http://www.dgldb.org>) is a web-based resource that weaves the genome of druggable genomes into known drug interactions and potential druggable targets. We input the co-expressed DEGs and evaluate the DEGs in DGIdb to find potential therapeutic drugs.

RNA isolation and qRT-PCR

The human endometrial cancer cell line (HEC - 1B and Ishikawa) and the normal endometrial epithelial cell line (HEEC) were both purchased from ATCC. These cell lines were cultured in a culture flask containing DMEM medium supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin in an incubator set at 37°C with 5% CO₂. Cells from the endometrial cancer cell lines and the normal endometrial cell line, both in the logarithmic growth phase, were collected separately. Total RNA was isolated from the cells following the manufacturer's instructions using the PureLink RNA Mini Kit (Invitrogen, Carlsbad, CA, USA), and the RNA was reverse transcribed using the PrimeScript RT Kit with gDNA Eraser (Takara, Dalian, China). Subsequently, RT-qPCR amplification was conducted using PowerUp SYBR Green Master Mix (Thermo) according to the manufacturer's protocol, with GAPDH serving as the reference gene. The PCR protocol consisted of an initial denaturation at 95°C for 15 seconds followed by 40 cycles at 95°C for 5 seconds, and 60°C for 1 minute. The primer sequences are presented in Table 2. Each sample was prepared in triplicate.

TABLE 2 The primers used for RT-qPCR.

Target gene	Forward primer	Reverse primer
AOPE	5'-GTTGCTGGTCAC ATTCTGG-3'	5'-GCAGGTAATC CCAAAAGCGAC-3'
BGN	5'-TGACTGGCATC CCCAAAGAC-3'	5'-GAGTAGCGAA GCAGGTCCTC-3'
BST1	5'-GGAGGAGCTTG AGGGAGAG-3'	5'-CTCAGTCGCT CCACCTCTG-3'
CIQB	5'-GCTCCTGGGCT AATCGATA-3'	5'-GCCGACTTTCC TGGATTCC-3'
GAPDH	5'- TCCTCTGACTCAAC AGCGACAC-3'	5'- CACCTGTTGCTGT AGCCAAATTC-3'

Results

Identification of DEGs and common genes between EMS and EC

All four datasets were analyzed for gene expression using the R limma package, and genes with P values <0.05 and Log [FC]>1.5 were designated as DEGs. As depicted in the volcano plot of Figure 1, specific information about the four datasets is summarized in Table 1.

In the GSE7305 dataset, 2595 DEGs were identified, including 1278 upregulated and 1317 downregulated genes. In GSE2339, 2171 DEGs were identified, with 1016 upregulated and 1155 downregulated genes. In GSE25628, 3842 DEGs were identified, comprising 1819 upregulated and 2023 downregulated genes. In GSE17025, 4790 DEGs were identified, including 2246 upregulated and 2544 downregulated genes (Table 3). The DEGs for each dataset were visualized using volcano plots (Figures 1A–D). After standardization, we identified 141 differentially expressed genes (Figure 1E) by comparing the DEGs present across the four datasets, with 56 differentially expressed genes being consistent. Among these, 33 were upregulated and 23 were downregulated (Table 4). The top 40 DEGs from all four datasets were visualized on a heat map (Figures 2A–D).

Enrichment analysis of differential genes

KEGG and GO analyses were conducted on the 141 DEGs to explore the biological functions of the integrated DEGs. In Biological Process (BP), DEGs were significantly enriched in processes related to growth and development, regulation of the Wnt signaling pathway, and proteoglycan metabolism (Figure 3A). In Cellular Component (CC), DEGs were primarily involved in the collagen-rich extracellular matrix, neuronal cell bodies, and membrane structures (Figure 3B). Molecular Function (MF) analysis revealed significant enrichment in

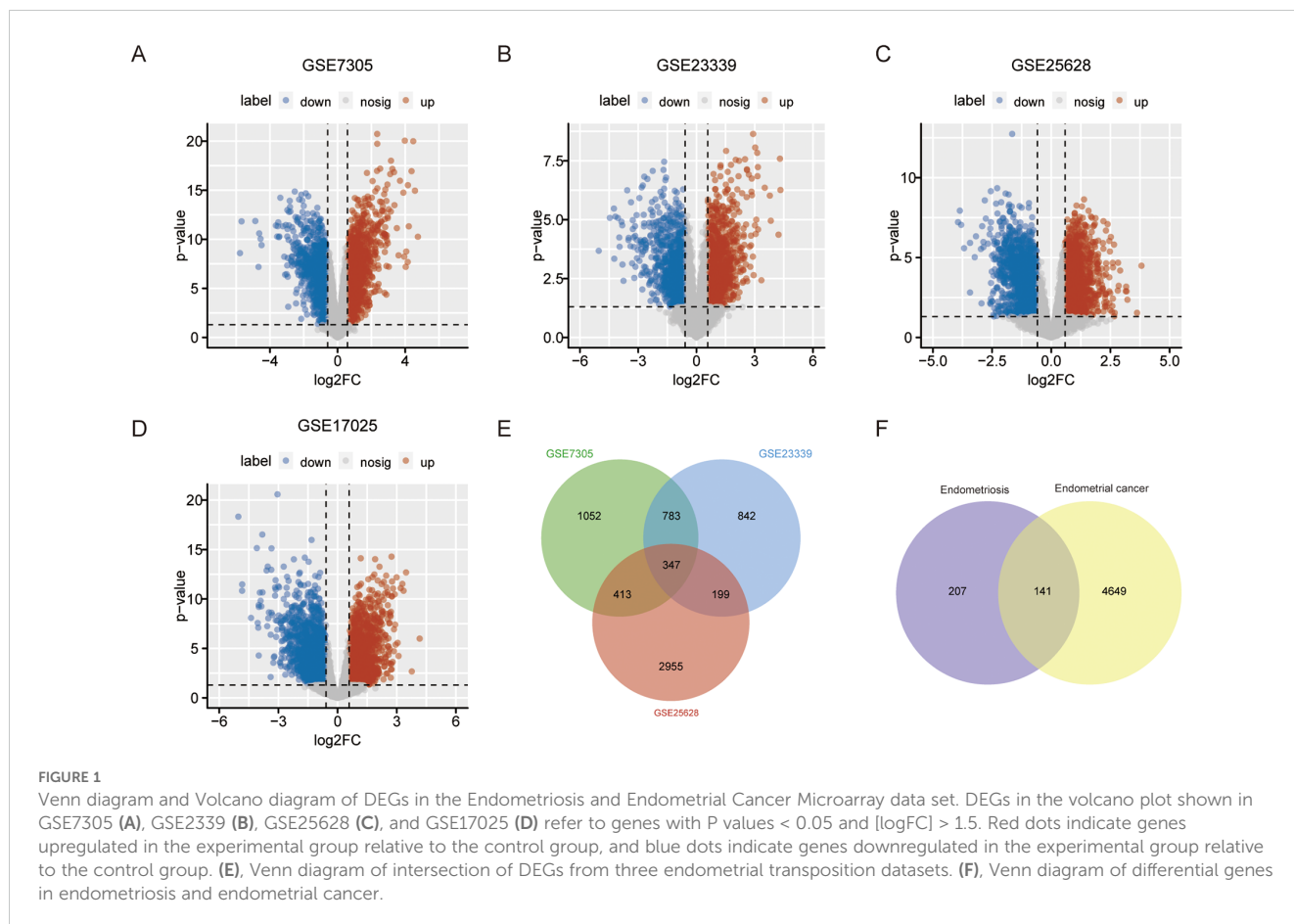


TABLE 3 The number of up-regulated and down-regulated genes and the number of differentially expressed genes in the differential gene analysis of each expression profile.

GEO No.	Upregulated genes	Downregulated genes	Differential genes
GSE7305	1278	1317	2595
GSE17025	2246	2544	4790
GSE23339	1016	1155	2171
GSE25628	1819	2023	3842

glycosaminoglycan binding, cholesterol, and lipid transfer activities (Figure 3C). KEGG pathway analysis indicated that DEGs were mainly enriched in the JAK-STAT signaling pathway and leukocyte migration through endothelial cells (Figure 3D). Significantly enriched terms were displayed in bubble plots using the R package, with a criterion of $P < 0.05$.

