Bioinformatical analysis identifies PDLIM3 as a Potential biomarker associated with Immune Infiltration in Patients with Endometriosis

Lei Gan [Equal first author, 1] , **Jiani Sun** [Equal first author, 1] , **Jing Sun** [Corresp. 1]

[1] Shanghai First Maternity and Infant Hospital, Tongji university, Shanghai, Shanghai, 中国

Corresponding Author: Jing Sun
Email address: sunjing61867@163.com

Background. Endometriosis is a chronic systemic disease, whose classic symptoms are pelvic pain and infertility. This disease seriously reduces the life quality of patients. The pathogenesis, recognition and treatment of endometriosis is still unclear, and cannot be over emphasized. The aim of our study was to investigate the potential biomarker of endometriosis for the mechanism and treatment.

Methods. Using GSE11691 , GSE23339 and GSE5108 datasets, differentially expressed genes (DEGs) were identified between endometriosis and normal samples. The functions of DEGs were reflected by the analysis of gene ontology (GO), pathway enrichment and gene set enrichment analysis (GSEA). The LASSO regression model was performed to identify candidate biomarkers. The receiver operating characteristic curve (ROC) was used to evaluate discriminatory ability of candidate biomarkers. The predictive value of the markers in endometriosis were further validated in the GSE 120103 dataset. Then, the expression level of biomarkers was detected by qRT-PCR and Western blot. Finally, the relationship between candidate biomarker expression and immune infiltration was estimated using CIBERSORT.

Results. A total of 42 genes were identified, which were mainly involved in cytokine–cytokine receptor interaction, systemic lupus erythematosus and chemokine signaling pathway. We confirmed PDLIM3 was a specific biomarker in endometriosis (AUC = 0.955) and validated in the GSE120103 dataset (AUC = 0.836). And the mRNA and protein expression level of PDLIM3 in endometriosis tissue was significantly higher than normal. Immune cell infiltration analysis revealed that PDLIM3 was correlated with M2 macrophages, neutrophils, CD4+ memory resting T cells, gamma delta T cells, M1 Macrophages, resting mast cells, follicular helper T cells, activated NK cells, CD8+ T cells, regulatory T cells (Tregs), naive B cells, plasma cells and resting NK cells.

1 **Bioinformatical analysis identifies PDLIM3 as a**
2 **Potential gene biomarker associated with Immune**
3 **Infiltration in Patients with Endometriosis**

5 Lei Gan[†], Jia-ni Sun[†], Jing Sun
6
7 Department of Gynecology, Shanghai First Maternity and Infant Hospital, School of Medicine,
8 Tongji University, Shanghai, China
9
10 [†]These authors have contributed equally to this work and share first authorship
11
12 Corresponding Author:
13 Jing Sun
14 Shanghai First Maternity and Infant Hospital, School of Medicine, Tongji University, Shanghai
15 200092, China
16 Email address: sunjing61867@tongji.edu.cn

18 Abstract

19 **Background.** Endometriosis is a chronic systemic disease, whose classic symptoms are pelvic
20 pain and infertility. This disease seriously reduces the life quality of patients. The pathogenesis,
21 recognition and treatment of endometriosis is still unclear, and cannot be over emphasized. The
22 aim of our study was to investigate the potential biomarker of endometriosis for the mechanism
23 and treatment.
24 **Methods.** Using GSE11691, GSE23339 and GSE5108 datasets, differentially expressed genes
25 (DEGs) were identified between endometriosis and normal samples. The functions of DEGs
26 were reflected by the analysis of gene ontology (GO), pathway enrichment and gene set
27 enrichment analysis (GSEA). The LASSO regression model was performed to identify candidate
28 biomarkers. The receiver operating characteristic curve (ROC) was used to evaluate
29 discriminatory ability of candidate biomarkers. The predictive value of the markers in
30 endometriosis were further validated in the GSE120103 dataset. Then, the expression level of
31 biomarkers was detected by qRT-PCR and Western blot. Finally, the relationship between
32 candidate biomarker expression and immune infiltration was estimated using CIBERSORT.
33 **Results.** A total of 42 genes were identified, which were mainly involved in cytokine–cytokine
34 receptor interaction, systemic lupus erythematosus and chemokine signaling pathway. We
35 confirmed PDLIM3 was a specific biomarker in endometriosis (AUC = 0.955) and validated in
36 the GSE120103 dataset (AUC = 0.836). And the mRNA and protein expression level of
37 PDLIM3 in endometriosis tissue was significantly higher than normal. Immune cell infiltration
38 analysis revealed that PDLIM3 was correlated with M2 macrophages, neutrophils, CD4+
39 memory resting T cells, gamma delta T cells, M1 Macrophages, resting mast cells, follicular

40 helper T cells, activated NK cells, CD8+ T cells, regulatory T cells (Tregs), naive B cells, plasma
41 cells and resting NK cells.
42
43

Introduction

45 Recently, endometriosis is defined as a chronic systemic disease, characterized by the presence
46 and infiltration of ectopic endometrial tissue (1). It affects nearly 10% of women who are of
47 reproductive age. The most common symptoms are pelvic pain and infertility. Nearly 70% of
48 teens with pelvic pain are later diagnosed with endometriosis (2). Endometriosis seriously
49 reduces the life quality of patients, increases incidence of depression and causes a heavy burden
50 on the direct and indirect healthcare (3). The gold standard for endometriosis diagnosis is
51 invasive laparoscopy examination. It is difficult to recognize because of patients' nonspecific
52 descriptions of symptoms, especially at early stages. Despite its prevalence, the average time
53 from symptom onset to diagnosis is five to ten years (4, 5). The new opinion about diagnosis of
54 endometriosis is emphasize disease symptoms and their origins rather than lesion presence or
55 absence (2). Specific circulating biomarkers, such as miR-125b-5p, miR-150-5p, miR-61, have
56 been studied as promising early clinical diagnostic tools, although functional researches are
57 needed to confirm their effect on this disease. Currently, medical therapy and surgical resection
58 are the gold standard treatment for endometriosis (6) while both are associated with high
59 recurrence rates and limited by high occurrence of side effects. The mechanisms underlying the
60 pathogenesis and development of endometriosis will open the door to a new approach to
61 treatment.
62
63 Various hypotheses have been proposed to explain the underlying molecular mechanisms of
64 endometriosis, consisting of the endometrial implantation theory of Sampson (7), eutopic
65 endometrium determinism, stem cell factors (8), epigenetic factors, immunodeficiency as well as
66 immune and inflammatory factors (9).
67
68 Recent clinical and molecular studies have suggested that innate properties or acquired properties
69 of the endometrium and impaired immune homeostasis are systems of significance in explaining
70 the mechanism of endometriosis (8). Immune imbalance is associated with increased
71 implantation, proliferation, and angiogenesis of ectopic endometrium. Previous studies have
72 shown that immune dysfunction happens in peritoneal fluid of endometriosis, which may
73 contribute to the progression of endometriosis. Functional changes in immune components have
74 been indicated, consisting of natural killer (NK) cells, monocytes/macrophages, B cells, T
75 lymphocytes, and cytokines (10). Although numerous studies have described occurrence of
76 immune abnormalities in endometriosis, the role of the immune system is not clearly understood.
77 Hence, it is essential to explore mechanism of immune dysfunction in endometriosis, which can
78 contribute to elucidate the role of immune system in pathogenesis of endometriosis and generate