PPI network construction and hub gene selection

The 56 DEGs were uploaded to the STRING database (<https://string-db.org/cgi/>) with a confidence score threshold of >0.4 , and the resulting PPI network was visualized using Cytoscape software. The PPI network comprised 56 nodes and 17 edges (Figure 4A), and the CytoHubba plugin identified the hub genes within the PPI

network. Using the MCC algorithm for sorting, the top 10 hub genes selected were APOE, FGF9, TIMP1, BGN, C1QB, MX1, SIGLEC1, BST2, ICAM1, and MME (Figure 4B).

TF-gene interaction and TF-miRNA co-regulatory network

TF-gene interactions were collected using NetworkAnalyst. For the hub genes (APOE, FGF9, TIMP1, BGN, C1QB, MX1, SIGLEC1, BST2, ICAM1, MME), TF-gene interactions were analyzed. The interactions between TFs and common DEGs are shown in Figure 5A. The network contains 130 nodes and 171 edges. Among them, ICAM1 is regulated by 69 TFs, and APOE and BST2 are each regulated by 29 TFs. These TFs regulate more than one interaction in the network with a common DEG. TF-miRNA co-regulatory network analysis reveals the interactions of miRNAs and TFs with common DEGs, which may be responsible for regulating the expression of DEGs. The network created by the TF-miRNA co-regulatory analysis consisted of 235 nodes and 307 edges. Figure 5B shows the co-regulatory network of TF-miRNAs.

Validation of hub genes in the TCGA database

The top 10 hub genes were used for subsequent analyses. The GEPIA database was used to investigate differences in gene

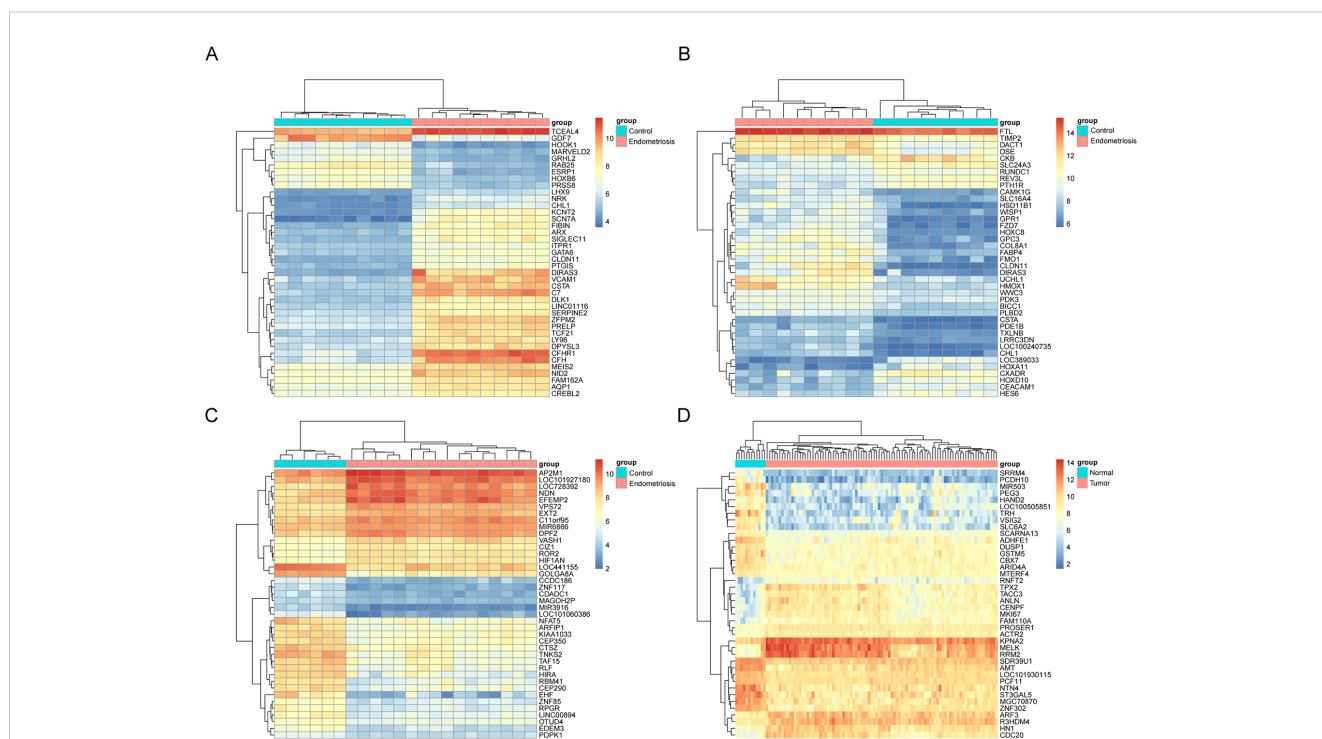


FIGURE 2 Heat map of the top 40 upregulated and downregulated degs in the microarray dataset. Top 40 DEGs in microarray datasets GSE7305 (A), GSE2339 (B), GSE25628 (C), and GSE17025 (D), DEGs refer to genes with P values <0.05 and $\text{Log}[\text{FC}] > 1.5$. Red indicates upregulation and blue indicates downregulation. Each column represents a GEO ID, and each row represents a gene name.

TABLE 4 DEGs co-expressed in the four data sets.

Category	Gene symbol
Upregulated	BST2 RPP25 HOXC6 FHL2 PAPSS2 CDH3 TIMP1 IL4R CHST15 RNASE1 QSOX1 PLTP HS3ST3A1 ST6GALNAC5 GGT5 BGN ENO2 SIGLEC1 RNASE6 ICAM1 NCF4 SPSB1 HCLS1 MX1 ADAMTSL2 ENC1 APOE OSMR C1QB TNFAIP3 SULF1 THBD PIM1
Downregulated	GABRP P2RY14 MAP3K1 REV3L ARSJ COBL MME ACSL CTNNA2 TOM1L1 HEY2 GRAMD1C SEMA5A CWH43 IL20RA TRH PGGHG ADGRG2 CRYZ ANXA3 FAM189A2 SEMA3E FGF9
Inconsistency	ESRP1 NXN PAMR1 CLDN7 CD248 TFCP2L1 RAMP1 CD24 DSP HN1L PRRG4 TSPAN13 MAP7 STIL NPR2 SCARA3 PRR15L PLS1 SPAG1 HYOU1 MYO6 TMEM158 FZD5 SFN SLC25A1 MYLIP CDK2AP2 FBLN1 KRT18 KLF5 PRKX TCN1 TCF21 GATA6 PRELP TCEAL2 BNC2 PDLIM3 ITM2A PLN FBXL7 JADE2 KLF2 PDGFRL ROBO3 NEK7 TAGLN FMOD ASAP3 CLDN5 CYBRD1 TSPAN4 EFEMP1 WISP1 PLOD2 CTSF RARRES2 LIMS2 OLFML1 ACTA2 THBS4 MYL9 DCN NSG1SERPINF1 CMAHP ABI3BP COLEC12 ZNF395 EMILIN1 THBS1 COL16A1 JAM2 GLT8D2 CPQ CXCL12 MORC3 ACOX2 BST1 GFPT2 CNTN1 NR4A1 MFAP4 ADAMDEC1 GADD45B

expression levels between early tumor tissues and normal tissues (Figure 6A). The results showed that four hub genes, APOE, BGN, BST2, and C1QB, were abnormally differentially expressed in normal tissues and endometrial cancer. Additionally, the

immunohistochemical (IHC) staining extracted from the Human Protein Atlas Database (HPA) also showed differential protein expression levels of hub genes, as shown in Figure 6B. Among the 10 hub genes, only the down-regulated genes APOE and BGN were associated with EC prognosis, confirming the reliability of the hub genes we identified (Figure 6C).

Immune infiltration analysis of hub genes

The TIMER database was used to analyze the correlation between mRNA expression of APOE, BGN, BST2, and C1QB and infiltrating immune cells in cancer tissues. APOE was correlated with the abundance of B cells (cor=0.419, P=1.00e-13), CD4+ T cells (cor=0.458, P=1.86e-16), macrophages (cor=0.332, P=6.25e-09), and neutrophils (cor=0.246, P=0.005). Dendritic cell abundance (cor=0.258, P=7.66e-06) also showed significant correlation (Figure 7A). BGN showed a significant correlation with the abundance of CD4+ T cells (cor=0.173, P=3.08e-03) (Figure 7B). BST2 was significantly correlated with the abundance of B cells (cor=0.136, P=2.09e-02), CD8+ T cells (cor=-0.238, P=4.28e-05), CD4+ T cells (cor=0.181, P=1.98e-03), neutrophils (cor=0.17, P=0.017), and CD4+ T cells (cor=0.17, P=0.016) (Figure 7C). C1QB was significantly correlated with B cells (cor=0.633, P=1.10e-33), CD8+ T cells (cor=0.255, P=1.15e-05), CD4+ T cells (cor=0.502, P=5.06e-20), macrophages (cor=0.464,

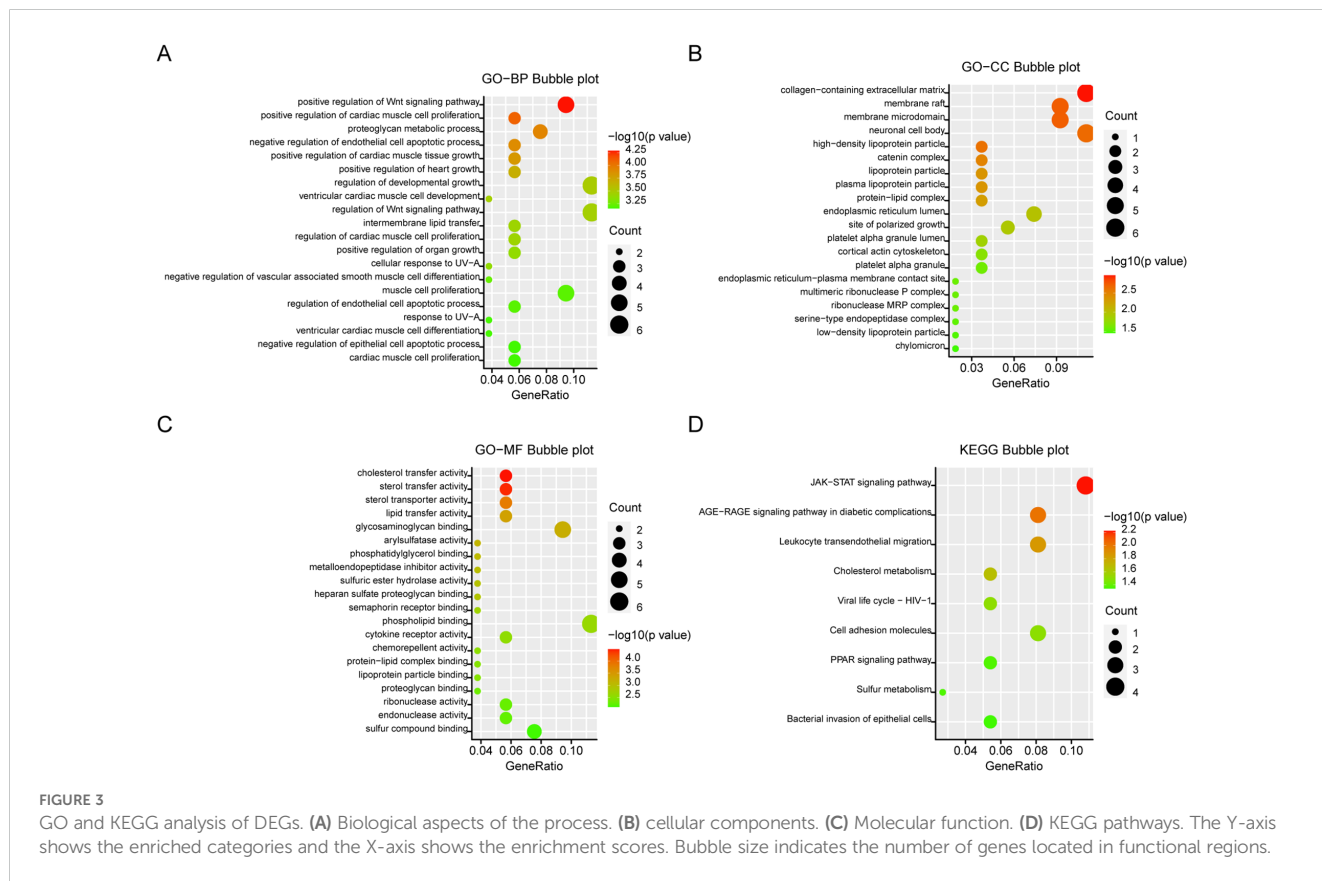


FIGURE 3 GO and KEGG analysis of DEGs. (A) Biological aspects of the process. (B) cellular components. (C) Molecular function. (D) KEGG pathways. The Y-axis shows the enriched categories and the X-axis shows the enrichment scores. Bubble size indicates the number of genes located in functional regions.

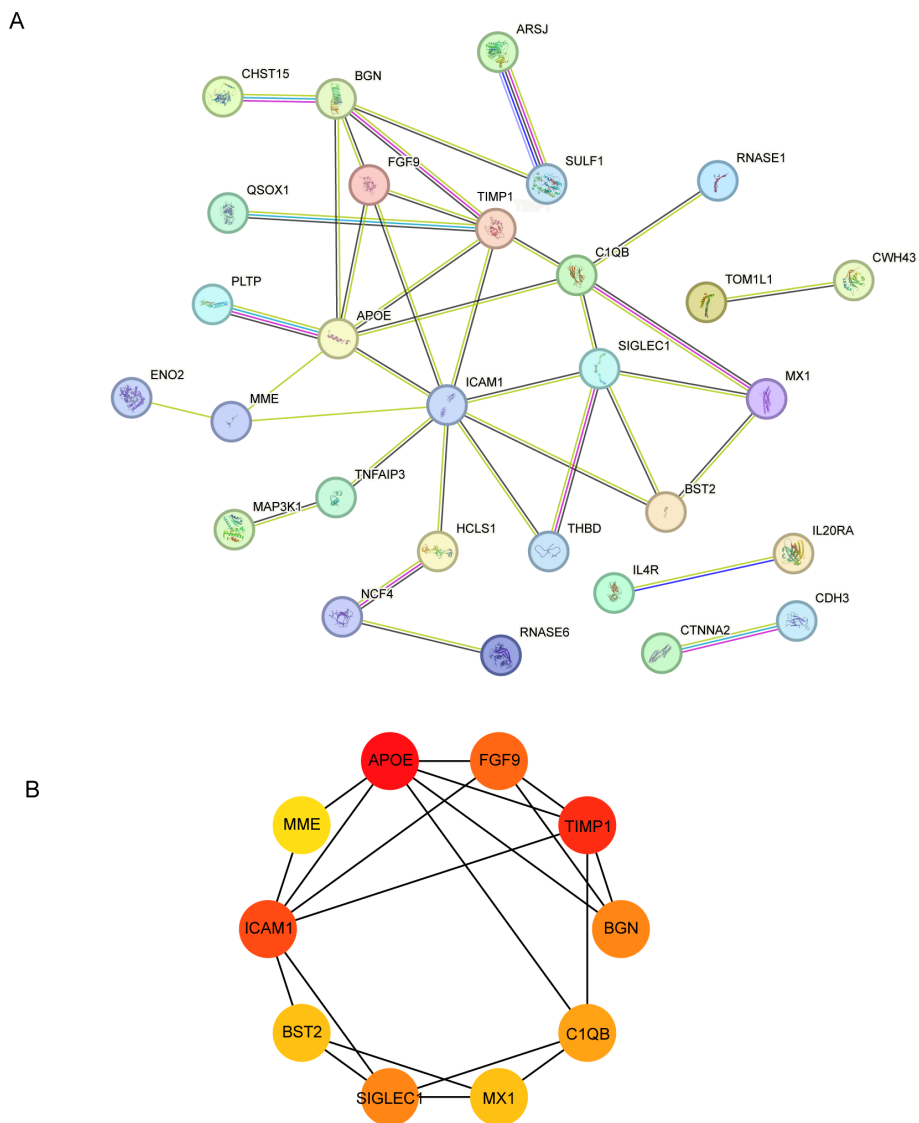


FIGURE 4
PPI network of DEGs and interaction network of top 10 hub genes. **(A)** PPI network of hub genes constructed from the STRING online database. **(B)** the top 10 hub genes were examined from the PPIs network of common differentially expressed genes, and based on the topological analysis of these 10 genes, the degree of ICAM1, APOE and TIMP1 was 7,6,5.