79 applicable insights to develop preventive and control strategies, original non-invasive diagnostic
80 methods and targeted therapeutics.
81
82 In the present study, we downloaded three microarray datasets (GSE11691, GSE23339,
83 GSE5108) from the Gene Expression Omnibus (GEO) database to detect the potential
84 biomarkers for further analysis. We identified candidate genes, explore immune mechanism and
85 key signaling pathways regulating occurrence, development and progression of endometriosis
86 through bioinformatics method. In conclusion, these findings provide novel insights into
87 understanding the underlying immune mechanisms and seeking for potential molecular
88 treatment.
89
90

Materials & Methods
91
92 **Gene expression profile**
93 We searched the associated gene expression profiles of endometriosis in NCBI-GEO
94 (https://www.ncbi.nlm.nih.gov/geo/) (11), using the keywords "endometriosis" and "Homo
95 sapiens". Three gene expression datasets containing microarray data from normal endometrial
96 tissues and ectopic endometrial tissues, GSE11691 (12), GSE23339 (13) and GSE5108 (14),
97 were selected for further analysis. The GSE11691 microarray data consisted of 9 normal human
98 endometrium and 9 ectopic human endometrium (2 from the peritoneal wall, 7 from the broad
99 ligament) and was generated using the GPL96 platform Affymetrix Human Genome U133A
100 Array. The GSE23339 microarray data consisted of 9 normal human endometrium and 10
101 ectopic human endometrium (all from the endometriotic cysts of the ovaries) and was generated
102 using the GPL6102 Illumina human-6 v2.0 expression beadchip. The GSE5108 microarray data
103 consisted of 11 normal human endometrium and 11 ectopic human endometrium (5 from the
104 peritoneal wall, 6 from the endometriotic cysts of the ovaries) and was generated using the
105 GPL2895 GE Healthcare/Amersham Biosciences CodeLink Human Whole Genome Bioarray. In
106 addition, the GSE120103 (15) dataset, containing 18 endometriosis (all from the endometriotic
107 cysts of the ovaries) and 18 control samples, was used as the validation cohort using the
108 GPL6480 platform Agilent-014850 Whole Human Genome Microarray 4x44K G4112F.
109 GSE12768 (16) dataset, containing 2 endometriosis (both from the endometriotic cysts of the
110 ovaries) and 2 control samples, was used as another validation cohort using the GPL7304 Institut
111 Cochin HG18 60mer expression array 47K.
112
113 **Identification of differentially expressed genes (DEGs)**
114 The raw data of three datasets, GSE11691, GSE23339 and GSE5108, was merged and
115 removed batch effects by "sva" package (http://www.bioconductor.org/) (17). Then, the "limma"
116 package (http://www.bioconductor.org/) was used to attain DEGs between endometriosis
117 samples and normal samples. The raw data of GSE12768 was analyzed separately to acquired
118 the DEGs list, and compared with the list of three datasets in order to verify the reliability of our

119   analysis. The significant DEGs were identified according to adjusted P value < 0.05 and ｜log

120   fold change (FC)｜>2.

121

122   **Functional and pathway enrichment analysis of DEGs**

123   To reflect gene functions, gene ontology (GO) (http://geneontology.org/) (18) biological

124   processes were shown in three aspects: biological processes (BP), cellular component (CC) and

125   molecular function (MF). And the analysis process was achieved by "clusterProfiler" package

126   (19, 20) in the present study, using the cut-off criterion P value < 0.05. ToppGene

127   (https://toppgene.cchmc.org/enrichment.jsp) (21) is a free tool for uncovering the biological

128   significance behind a cluster of genes. The REACTOME (https://reactome.org/) (22) pathway

129   enrichment analysis, was exactly performed using this tool. A P value < 0.05 was considered as

130   the threshold points.

131

132   Moreover, to get more information of mechanism behind DEGs, gene set enrichment analysis

133   (GSEA) was employed to explore the potential feature between the endometriosis and control

134   groups. The "c2.cp.kegg.v7.4.symbols.gmt" from the Molecular Signatures Database (MSigDB)

135   was downloaded and used as the reference gene set. An adjusted P < 0.05 and false discovery

136   rate (FDR) < 0.05 were considered as significant.

137

138   **Protein-protein interaction (PPI) networks construction**

139   STRING (https://string-db.org/) (23) is a tool designed to achieve a comprehensive and objective

140   global network and visualize PPI information. Proteins associated with DEGs were chosen based

141   on information in STRING (confidence score > 0.4), and then PPI networks were constructed

142   using Cytoscape software (http://cytoscape.org/) (24).

143

144   **Screening of specific gene biomarkers**

145   The least absolute shrinkage and selection operator (LASSO), a regression analysis algorithm,

146   was applied to evaluate the value of DEGs in this study. The "glmnet" package in R did the

147   above analysis and returned candidate potential genes. And the expression levels of candidate

148   biomarkers were further tested in the GSE120103 dataset.

149

150   **Receiver operating characteristic (ROC) curve analysis**

151   To test the predictive accuracy of the identified biomarkers, we generated ROC curves using

152   "pROC" package (25) based on the candidate genes and compare the predictive value of

153   candidate genes by the area under the ROC curve (AUC). And then, the predictive value of these

154   genes was also validated in the GSE120103 dataset.