$P=0.005$), neutrophils ($cor=0.459$, $P=1.21e-16$), and dendritic cells ($cor=0.624$, $P=7.24e-36$). These results provide strong evidence that APOE, BGN, BST2, and C1QB play crucial roles in infiltrating immune cells, including B cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils, and dendritic cells.

Candidate drug identification

We entered the 10 most significantly expressed DEGs into the DGIdb database to identify drug-gene interactions and potential drug targets. APOE was associated with 17 drugs for treating endometriosis, ICAM1 with 5 drugs, and MME with 8 drugs for treating endometriosis (Table 5). The potential applicability of these

drugs as targets awaits further investigation through future cellular and animal studies.

In vitro validation of qRT-PCR

To further validate our bioinformatics findings, we performed *in vitro* experiments. To further validate our bioinformatics findings, we conducted *in vitro* experiments. Endometrial epithelial cells HEEC were used as a control. We confirmed by qRT-PCR that the mRNA expression levels of APOE, BGN, BST1, and C1QB in endometrial cancer cell lines (HEC-1B and Ishikawa) were consistent with our previous results. The mRNA expression levels of APOE, BST1, and C1QB in endometrial cancer cell lines were significantly higher than

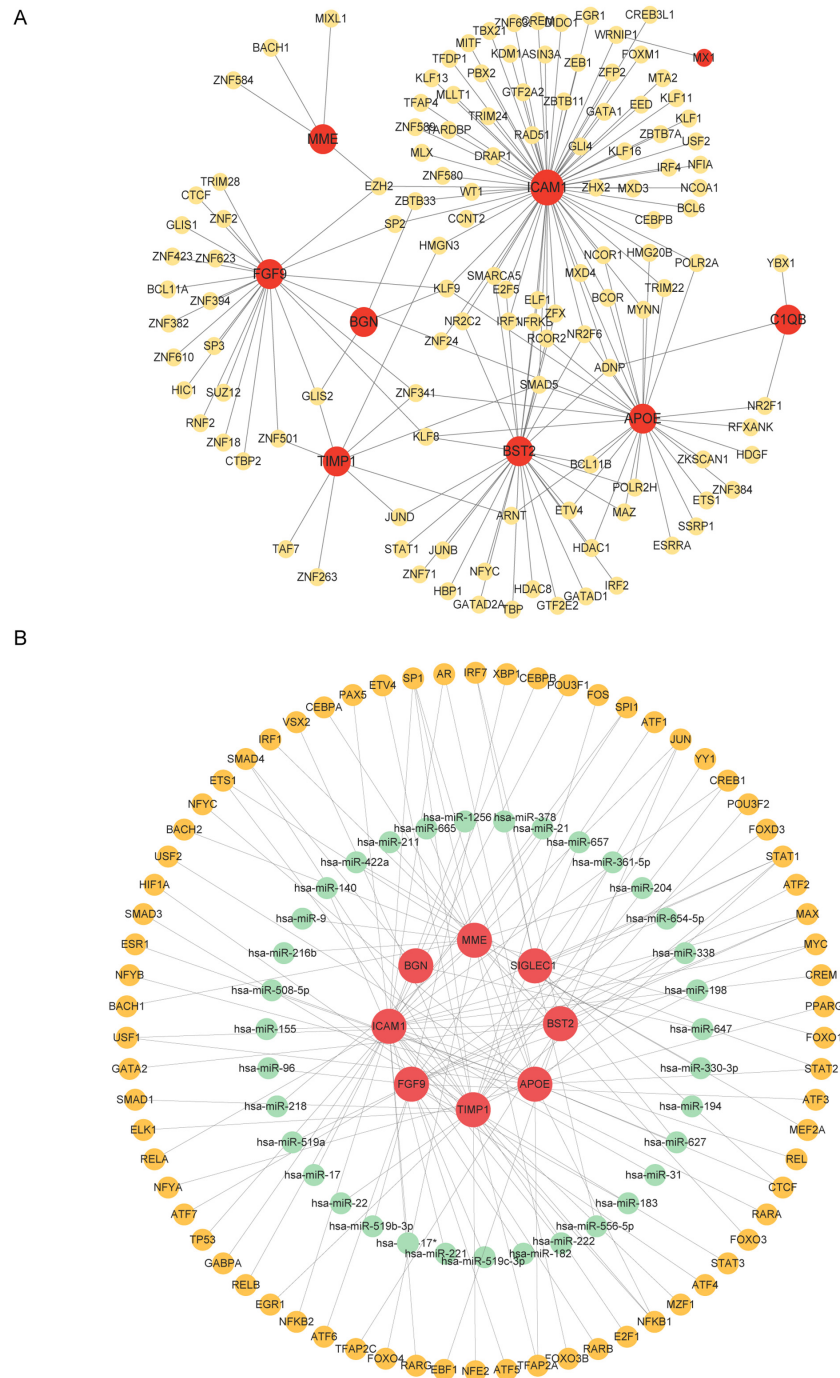


FIGURE 5
 Network of TF genes interacting with common DEGs and TF-miRNA co-regulatory network. **(A)** Highlighted red nodes represent common genes and other nodes represent tf genes. **(B)** The TF-miRNA co-regulatory network consists of 235 nodes and 307 pieces, which include 70 TF-genes, 37 miRNAs, and 8 DEGs. The red color is the DEGs, the green node represents the miRNA, and the other nodes represent the TF-genes.

those in normal endometrial epithelial cells, whereas BGN expression was higher in normal endometrial cell lines than in endometrial cancer cell lines (Figures 8A–D).

Discussion

Endometriosis is a hormone-dependent chronic inflammatory disease, the exact cause of which remains unclear. It is influenced by

retrograde menstruation, coelomic metaplasia, lymphovascular dissemination, as well as genetic and environmental factors (14). Endometriosis is a risk factor for endometrial cancer (13, 15), and the two share common etiological mechanisms, including estrogen stimulation and chronic inflammation (16, 17). Therefore, in this study, we adopted bioinformatics-related methods and utilized the similarities between endometriosis and endometrial cancer to search for targets for the treatment of endometriosis.