155

156   **Relationship between identified genes and infiltrating immune cells in endometriosis**

157   We used the CIBERSORT (https://cibersortx.stanford.edu/) (25) to calculate immune cell

158   infiltrations of endometriosis and control tissues. Then, we applied "corrplot" and "vioplot" R

159   package to visualize the results and explore correlation of infiltrating immune cells. In addition,
160   the relation of the candidate gene markers with the levels of infiltrating immune cells was
161   calculated and visualized by "ggpubr" package.
162

Quantitative real-time PCR (qRT-PCR)

164   Total RNA was isolated from ovarian endometriosis and eutopic endometrium (NC). These
165   operations were approved by Ethics committee of Shanghai First Maternity and Infant Hospital
166   (KS21198). We used Trizol (RNAiso Plus; Takara, Japan) to isolate RNA from these two kinds
167   of tissues according to the manufacturer's instructions. Then, one microgram of RNA was
168   reversely transcribed into cDNA using a PrimeScript™ RT reagent kit (Takara). Amplification
169   was performed using gene-specific primers (Sangon, Shanghai, China) and TB Green® Premix
170   EX TaqTM II (Takara) on a qRT-PCR device (QuantStudio5, Thermo Fisher Scientific,
171   Waltham, MA, USA). β-actin was used as an internal control. The relative expression of the
172   genes was calculated using the 2-ΔΔCT method. The primers used were as follows: human β-
173   actin-Forward: 5′- GTGCTATGTTGCTCTAGACTTCG-3′; human β-actin-Reverse: 5′-
174   ATGCCACAGGATTCCATACC-3′; human PDLIM3-Forward: 5′-
175   GAGTCGGACGTGTACCGGA-3′; human PDLIM3-Reverse: 5′-
176   TGCCACTCCCACATTTGTCAC-3′.
177

Protein Extraction and Western Blot

179   Normal endometrium tissues and ovarian endometriosis tissues were lysed in RIPA lysis buffer
180   (PC101, EpiZyme, Shanghai) with protease and phosphatase inhibitor cocktail. The protein
181   concentration was detected by a BCA protein assay. The same amounts of total protein (32 μg)
182   were loaded and separated onto 12.5% SDS-PAGE and was then transferred to PVDF
183   membranes. After blocking with protein-free rapid blocking buffer (PS108P, EpiZyme,
184   Shanghai) for 0.5 h at room temperature, the membranes were incubated with primary antibodies
185   at 4 °C overnight. The primary antibodies rabbit anti-PDLIM3 antibody (1:2000; Servicebio)
186   was used and rabbit anti-β-Actin (1:5000, ABclonal) served as the internal control.
187   The membranes were probed with secondary antibodies (1:3000, Servicebio) for 1.5 h at room
188   temperature and three washes in TBST were processed to remove excess secondary antibody.
189   We then used an ECL western blotting kit (NCM Biotech, China) to visualize and image the
190   targeted protein bands.
191

Statistical Analysis

193   Most statistical calculations were conducted using R (version 4.1.1). The results of qRT-PCR
194   and western blot were shown as the mean ± SD and GraphPad Prism 8.0 (GraphPad Software,
195   USA) were used for statistical analysis. All the comparisons between two groups were analyzed
196   using Student's t-test for normally distributed variables. If the variables distribution was
197   abnormal, the Mann–Whitney U-test would be used. The relationship between the expression of
198   biomarkers and infiltrating immune cells was analyzed using Spearman's rank correlation

199 analysis. Two-sided with P < 0.05 (95% confidence interval) and false discovery rate (FDR) <
200 0.05 were considered as statistically significant.
201

Results

203 **Identification of DEGs in endometriosis**
204 Data from a total of 30 endometriosis samples (14 peritoneal endometriosis samples, 16 ovarian
205 endometriosis samples) and 29 control endometrium samples from three GEO datasets
206 (GSE11691, GSE23339 and GSE5108) were analyzed. The DEGs were calculated using the
207 "limma" package after removing the batch effects. We obtained a total of 42 DEGs, including
208 upregulated 23 genes and downregulated 19 genes (Table 1). All DEGs were shown in the
209 heatmap and volcano map (Figure 1). In addition, the DEGs list from GSE12768 dataset was
210 displayed in Supplementary Table 1.
211

212 **GO and pathway enrichment of DEGs in endometriosis**
213 The above list of DEGs was used to evaluate if they were enriched into specific GO or reactome
214 pathways. GO analysis identified that the DEGs were significantly enriched in BP, including
215 humoral immune response, complement activation, alternative pathway, antimicrobial humoral
216 response, and so on (Table 2). In terms of MF, DEGs were mainly enriched in Toll-like receptor
217 binding, G protein-coupled receptor binding, long-chain fatty acid binding, and chemokine
218 receptor binding (Table 2). Moreover, CC demonstrated that filamentous actin, lipid droplet,
219 sperm flagellum and actin filament were the most significantly enriched GO term (Table 2). The
220 results were displayed in Figure 2A, 2B. In addition, the GSEA results illustrated that the
221 enriched pathways mainly involved cytokine–cytokine receptor interaction, systemic lupus
222 erythematosus, chemokine signaling pathway, complement and coagulation cascades, and
223 leishmania infection (Figure 2C, 2D).
224

225 REACTOME pathway enrichment analysis was also used to obtain the signaling pathways for
226 different genes. In this study, these DEGs were mainly involved in the regulation of complement
227 cascade, suprofen pathway, indomethacin pathway and mefenamic acid pathway (Table 3).
228

229 **PPI networks construction and validation of predictive feature biomarkers**
230 We uploaded a list of proteins associated with DEGs, removed the disconnected nodes and then
231 acquired a PPI network using Cytoscape software (Figure 3A). There were 26 nodes and 50
232 edges on this network.
233

234 Furthermore, the LASSO algorithm was used to narrowed down the DEGs, leading to the
235 identification of seven variables as specific biomarkers for endometriosis (Figure 3B). The seven
236 genes were FMO1, PDLIM3, FAM129A, CLDN11, C3, TMPRSS4 and DEFB1. And then, we
237 used ROC curve to estimate the predictive value of genes. As shown in Figure 3C, the ability of
238 PDLIM3 in discriminating endometriosis from the control samples demonstrated a favorable

239 diagnostic value, with an AUC of 0.955 (95% CI 0.894–0.997). Moreover, this result was
240 confirmed in the GSE120103 dataset with an AUC of 0.836 (95% CI 0.691–0.948). Other six
241 genes showed no significant difference in the GSE120103 dataset (Figure 3D).

243 At the same time, in order to generate more accurate and reliable results, the GSE120103 dataset
244 was employed to verify the expression levels of the seven features. The expression levels of
245 PDLIM3 in endometriosis tissue were notably higher than those in the normal tissue (Figures
246 4A). There was no significant difference between the two groups in other six gene expression
247 (Figure 4A). Therefore, PDLIM3 became one of the potential specific biomarkers in
248 endometriosis.

250 **QRT-PCR and Western Blot**
251 To validate the expression of PDLIM3, qRT-PCR and western blot were used to measure
252 transcriptional and protein expression in both ovarian endometriosis and normal eutopic
253 endometrium tissue from endometriosis-free women. We found PDLIM3 had higher expression
254 level in endometriosis group than normal control (qRT-PCR: n=6 vs. 6; Western blot: n=7 vs. 4)
255 (Figure 4B, 4C). The patients of endometriosis ranged in age from 26 to 43 years (average age
256 32 years). The BMI was 21.76 ± 2.42 (mean $\pm$ SD). And, their rASRM endometriosis stage was
257 III or IV. Normal group samples were from non-endometriosis patients, ranged in age from 31 to
258 41 years (average age 34 years). The BMI was 21.23 ± 1.21. There was no significant different in
259 their general characteristics.

261 **Correlation analysis between PDLIM3 and infiltrating immune cells**
262 The composition of immune cells in endometriosis tissues and normal control tissues were
263 shown in Figure 5A. The proportions of naive B cells (P < 0.001), plasma cells (P = 0.012),
264 CD8+ T cells (P = 0.013), follicular helper T cells (P < 0.001), regulatory T cells (Tregs) (P =
265 0.004), resting NK cells (P = 0.037), and activated NK cells (P < 0.001) in endometriosis tissues
266 were significantly lower than those in control tissues. However, the proportions of CD4+
267 memory resting T cells (P < 0.001), gamma delta T cells (P = 0.031), M1 Macrophages (P =
268 0.020), M2 Macrophages (P < 0.001), and neutrophils (P = 0.002) in endometriosis tissues were
269 significantly higher than those in normal tissues (Figure 5B). However, in GSE120103, only the
270 proportion of Tregs cell (P = 0.037) in endometriosis tissues were significantly lower than that in
271 control tissues (Supplementary Figure 1).

274 The association of 22 types of immune cells was calculated and drawn in Figure 5C. Then, we
275 found PDLIM3 was positively correlated with M2 Macrophages (r = 0.70, P < 0.001),
276 neutrophils (r = 0.45, P < 0.001), CD4+ memory resting T cells (r = 0.44, P < 0.001), gamma
277 delta T cells (r = 0.34, P = 0.0088), resting mast cells (r = 0.31, P = 0.019) and M1 Macrophages
278 (r = 0.28, P = 0.037). However, PDLIM3 was negatively correlated with follicular helper T cells

279 (r = -0.66, P < 0.001), activated NK cells (r = -0.46, P < 0.001), CD8+ T cells (r = -0.41, P =
280 0.0015), Tregs cells (r = -0.39, P = 0.0028), naive B cells (r = -0.37, P = 0.0047), plasma cells (r
281 = -0.31, P = 0.018) and resting NK cells (r = -0.29, P = 0.031) (Figures 6). Besides, PDLIM3
282 was significantly negatively correlated with activated mast cells (r = -0.69, P = 0.038) in
283 GSE120103 (Supplementary Figure 1).

284

285

286

287 **Discussion**

288 The most widly accepted theory about pathogenesis of endometriosis is retrograde menstruation
289 and endometrial implantation reported by Sampson et al. However, it cannot perfectly explain all
290 aspects of endometriosis. More evidence determined a combination of genetic or epigenetic
291 alterations leads to endometriosis (27-30). Therefore, it is critical to acquire a clear
292 understanding of the molecular mechanisms underlying endometriosis if we are to improve the
293 diagnosis and treatment of endometriosis. The main purpose of our study was to detect the
294 potential biomarker of endometriosis for the mechanism and treatment.

295

296 In this investigation, we compared three public datasets and a total of 42 DEGs were identified,
297 including 23 upregulated genes and 19 downregulated genes. Another DEGs list of the validation
298 set included most of 42 DEGs, which strengthen current results. The results of enrichment
299 analyses suggested that gene function enriched by DEGs were mainly associated with immune
300 system, such as humoral immune response, complement activation, alternative pathway,
301 antimicrobial humoral response, cytokine–cytokine receptor interaction, systemic lupus
302 erythematosus, chemokine signaling pathway, and regulation of complement cascade.
303 Endometriosis is a disease with prolonged inflammatory response. The presence of ectopic
304 tissues in peritoneal cavity, without clear mechanism, reaches a state of balance with a high level
305 of proinflammatory cytokines and immune cells (31). This evidence is consistent with our
306 findings, proving that the results in the present study are accurate, as well as confirming that the
307 immune response plays a vital role in endometriosis.

308

309 Based on a machine-learning algorithm, seven variables were identified from 42 DEGs and then
310 verified their diagnosis value by ROC curve. Finally, PDLIM3 was screened out due to it is
311 highly expressed in ectopic endometrium and has a favorable diagnostic value in all datasets.
312 There were totally 14 peritoneal endometriosis samples, 16 ovarian endometriosis samples in
313 three train sets and 18 ovarian endometriosis samples in the fourth verification dataset. Pelvic
314 endometriosis includes ovarian endometriosis (OME), deep infiltrating (DIE) and superficial
315 (SUP) endometriosis. OME is the most common lesion of endometriosis (28), which mostly
316 accounted for our training data and totally accounted for our validating data. Therefore, we
317 believed four datasets used in our study could effectively reflect the clinical proportions of
318 endometriosis of different types. On the other hand, indeed, the origin of samples, which was

319 similar but not identical, between meta-analysis and validation could partly explain the no
320 significant different gene expression for six of the seven genes initially identified.
321
322 PDLIM3, also known as ALP (actin-associated LIM protein), is a member of the PDZ-LIM
323 domain family (32). It is mainly contained a N-terminal PDZ domain and a C-terminal LIM
324 domain. Previous studies revealed PDLIM3 was mainly expressed in skeletal muscle (33) and
325 could regulate skeletal muscle development (34, 35). Lak et al. identified PDLIM3 was one of
326 the rhabdomyosarcoma markers and associated with clinical outcome (36). Wang et al. found
327 that polymorphisms in PDLIM3 gene increased risk of idiopathic dilated cardiomyopathy (37).
328 Other studies indicated PDLIM3 was also correlated to hypertrophic cardiomyopathy (38, 39),
329 atrial fibrillation (40), myofibrillar myopathy (41), the prognosis of hepatocellular carcinoma
330 (42), early regional metastasis of tongue cancer (43), muscle-invasive bladder urothelial
331 carcinoma (44) and medulloblastoma (45). It is the first time that PDLIM3 has been linked to
332 endometriosis. We believe it will play an essential role in better understanding endometriosis
333 mechanism and acting a potential target of therapy.
334
335 Furthermore, we used CIBERSORT to assess the types of immune cell infiltration in
336 endometriosis and normal samples. A variety of immune cell subtypes were potentially related to
337 the occurrence and development of endometriosis. We performed correlation analysis between
338 PDLIM3 and immune cells, PDLIM3 was found to be correlated with M2 macrophages,
339 neutrophils, CD4+ memory resting T cells, gamma delta T cells, M1 Macrophages, resting mast
340 cells, follicular helper T cells, activated NK cells, CD8+ T cells, Tregs cells, naive B cells,
341 plasma cells and resting NK cells. And, we analyzed the immune microenvironment of
342 validation dataset. The tissue compositions and sample size of the meta-analysis and GSE120103
343 were different, which may be able to explain the different results. Notably, these two datasets
344 both showed the specificity of the proportion of Tregs cells. Previous studies indicated the
345 alteration of Tregs cell and Th17 cell may lead to the occurrence and development of ectopic
346 endometrium implantation and accelerate the progression of endometriosis towards advanced
347 stage (46-48).
348
349 Endometriosis is more than a pelvic disease. Changes in multiple systems, such as cardiovascular
350 system, immune system and nervous system, have been found in patients with endometriosis (1).
351 The shifts in circulating immune cell populations as well as the presence of proinflammatory
352 cytokines create a widespread inflammatory environment. A precise control over various types
353 of immune cells is demanded to achieve a safe and effective treatment on endometriosis. The
354 novel biomarkers of endometriosis correlated with the magnitude of immune cell infiltration may
355 be targets of therapy or predictors of therapeutic response.
356
357 The limitations of this study should also be acknowledged. First, the study was retrospective;
358 thus, important clinical information was unavailable. Second, immune cell infiltration and the