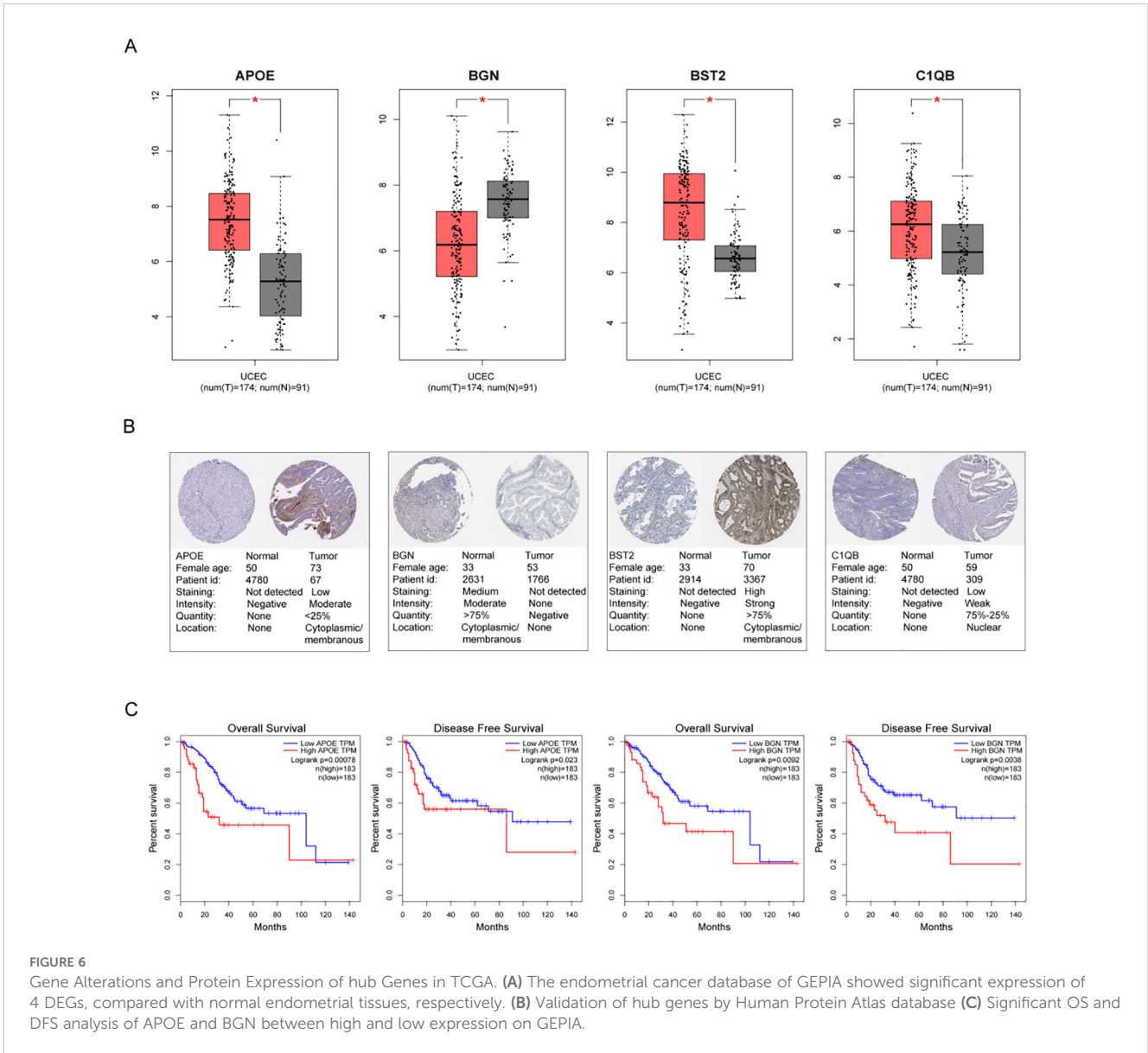


FIGURE 6 Gene Alterations and Protein Expression of hub Genes in TCGA. **(A)** The endometrial cancer database of GEPIA showed significant expression of 4 DEGs, compared with normal endometrial tissues, respectively. **(B)** Validation of hub genes by Human Protein Atlas database **(C)** Significant OS and DFS analysis of APOE and BGN between high and low expression on GEPIA.

By intersecting the datasets of endometriosis and endometrial cancer, we obtained 114 commonly expressed differential genes. Through analyzing the GO functions of DEGs, we found that DEGs were significantly enriched in the regulation of the wnt signaling pathway, growth and development, collagen-related extracellular matrix, and glycosaminoglycan binding. Among them, the abnormality of growth and development-related genes may disrupt the normal physiological functions of endometrial tissue, and the dysregulation of the wnt signaling pathway may lead to endometrial hyperplasia and then trigger endometrial cancer (18). In many solid tumors, due to the massive production of extracellular matrix (ECM) proteins such as collagen by stromal cells, a desmoplastic reaction occurs, resulting in the hardening of the tumor tissue. This fibrotic reaction can inhibit drug delivery and enhance tumor progression and metastasis (19), and fibrotic characteristics have also been found in endometriosis (20, 21). Glycosaminoglycan (GAG) is an important component of the tumor microenvironment (TME). GAG can interact

with multiple binding partners, thus affecting cancer progression at multiple levels (22).

Analysis of DEGs in KEGG pathways was conducted to identify the common pathways involved in endometriosis and endometrial cancer. The results showed that the main JAK - STAT signaling pathway and leukocyte transendothelial migration were involved. The inflammation caused by endometriosis is considered to be related to the occurrence of endometrial cancer. The expression of inflammatory mediators may activate signal pathways that promote malignancy in cancer cells and cancer - related inflammatory cells (23, 24). Many inflammatory cytokines transmit their information through the JAK/STAT pathway, which is crucial for immune cell signaling (25). This signaling pathway is involved in the proliferation, migration, and invasion of endometrial cancer cells (2, 26). Recent studies have also shown that the JAK/STAT3 pathway is upregulated in endometriosis and can be used as a treatment target for endometrial cancer (27).

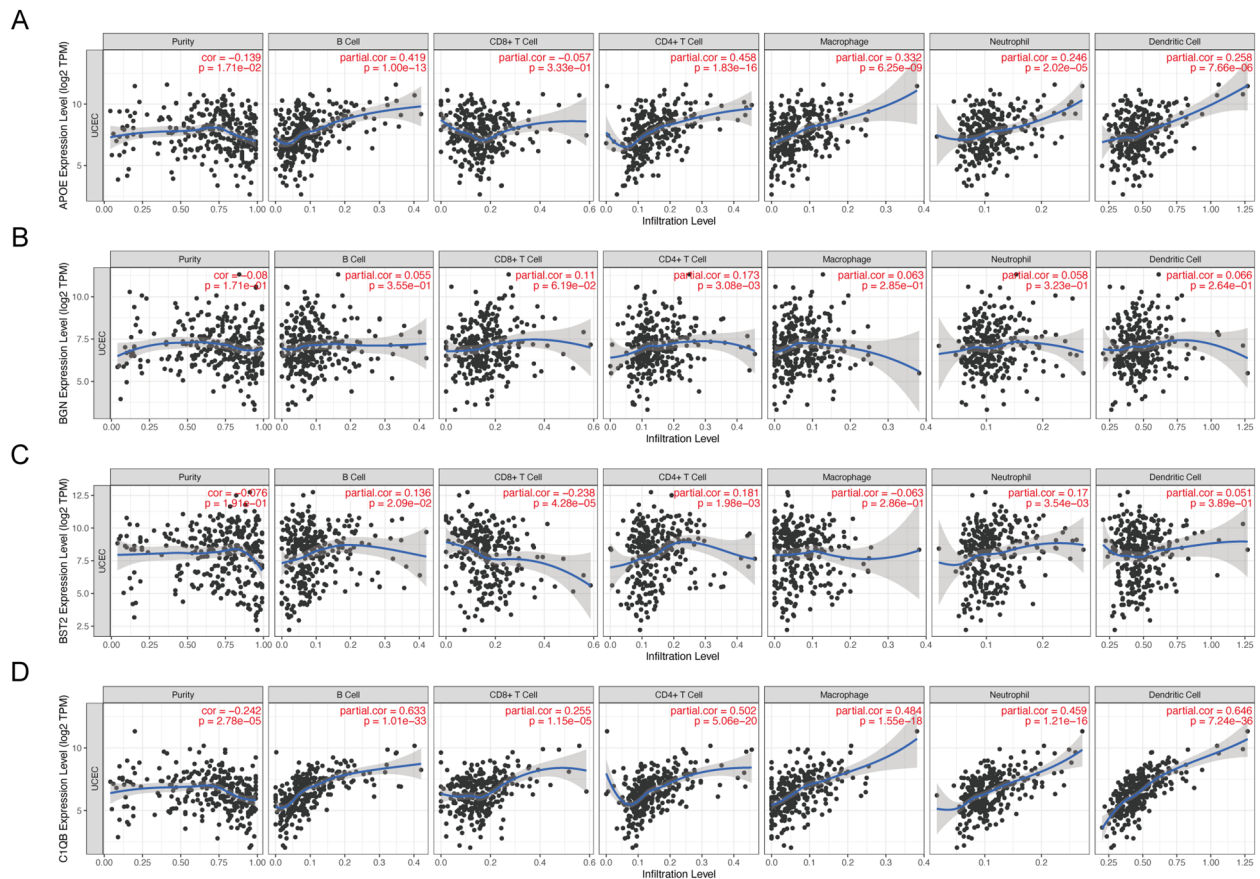


FIGURE 7 Correlation between hub gene expression and immune cell infiltration level in 4 endometrial carcinomas. APOE (A), BGN (B), BST2 (C), and C1QB (D) expression correlated with tumor purity and infiltration levels of B cells, CD8+T cells, CD4+T cells, macrophages, neutrophils, and dendritic cells in endometrial cancer.