359 functions of PDLIM3 in endometriosis were inferred by bioinformatics analysis, and future
360 prospective studies with larger sample sizes should be employed to validate our findings. Third,
361 the biomarker PDLIM3 in this study is derived from endometriosis tissue, which indicates that
362 this biomarker is not suitable to be used for early clinical diagnosis. Further study should be
363 performed to confirm the function of PDLIM3 in endometriosis.
364
365

Conclusions
367 In a word, our results strongly suggest that PDLIM3 plays a previously unidentified but critical
368 role in pathogenesis and immune environment of endometriosis. Better understanding of
369 PDLIM3 could have a highly beneficial impact on the clinical treatment and outcomes of
370 endometriosis patients.
371
372

Acknowledgements
374 All the authors are thankful to Shanghai First Maternity and Infant Hospital and Tongji
375 University for successful completion of this research.

376

References
378 1.   Taylor HS, Kotlyar AM, Flores VA. Endometriosis is a chronic systemic disease: clinical
379 challenges and novel innovations. Lancet. 2021;397(10276):839-52.
380 2.   Yeung P, Jr., Sinervo K, Winer W, Albee RB, Jr. Complete laparoscopic excision of
381 endometriosis in teenagers: is postoperative hormonal suppression necessary? Fertil Steril.
382 2011;95(6):1909-12, 12.e1.
383 3.   Agarwal SK, Chapron C, Giudice LC, Laufer MR, Leyland N, Missmer SA, et al.
384 Clinical diagnosis of endometriosis: a call to action. Am J Obstet Gynecol. 2019;220(4):354.e1-
385 .e12.
386 4.   Hudelist G, Fritzer N, Thomas A, Niehues C, Oppelt P, Haas D, et al. Diagnostic delay
387 for endometriosis in Austria and Germany: causes and possible consequences. Hum Reprod.
388 2012;27(12):3412-6.
389 5.   Greene R, Stratton P, Cleary SD, Ballweg ML, Sinaii N. Diagnostic experience among
390 4,334 women reporting surgically diagnosed endometriosis. Fertil Steril. 2009;91(1):32-9.
391 6.   Vercellini P, Viganò P, Somigliana E, Fedele L. Endometriosis: pathogenesis and
392 treatment. Nat Rev Endocrinol. 2014;10(5):261-75.
393 7.   Kruitwagen RF, Poels LG, Willemsen WN, de Ronde IJ, Jap PH, Rolland R. Endometrial
394 epithelial cells in peritoneal fluid during the early follicular phase. Fertil Steril. 1991;55(2):297-
395 303.
396 8.   Burney RO, Giudice LC. Pathogenesis and pathophysiology of endometriosis. Fertil
397 Steril. 2012;98(3):511-9.

398   9.   Ahn SH, Monsanto SP, Miller C, Singh SS, Thomas R, Tayade C. Pathophysiology and
399   Immune Dysfunction in Endometriosis. Biomed Res Int. 2015;2015:795976.

400   10.   Zou G, Wang J, Xu X, Xu P, Zhu L, Yu Q, et al. Cell subtypes and immune dysfunction
401   in peritoneal fluid of endometriosis revealed by single-cell RNA-sequencing. Cell Biosci.
402   2021;11(1):98.

403   11.   Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, Ledoux P, et al. NCBI GEO:
404   mining millions of expression profiles--database and tools. Nucleic acids research.
405   2005;33(Database issue):D562-6.

406   12.   Hull ML, Escareno CR, Godsland JM, Doig JR, Johnson CM, Phillips SC, et al.
407   Endometrial-peritoneal interactions during endometriotic lesion establishment. The American
408   journal of pathology. 2008;173(3):700-15.

409   13.   Hawkins SM, Creighton CJ, Han DY, Zariff A, Anderson ML, Gunaratne PH, et al.
410   Functional microRNA involved in endometriosis. Molecular endocrinology (Baltimore, Md).
411   2011;25(5):821-32.

412   14.   Eyster KM, Klinkova O, Kennedy V, Hansen KA. Whole genome deoxyribonucleic acid
413   microarray analysis of gene expression in ectopic versus eutopic endometrium. Fertil Steril.
414   2007;88(6):1505-33.

415   15.   Bhat MA, Sharma JB, Roy KK, Sengupta J, Ghosh D. Genomic evidence of Y
416   chromosome microchimerism in the endometrium during endometriosis and in cases of
417   infertility. Reproductive biology and endocrinology : RB&E. 2019;17(1):22.

418   16.   Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential
419   expression analyses for RNA-sequencing and microarray studies. Nucleic acids research.
420   2015;43(7):e47.

421   17.   Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential
422   expression analyses for RNA-sequencing and microarray studies. Nucleic acids research.
423   2015;43(7):e47.

424   18.   Thomas PD. The Gene Ontology and the Meaning of Biological Function. Methods Mol
425   Biol. 2017;1446:15-24.

426   19.   Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal
427   enrichment tool for interpreting omics data. Innovation (New York, NY). 2021;2(3):100141.

428   20.   Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological
429   themes among gene clusters. Omics : a journal of integrative biology. 2012;16(5):284-7.

430   21.   Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment
431   analysis and candidate gene prioritization. Nucleic acids research. 2009;37(Web Server
432   issue):W305-11.

433   22.   Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome
434   pathway knowledgebase. Nucleic acids research. 2020;48(D1):D498-d503.

435   23.   Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al.
436   STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic acids
437   research. 2015;43(Database issue):D447-52.