We verified the expression of hub genes through GEPIA and HPA databases. Among them, APOE, BGN, BST1, and BGN showed expression differences in endometrial cancer. Survival analysis indicated that only APOE and BGN were related to the overall survival (OS) and disease-free survival (DFS) of endometrial cancer patients, suggesting that they can be used as biomarkers for predicting patient survival. Additionally, all four genes are involved in regulating cancer cell proliferation, invasion, and migration, which is of great significance in the research of gynecological and obstetrical related tumors. APOE exerts biological functions during certain female fertility disorders (endometriosis) and other gynecological diseases (such as breast cancer, choriocarcinoma, endometrial adenocarcinoma/hyperplasia, and ovarian cancer) (28, 29). The expression level of APOE is related to the histological grade, lymph node metastasis, and FIGO stage in endometrial cancer (30) and stimulates the malignant progression of ovarian cancer by inducing FAK - ERK activation in cell/matrix adhesion (31). BGN enhances the migration and invasion abilities of endometrial cancer (32). Previous studies have found that the expression of BGN is associated with the presence of PTEN loss and KRAS mutation (33, 34). BST2 can promote the metastasis, invasion, and proliferation of oral squamous cell carcinoma through the AKT/ERK1/2 signaling pathway (35). In colorectal

cancer, BST2 induces macrophage M2 polarization and promotes tumor progression (36). And C1QB can be used as a potential prognostic biomarker for multiple tumors (37). In cervical cancer, the expression of C1QB protein is related to P16 expression (38). However, their specific relationship with endometrial cancer still needs to be further explored in depth by further studies.

We input the DEGs into the NetworkAnalyst platform to determine the TF gene interactions and TF - miRNA co-regulatory networks. TF genes play a regulatory role according to gene expression, which may lead to the production of cancer cells. From the network, ICAM1 exhibits high interaction rates with other TF-genes, with a degree value of 69 in the TF-genes interaction network. Visualizing the core regulatory network in the TF-miRNA core regulation network helps identify key miRNAs influencing the prognosis of endometriosis. The study identified 70 TF-genes and 37 miRNAs. Among the TFs with the most interactions, ICAM1 has a high degree value of 31, associating with 8 miRNAs. ICAM1 and TIMP1, as regulators of cell adhesion and migration (39), were also highlighted in our analysis. These genes may promote the infiltration and metastasis of tumor cells in the tumor microenvironment by affecting intercellular adhesion and matrix degradation (40–42), similarly in recent years studies have shown that TIMP1 is highly expressed in endometrial cancer and endometriosis tissues (43, 44).

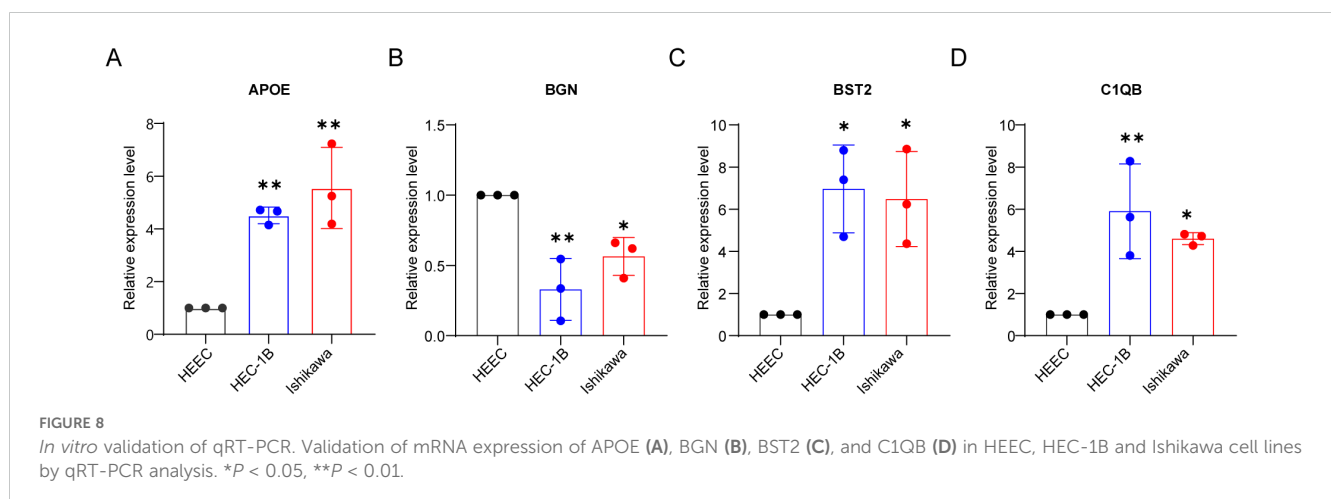
TABLE 5 DEGs-drug interactions in the DGldb database.

Gene	Drug	Score	Types	Sources	PMIDs
APOE	TROGLITAZONE	0.29	Definite	NCI	15057551
	SOYBEAN OIL	1.45	Definite	NCI	3021887
	ACENOCOUMAROL	1.82	Definite	PharmGKB	29432897
	PREDNISON	0.23	Definite	NCI	3185288
	PRAVASTATIN	0.31	Definite	PharmGKB	19667110 17289397
	STAUROSPORINE	0.27	Definite	NCI	11948667
	SIMVASTATIN	0.52	Definite	PharmGKB	10736278
	TRIAMCINOLONE	0.32	Definite	NCI	9150415
	ATORVASTATIN	0.30	Definite	PharmGKB	16103896 19667110 20031582 17289397
	LUTEIN	3.64	Definite	NCI	11413081
	LORAZEPAM	0.33	Definite	NCI	15699298
	FLUVASTATIN	0.52	Definite	PharmGKB	17289397 30363031
	GANCICLOVIR	1.82	Definite	NCI	16322528
	IRBESARTAN	0.73	Definite	NCI	12827021
	FENOFIBRATE	0.25	Definite	PharmGKB	12042669
	RITONAVIR	0.57	Definite	PharmGKB	17700364 15809899
WARFARIN	0.30	Definite	PharmGKB	21923605 29432897	
ICAM1	ENLIMOMAB	12.37	Definite	ChemblInteractions	-
	BI-505	18.55	Definite	TdgClinicalTrial	-
	ALICAFORSEN	27.73	Definite	ChemblInteractions TTD	-
	HYALURONAN	4.12	Definite	TdgClinicalTrial TTD	-
	LIFITEGRAST	8.24	Definite	TdgClinicalTrial DTC TTD	-
MME	CANOXATRIL	34.78	Definite	TdgClinicalTrial TEND TTD	-
	SACUBITRIL	7.73	Definite	ChemblInteractions	-
	GALLOPAMIL	2.58	Definite	TTD	-
	SAMPATRILAT	7.73	Definite	TTD	-
	LCZ696	7.73	Definite	TTD	-
	PEPINEMAB	3.86	Definite	TTD	-
	ILEPATRIL	3.86	Definite	TdgClinicalTrial	-
	SLV-334	3.86	Definite	TTD	-

Research indicates that ICAM1 can serve as a biomarker for endometriosis (45). Comprehensive analysis of the PPI network, TF-miRNA core regulatory network provides valuable information for understanding the regulatory mechanisms and potential therapeutic targets of endometriosis and its progression.