438 24.   Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a
439 software environment for integrated models of biomolecular interaction networks. Genome
440 research. 2003;13(11):2498-504.
441 25.   Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-
442 source package for R and S+ to analyze and compare ROC curves. BMC bioinformatics.
443 2011;12:77.
444 26.   Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration
445 of cell subsets from tissue expression profiles. Nature methods. 2015;12(5):453-7.
446 27.   Saha R, Pettersson HJ, Svedberg P, Olovsson M, Bergqvist A, Marions L, et al.
447 Heritability of endometriosis. Fertil Steril. 2015;104(4):947-52.
448 28.   Dalsgaard T, Hjordt Hansen MV, Hartwell D, Lidegaard O. Reproductive prognosis in
449 daughters of women with and without endometriosis. Human reproduction (Oxford, England).
450 2013;28(8):2284-8.
451 29.   Bulun SE, Yilmaz BD, Sison C, Miyazaki K, Bernardi L, Liu S, et al. Endometriosis.
452 Endocr Rev. 2019;40(4):1048-79.
453 30.   Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding
454 the missing heritability of complex diseases. Nature. 2009;461(7265):747-53.
455 31.   Li WN, Hsiao KY, Wang CA, Chang N, Hsu PL, Sun CH, et al. Extracellular vesicle-
456 associated VEGF-C promotes lymphangiogenesis and immune cells infiltration in endometriosis.
457 Proceedings of the National Academy of Sciences of the United States of America.
458 2020;117(41):25859-68.
459 32.   Xia H, Winokur ST, Kuo WL, Altherr MR, Bredt DS. Actinin-associated LIM protein:
460 identification of a domain interaction between PDZ and spectrin-like repeat motifs. The Journal
461 of cell biology. 1997;139(2):507-15.
462 33.   Passier R, Richardson JA, Olson EN. Oracle, a novel PDZ-LIM domain protein
463 expressed in heart and skeletal muscle. Mechanisms of development. 2000;92(2):277-84.
464 34.   Pomiès P, Pashmforoush M, Vegezzi C, Chien KR, Auffray C, Beckerle MC. The
465 cytoskeleton-associated PDZ-LIM protein, ALP, acts on serum response factor activity to
466 regulate muscle differentiation. Molecular biology of the cell. 2007;18(5):1723-33.
467 35.   Ohsawa N, Koebis M, Suo S, Nishino I, Ishiura S. Alternative splicing of PDLIM3/ALP,
468 for α-actinin-associated LIM protein 3, is aberrant in persons with myotonic dystrophy.
469 Biochemical and biophysical research communications. 2011;409(1):64-9.
470 36.   Lak NSM, Voormanns TL, Zappeij-Kannegieter L, van Zogchel LMJ, Fiocco M, van
471 Noesel MM, et al. Improving Risk Stratification for Pediatric Patients with Rhabdomyosarcoma
472 by Molecular Detection of Disseminated Disease. Clinical cancer research : an official journal of
473 the American Association for Cancer Research. 2021;27(20):5576-85.
474 37.   Wang D, Fang J, Lv J, Pan Z, Yin X, Cheng H, et al. Novel polymorphisms in PDLIM3
475 and PDLIM5 gene encoding Z-line proteins increase risk of idiopathic dilated cardiomyopathy. J
476 Cell Mol Med. 2019;23(10):7054-62.

477   38.   Lopes LR, Murphy C, Syrris P, Dalageorgou C, McKenna WJ, Elliott PM, et al. Use of
478   high-throughput targeted exome-sequencing to screen for copy number variation in hypertrophic
479   cardiomyopathy. European journal of medical genetics. 2015;58(11):611-6.
480   39.   Bagnall RD, Yeates L, Semsarian C. Analysis of the Z-disc genes PDLIM3 and MYPN
481   in patients with hypertrophic cardiomyopathy. Int J Cardiol. 2010;145(3):601-2.
482   40.   Vad OB, Paludan-Müller C, Ahlberg G, Kalstø SM, Ghouse J, Andreasen L, et al. Loss-
483   of-Function Variants in Cytoskeletal Genes Are Associated with Early-Onset Atrial Fibrillation.
484   J Clin Med. 2020;9(2).
485   41.   Williams ZJ, Velez-Irizarry D, Gardner K, Valberg SJ. Integrated proteomic and
486   transcriptomic profiling identifies aberrant gene and protein expression in the sarcomere,
487   mitochondrial complex I, and the extracellular matrix in Warmblood horses with myofibrillar
488   myopathy. BMC genomics. 2021;22(1):438.
489   42.   Wu C, Zhou Y, Wang M, Dai G, Liu X, Lai L, et al. Bioinformatics Analysis Explores
490   Potential Hub Genes in Nonalcoholic Fatty Liver Disease. Frontiers in genetics. 2021;12:772487.
491   43.   Lee DY, Kang Y, Im NR, Kim B, Kwon TK, Jung KY, et al. Actin-Associated Gene
492   Expression is Associated with Early Regional Metastasis of Tongue Cancer. The Laryngoscope.
493   2021;131(4):813-9.
494   44.   Feng Y, Jiang Y, Wen T, Meng F, Shu X. Identifying Potential Prognostic Markers for
495   Muscle-Invasive Bladder Urothelial Carcinoma by Weighted Gene Co-Expression Network
496   Analysis. Pathology oncology research : POR. 2020;26(2):1063-72.
497   45.   Shou Y, Robinson DM, Amakye DD, Rose KL, Cho YJ, Ligon KL, et al. A five-gene
498   hedgehog signature developed as a patient preselection tool for hedgehog inhibitor therapy in
499   medulloblastoma. Clinical cancer research : an official journal of the American Association for
500   Cancer Research. 2015;21(3):585-93.
501   46.   Vallvé-Juanico J, Houshdaran S, Giudice LC. The endometrial immune environment of
502   women with endometriosis. Human reproduction update. 2019;25(5):564-91.
503   47.   Khan KN, Yamamoto K, Fujishita A, Muto H, Koshiba A, Kuroboshi H, et al.
504   Differential Levels of Regulatory T Cells and T-Helper-17 Cells in Women With Early and
505   Advanced Endometriosis. J Clin Endocrinol Metab. 2019;104(10):4715-29.
506   48.   Le NXH, Loret de Mola JR, Bremer P, Groesch K, Wilson T, Diaz-Sylvester P, et al.
507   Alteration of systemic and uterine endometrial immune populations in patients with
508   endometriosis. American journal of reproductive immunology (New York, NY : 1989).
509   2021;85(3):e13362.