Tumor-infiltrating immune cells play a crucial role in tumor progression, affecting anticancer treatments and eventually leading to the development of immune tolerance (46). Endometrial cancer is considered an immunogenic disease, and increasing evidence suggests the involvement of the immune system in its progression

and patient prognosis (47). Similarly, patients with endometriosis exhibit abnormal distribution of immune cells, and immune system dysregulation holds clinical significance in the pathogenesis (48). Our results showed that the expressions of APOE, BGN, BST1, and BGN were positively correlated with the infiltration status of immune cells such as B cells, CD8 + T cells, and neutrophils in cancer, demonstrating the significance of APOE, BGN, BST1, and BGN in recruiting and regulating infiltrating immune cells in cancer. In our qRT - PCR verification, the expressions of APOE, BGN, BST1, and C1QB in endometrial cancer cell lines were different from



those in normal endometrial epithelial cells, which was consistent with our analysis. However, through literature data, we found that BGN may promote the disease progression of most cancers. therefore, there may be some other regulatory relationships between BGN and endometriosis, and we cannot say that it has a certain relationship with the prognosis of endometriosis. Three of the 10 DEGs were found to be associated with drugs in the DGIdb drug database (Table 5). The targets of these potentially applicable drugs await further investigation through future cellular and animal studies.

In summary, by integrating data from four databases and employing various bioinformatics methods to explore important gene modules, we identified common genes between endometriosis and endometrial cancer, revealing relative similarities in immune dysfunction between these two diseases. The four hub genes (APOE, BST1, BGN, C1QB) with differential expression in endometrial cancer tissues may serve as potential therapeutic targets, warranting further research to validate their effectiveness as potential prognostic markers and treatment targets for endometrial cancer.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

RYM: Data curation, Funding acquisition, Writing – original draft. YZ: Investigation, Writing – original draft. JNW: Data

curation, Funding acquisition, Writing – original draft. HX: Writing – review & editing. RRZ: Writing – review & editing. ZJX: Writing – review & editing. LZ: Formal analysis, Writing – review & editing. RHZ: Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was funded by the hospital-level project of the Ninth People's Hospital of Suzhou (Grant No. YK202202), the science and education program for promoting health in Wujiang District, Suzhou City (Grant No.WWK202201), and the Youth Science and Technology program of Suzhou (Grant No. KJXW2021078).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Zondervan KT, Becker CM, Koga K, Missmer SA, Taylor RN, Vignani P. Endometriosis. *Nat Rev Dis Primers*. (2018) 4:9. doi: 10.1038/s41572-018-0008-5
- Crosbie EJ, Kitson SJ, McAlpine JN, Mukhopadhyay A, Powell ME, Singh N. Endometrial cancer. *Lancet*. (2022) 399:1412–28. doi: 10.1016/S0140-6736(22)00323-3
- Oaknin A, Bosse TJ, Creutzberg CL, Giornelli G, Harter P, Joly F, et al. Endometrial cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*. (2022) 33:860–77. doi: 10.1016/j.annonc.2022.05.009
- Hermens M, Van Altena AM, Velthuis I, van de Laar DCM, Bulten J, van Vliet H, et al. Endometrial cancer incidence in endometriosis and adenomyosis. *Cancers (Basel)*. (2021) 13. doi: 10.3390/cancers13184592
- Terzic M, Aimagambetova G, Kunz J, Bapayeva G, Aitbayeva B, Terzic S, et al. Molecular basis of endometriosis and endometrial cancer: current knowledge and future perspectives. *Int J Mol Sci*. (2021) 22. doi: 10.3390/ijms22179274
- Taylor HS, Kotlyar AM, Flores VA. Endometriosis is a chronic systemic disease: clinical challenges and novel innovations. *Lancet*. (2021) 397:839–52. doi: 10.1016/S0140-6736(21)00389-5
- Othman Eel D, Hornung D, Salem HT, Khalifa EA, El-Metwally TH, Al-Hendy A. Serum cytokines as biomarkers for nonsurgical prediction of endometriosis. *Eur J Obstet Gynecol Reprod Biol*. (2008) 137:240–6. doi: 10.1016/j.ejogrb.2007.05.001
- Akoum A, Jolicoeur C, Boucher A. Estradiol amplifies interleukin-1-induced monocyte chemotactic protein-1 expression by ectopic endometrial cells of women with endometriosis. *J Clin Endocrinol Metab*. (2000) 85:896–904. doi: 10.1210/jc.85.2.896
- Zhu X, Xu Y, Wang J, Xue Z, Qiu T, Chen J. Loss of NLRP3 reduces oxidative stress and polarizes intratumor macrophages to attenuate immune attack on endometrial cancer. *Front Immunol*. (2023) 14:1165602. doi: 10.3389/fimmu.2023.1165602
- Pena CG, Nakada Y, Saatcioglu HD, Aloisio GM, Cuevas I, Zhang S, et al. LKB1 loss promotes endometrial cancer progression via CCL2-dependent macrophage recruitment. *J Clin Invest*. (2015) 125:4063–76. doi: 10.1172/JCI82152
- Lu W, He F, Lin Z, Liu S, Tang L, Huang Y, et al. Dysbiosis of the endometrial microbiota and its association with inflammatory cytokines in endometrial cancer. *Int J Cancer*. (2021) 148:1708–16. doi: 10.1002/ijc.v148.7
- Abramiuk M, Grywalska E, Malkowska P, Sierawska O, Hryniewicz R, Niedzwiedzka-Rystwej P. The role of the immune system in the development of endometriosis. *Cells*. (2022) 11. doi: 10.3390/cells11132028
- Yu JJ, Sun HT, Zhang ZF, Shi RX, Liu LB, Shang WQ, et al. IL15 promotes growth and invasion of endometrial stromal cells and inhibits killing activity of NK cells in endometriosis. *Reproduction*. (2016) 152:151–60. doi: 10.1530/REP-16-0089
- Gruber TM, Mechsner S. Pathogenesis of endometriosis: the origin of pain and subfertility. *Cells*. (2021) 10. doi: 10.3390/cells10061381
- Kim H, Kim HJ, Ahn HS. Does endometriosis increase the risks of endometrial hyperplasia and endometrial cancer? *Gynecol Oncol*. (2023) 169:147–53. doi: 10.1016/j.ygyno.2022.06.021
- Adilbayeva A, Kunz J. Pathogenesis of endometriosis and endometriosis-associated cancers. *Int J Mol Sci*. (2024) 25. doi: 10.3390/ijms25147624
- Farolfi A, Altavilla A, Morandi L, Capelli L, Chiadini E, Prisinzano G, et al. Endometrioid cancer associated with endometriosis: from the seed and soil theory to clinical practice. *Front Oncol*. (2022) 12:859510. doi: 10.3389/fonc.2022.859510
- Chen JJ, Xiao ZJ, Meng X, Wang Y, Yu MK, Huang WQ, et al. MRP4 sustains Wnt/beta-catenin signaling for pregnancy, endometriosis and endometrial cancer. *Theranostics*. (2019) 9:5049–64. doi: 10.7150/tno.32097