The statistical metrics for key differentially expressed genes (DEGs)

1   Table 1. The statistical metrics for key differentially expressed genes (DEGs).

Gene Symbol	logFC	P.Value	adj.P.Val	t value	Regulation
C7	3.783999053	2.71E-13	1.35E-10	9.285264	up
PTGIS	3.604098022	3.16E-20	3.15E-16	13.64407	up
FABP4	3.468506892	3.05E-16	4.17E-13	11.06606	up
PLA2G2A	3.420962661	2.02E-10	2.06E-08	7.609885	up
MYH11	2.919064902	6.89E-10	5.46E-08	7.301042	up
WISP2	2.750733339	1.73E-10	1.81E-08	7.648918	up
CFH	2.682022262	2.63E-16	4.17E-13	11.10573	up
PDLIM3	2.630060076	1.29E-14	8.06E-12	10.07301	up
C3	2.621335795	2.33E-12	6.85E-10	8.735987	up
CLDN11	2.61447217	6.34E-13	2.53E-10	9.067731	up
PLN	2.549876617	2.77E-11	4.53E-09	8.110245	up
FZD7	2.50539562	2.88E-16	4.17E-13	11.08116	up
DPYSL3	2.334733121	2.79E-18	1.39E-14	12.35555	up
FMO1	2.287811528	1.46E-15	1.62E-12	10.64648	up
FAM129A	2.240230766	1.77E-14	1.04E-11	9.990994	up
S100A8	2.237119405	9.13E-10	6.96E-08	7.230211	up
LY96	2.195123038	3.34E-16	4.17E-13	11.04102	up
CPE	2.160535979	1.47E-12	4.75E-10	8.852691	up
COL8A1	2.148521844	7.63E-12	1.73E-09	8.435667	up
CXCL2	2.082497549	3.79E-06	5.78E-05	5.0803	up
CD163	2.06809014	2.63E-11	4.38E-09	8.12315	up
AOX1	2.026180382	1.98E-09	1.32E-07	7.034879	up
CCL2	2.006001247	2.29E-10	2.20E-08	7.578566	up
RORB	-2.07036228	2.91E-08	1.12E-06	-6.3538	down
TMPRSS4	-2.12081688	4.98E-09	2.79E-07	-6.80216	down
EHF	-2.13597521	7.95E-10	6.25E-08	-7.26498	down
SCNN1A	-2.14340132	7.39E-11	9.84E-09	-7.86305	down
MSX1	-2.17101832	9.74E-10	7.30E-08	-7.21392	down
KIAA1324	-2.18148615	1.16E-10	1.37E-08	-7.74893	down
ASRGL1	-2.24612274	3.73E-10	3.42E-08	-7.45546	down
PRSS8	-2.27282731	2.86E-11	4.61E-09	-8.10186	down
RAB25	-2.35255482	1.18E-11	2.35E-09	-8.32591	down
HSD17B2	-2.36471296	4.09E-07	9.60E-06	-5.67198	down
DEFB1	-2.38849343	3.57E-08	1.32E-06	-6.3015	down
GNLY	-2.44390133	1.59E-08	6.89E-07	-6.50697	down
MMP26	-2.60194621	5.19E-07	1.15E-05	-5.60914	down
AGR2	-2.62956901	1.19E-08	5.42E-07	-6.58042	down

CLDN3	-2.67438295	2.53E-10	2.41E-08	-7.55304	down
CD24	-2.68174109	6.68E-11	9.13E-09	-7.88849	down
GABRP	-2.84571026	1.07E-09	7.82E-08	-7.18967	down
SCGB1D2	-2.97903101	1.58E-05	0.000193	-4.68827	down
SCGB2A1	-3.6734424	5.33E-09	2.93E-07	-6.78496	down

The enriched GO terms of differentially expressed genes

1   Table 2. The enriched GO terms of differentially expressed genes.

ID	Category	GO Name	P.Value	adj.P.Val	Count	Gene ID
GO:0006959	BP	humoral immune response	4.84E-08	5.04E-05	9	CFH/C7/C3/PLA2G2A/CCL2/S100A8/GNLY/DEFB1/CXCL2
GO:0006957	BP	complement activation, alternative pathway	6.03E-06	0.003141	3	CFH/C7/C3
GO:0019730	BP	antimicrobial humoral response	9.26E-06	0.003216	5	PLA2G2A/S100A8/GNLY/DEFB1/CXCL2
GO:0002237	BP	response to molecule of bacterial origin	6.14E-05	0.015984	6	LY96/FMO1/CD24/CCL2/S100A8/CXCL2
GO:0050727	BP	regulation of inflammatory response	8.37E-05	0.01744	6	PTGIS/FABP4/C3/PLA2G2A/S100A8/MMP26
GO:0050729	BP	positive regulation of inflammatory response	0.000145	0.025149	4	FABP4/C3/PLA2G2A/S100A8
GO:0097278	BP	complement-dependent cytotoxicity	0.000176	0.026206	2	CFH/C3
GO:0032496	BP	response to lipopolysaccharide	0.00047	0.059867	5	LY96/FMO1/CCL2/S100A8/CXCL2
GO:0061844	BP	antimicrobial humoral immune response mediated	0.000517	0.059867	3	GNLY/DEFB1/CXCL2

GO:00 16338	BP	by antimicrobial peptide calcium-independent cell-cell adhesion via plasma membrane cell-adhesion molecules	0.000 974	0.092 221	2	CLDN11/CLDN3
GO:00 31941	CC	filamentous actin	0.001 858	0.170 931	2	DPYSL3/PDLIM3
GO:00 05811	CC	lipid droplet	0.015 451	0.170 931	2	FABP4/CLDN11
GO:00 36126	CC	sperm flagellum	0.020 708	0.170 931	2	SCNN1A/DEFB1
GO:00 05884	CC	actin filament	0.021 41	0.170 931	2	DPYSL3/PDLIM3
GO:00 05923	CC	bicellular tight junction	0.022 846	0.170 931	2	CLDN11/CLDN3
GO:00 05788	CC	endoplasmic reticulum lumen	0.023 003	0.170 931	3	FMO1/C3/COL8A1
GO:00 97729	CC	9+2 motile cilium	0.023 948	0.170 931	2	SCNN1A/DEFB1
GO:00 70160	CC	tight junction	0.025 451	0.170 931	2	CLDN11/CLDN3
GO:19 02495	CC	transmembrane transporter complex	0.026 45	0.170 931	3	PLN/SCNN1A/GABRP
GO:00 05859	CC	muscle myosin complex	0.029 564	0.170 931	1	MYH11
GO:00 35325	MF	Toll-like receptor binding	0.000 258	0.014 279	2	LY96/S100A8