19. Stylianou A, Gkretsi V, Louca M, Zacharia LC, Stylianopoulos T. Collagen content and extracellular matrix cause cytoskeletal remodelling in pancreatic fibroblasts. *J R Soc Interface*. (2019) 16:20190226. doi: 10.1098/rsif.2019.0226
20. Chen CG, Iozzo RV. Extracellular matrix guidance of autophagy: a mechanism regulating cancer growth. *Open Biol*. (2022) 12:210304. doi: 10.1098/rsob.210304
21. Kusama K, Satoyoshi A, Azumi M, Yoshie M, Kojima J, Mizuno Y, et al. Toll-like receptor signaling pathway triggered by inhibition of serpin A1 stimulates production of inflammatory cytokines by endometrial stromal cells. *Front Endocrinol (Lausanne)*. (2022) 13:966455. doi: 10.3389/fendo.2022.966455
22. Wieboldt R, Laubli H. Glycosaminoglycans in cancer therapy. *Am J Physiol Cell Physiol*. (2022) 322:C1187–C200. doi: 10.1152/ajpcell.00063.2022
23. Zucchetto A, Serraino D, Polesel J, Negri E, De Paoli A, Dal Maso L, et al. Hormone-related factors and gynecological conditions in relation to endometrial cancer risk. *Eur J Cancer Prev*. (2009) 18:316–21. doi: 10.1097/CEJ.0b013e328329d830
24. Huang JY, Ma KS, Wang LT, Chiang CH, Yang SF, Wang CH, et al. The risk of endometrial cancer and uterine sarcoma following endometriosis or pelvic inflammatory disease. *Cancers (Basel)*. (2023) 15. doi: 10.3390/cancers15030833
25. Villarino AV, Kanno Y, O'shea JJ. Mechanisms and consequences of Jak-STAT signaling in the immune system. *Nat Immunol*. (2017) 18:374–84. doi: 10.1038/ni.3691
26. Shen L, Zhang C, Cui K, Liang X, Zhu G. Low-density lipoprotein contributes to endometrial carcinoma cell proliferation, migration, and invasion by activating the JAK-STAT signaling pathway. *Anal Cell Pathol (Amst)*. (2023) 2023:4015167. doi: 10.1155/2023/4015167
27. Kotlyar AM, Mamillapalli R, Flores VA, Taylor HS. Tofacitinib alters STAT3 signaling and leads to endometriosis lesion regression. *Mol Hum Reprod*. (2021) 27. doi: 10.1093/molehr/gaab016
28. Xu X, Yang H, Chen JN, Hua LJ, Wang RY, Liu T, et al. Moxibustion attenuates inflammation and alleviates axial spondyloarthritis in mice: Possible role of APOE in the inhibition of the Wnt pathway. *J Tradit Complement Med*. (2022) 12:518–28. doi: 10.1016/j.jtcme.2022.04.002
29. Kacperczyk M, Kmiecik A, Kratz EM. The role of apoE expression and variability of its glycosylation in human reproductive health in the light of current information. *Int J Mol Sci*. (2021) 22. doi: 10.3390/ijms22137197
30. Wu C, Li T, Cheng W. The correlation between APOE expression and the clinical characteristics and prognosis of patients with endometrial cancer. *Med (Baltimore)*. (2022) 101:e30536. doi: 10.1097/MD.00000000000030536
31. Lai H, Zhao X, Qin Y, Ding Y, Chen R, Li G, et al. FAK-ERK activation in cell/matrix adhesion induced by the loss of apolipoprotein E stimulates the Malignant progression of ovarian cancer. *J Exp Clin Cancer Res*. (2018) 37:32. doi: 10.1186/s13046-018-0696-4
32. Sun H, Wang X, Zhang Y, Che X, Liu Z, Zhang L, et al. Biglycan enhances the ability of migration and invasion in endometrial cancer. *Arch Gynecol Obstet*. (2016) 293:429–38. doi: 10.1007/s00404-015-3844-5
33. Jacobsen F, Kraft J, Schroeder C, Hube-Magg C, Kluth M, Lang DS, et al. Up-regulation of biglycan is associated with poor prognosis and PTEN deletion in patients with prostate cancer. *Neoplasia*. (2017) 19:707–15. doi: 10.1016/j.neo.2017.06.003
34. Subbarayan K, Massa C, Leisz S, Steven A, Bethmann D, Biehl K, et al. Biglycan as a potential regulator of tumorigenicity and immunogenicity in K-RAS-transformed cells. *Oncoimmunology*. (2022) 11:2069214. doi: 10.1080/2162402X.2022.2069214
35. Shan F, Shen S, Wang X, Chen G. BST2 regulated by the transcription factor STAT1 can promote metastasis, invasion and proliferation of oral squamous cell carcinoma via the AKT/ERK1/2 signaling pathway. *Int J Oncol*. (2023) 62. doi: 10.3892/ijo.2023.5502
36. He X, Chen H, Zhong X, Wang Y, Hu Z, Huang H, et al. BST2 induced macrophage M2 polarization to promote the progression of colorectal cancer. *Int J Biol Sci*. (2023) 19:331–45. doi: 10.7150/ijbs.72538
37. Jiang J, Ding Y, Wu M, Lyu X, Wang H, Chen Y, et al. Identification of TYROBP and C1QB as two novel key genes with prognostic value in gastric cancer by network analysis. *Front Oncol*. (2020) 10:1765. doi: 10.3389/fonc.2020.01765
38. Lin ST, Liang ZQ, Chen XY, Ye XQ, Pang YY, Luo JY, et al. Detection of complement C1q B chain overexpression and its latent molecular mechanisms in cervical cancer tissues using multiple methods. *Int J Genomics*. (2022) 2022:8775330. doi: 10.1155/2022/8775330
39. Bui TM, Wiesolek HL, Sumagin R. ICAM-1: A master regulator of cellular responses in inflammation, injury resolution, and tumorigenesis. *J Leukoc Biol*. (2020) 108:787–99. doi: 10.1002/JLB.2MR0220-549R
40. Taftaf R, Liu X, Singh S, Jia Y, Dashzeveg NK, Hoffmann AD, et al. ICAM1 initiates CTC cluster formation and trans-endothelial migration in lung metastasis of breast cancer. *Nat Commun*. (2021) 12:4867. doi: 10.1038/s41467-021-25189-z
41. Zhou Q, Xu J, Xu Y, Sun S, Chen J. Role of ICAM1 in tumor immunity and prognosis of triple-negative breast cancer. *Front Immunol*. (2023) 14:1176647. doi: 10.3389/fimmu.2023.1176647
42. Tian Z, Ou G, Su M, Li R, Pan L, Lin X, et al. TIMP1 derived from pancreatic cancer cells stimulates Schwann cells and promotes the occurrence of perineural invasion. *Cancer Lett*. (2022) 546:215863. doi: 10.1016/j.canlet.2022.215863
43. Luddi A, Marrocco C, Governini L, Semplici B, Pavone V, Luisi S, et al. Expression of matrix metalloproteinases and their inhibitors in endometrium: high levels in endometriotic lesions. *Int J Mol Sci*. (2020) 21. doi: 10.3390/ijms21082840
44. Shaco-Levy R, Sharabi S, Piura B, Sion-Vardy N. MMP-2, TIMP-1, E-cadherin, and beta-catenin expression in endometrial serous carcinoma compared with low-grade endometrial endometrioid carcinoma and proliferative endometrium. *Acta Obstet Gynecol Scand*. (2008) 87:868–74. doi: 10.1080/00016340802253775
45. Mosbah A, Nabel Y, Khashaba E. Interleukin-6, intracellular adhesion molecule-1, and glycodelin A levels in serum and peritoneal fluid as biomarkers for endometriosis. *Int J Gynaecol Obstet*. (2016) 134:247–51. doi: 10.1016/j.ijgo.2016.01.018
46. Suszczyk D, Skiba W, Pawlowska A, Polak G, Tarkowski R, Wertel I. Expression of gal-9 on dendritic cells and soluble forms of TIM-3/gal-9 in patients suffering from endometriosis. *Int J Mol Sci*. (2023) 24. doi: 10.3390/ijms24065948
47. Suszczyk D, Skiba W, Jakubowicz-Gil J, Kotarski J, Wertel I. The role of myeloid-derived suppressor cells (MDSCs) in the development and/or progression of endometriosis-state of the art. *Cells*. (2021) 10. doi: 10.3390/cells10030677
48. Vallve-Juanico J, Houshdaran S, Giudice LC. The endometrial immune environment of women with endometriosis. *Hum Reprod Update*. (2019) 25:564–91. doi: 10.1093/humupd/dmz018