GO ID		Term	p-value	p.adjust	Count	Genes
GO:0001664	MF	G protein-coupled receptor binding	0.000271	0.014279	5	FZD7/C3/CCL2/DEFB1/CXCL2
GO:0036041	MF	long-chain fatty acid binding	0.000305	0.014279	2	FABP4/S100A8
GO:0042379	MF	chemokine receptor binding	0.000362	0.014279	3	CCL2/DEFB1/CXCL2
GO:0005504	MF	fatty acid binding	0.002524	0.079747	2	FABP4/S100A8
GO:0048020	MF	CCR chemokine receptor binding	0.003877	0.086681	2	CCL2/DEFB1
GO:0005044	MF	scavenger receptor activity	0.004045	0.086681	2	CD163/TMPRSS4
GO:0008009	MF	chemokine activity	0.004389	0.086681	2	CCL2/CXCL2
GO:0033293	MF	monocarboxylic acid binding	0.00902	0.158341	2	FABP4/S100A8
GO:0038024	MF	cargo receptor activity	0.010022	0.158341	2	CD163/TMPRSS4

The enriched pathway terms of differentially expressed genes

1 Table 3. The enriched pathway terms of differentially expressed genes.

ID	Pathway Name	P.Value	Gene Count	Gene ID
1269250	Regulation of Complement cascade	4.35E-05	3	CFH,C3,C7
SMP00101	Suprofen Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00104	Indomethacin Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00109	Mefanamic acid Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00113	Oxaprozin Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00114	Nabumetone Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00120	Naproxen Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00289	Diflunisal Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00077	Piroxicam Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00085	Ketoprofen Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00086	Ibuprofen Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00094	Sulindac Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00093	Diclofenac Pathway	1.34E-04	2	PLA2G2A,PTGIS
SMP00098	Ketorolac Pathway	1.34E-04	2	PLA2G2A,PTGIS
1457777	Antimicrobial peptides	1.45E-04	4	S100A8,PLA2G2A,DEFB1,GNLY
SMP00102	Bromfenac Pathway	1.78E-04	2	PLA2G2A,PTGIS
M22072	Alternative Complement Pathway	2.28E-04	2	C3,C7
1269241	Complement cascade	3.52E-04	3	CFH,C3,C7
M39649	Complement and Coagulation	4.57E-	3	CFH,C3,C7

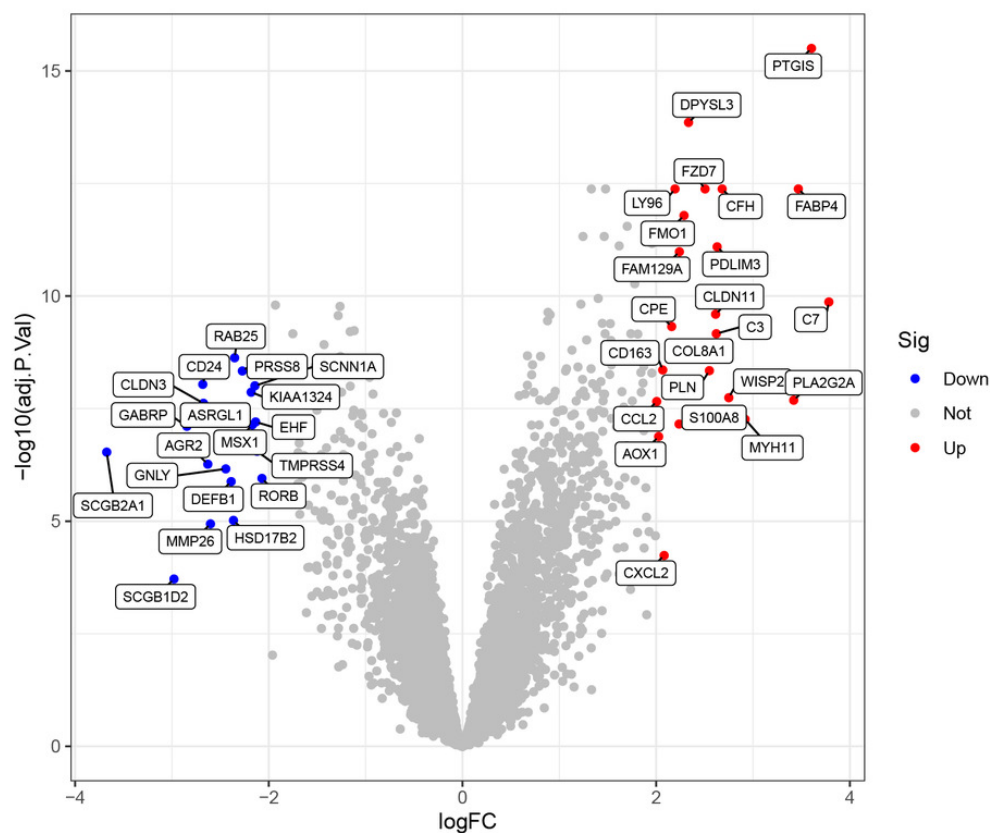
	Cascades	04		
M4732	Lectin Induced Complement Pathway	4.91E-04	2	C3,C7
M7146	Classical Complement Pathway	6.59E-04	2	C3,C7
M16894	Complement and coagulation cascades	7.24E-04	3	CFH,C3,C7
1427857	Regulation of TLR by endogenous ligand	7.52E-04	2	S100A8,LY96
M39770	Epithelial to mesenchymal transition in colorectal cancer	7.92E-04	4	FZD7,TMPRSS4,CLDN11,CLDN3
M3952	Cells and Molecules involved in local acute inflammatory response	8.51E-04	2	C3,C7
M39733	Cells and Molecules involved in local acute inflammatory response	8.51E-04	2	C3,C7
SMP00083	Acetylsalicylic Acid Pathway	8.51E-04	2	PLA2G2A,PTGIS
PW:0000485	eicosanoids metabolic	9.56E-04	2	PLA2G2A,PTGIS
MAP00590	MAP00590 Prostaglandin and leukotriene metabolism	1.07E-03	2	PLA2G2A,PTGIS
83073	Complement and coagulation cascades	1.07E-03	3	CFH,C3,C7
M917	Complement Pathway	1.18E-03	2	C3,C7
SMP00106	Meloxicam Pathway	1.18E-03	2	PLA2G2A,PTGIS
SMP00116	Valdecoxib Pathway	1.18E-03	2	PLA2G2A,PTGIS
SMP00084	Etodolac Pathway	1.18E-03	2	PLA2G2A,PTGIS
SMP00087	Rofecoxib Pathway	1.18E-03	2	PLA2G2A,PTGIS
SMP00096	Celecoxib Pathway	1.18E-03	2	PLA2G2A,PTGIS
M39502	Complement Activation	1.43E-03	2	C3,C7
M39398	Allograft Rejection	1.52E-03	3	C3,GNLY,C7

M264	Endogenous TLR signaling	1.71E-03	2	S100A8,LY96
1474301	IL-17 signaling pathway	1.72E-03	3	S100A8,CCL2,CXCL2
M39382	Eicosanoid Synthesis	2.00E-03	2	PLA2G2A,PTGIS
M39581	Human Complement System	2.06E-03	3	CFH,C3,C7
M40056	16p11.2 distal deletion syndrome	2.16E-03	2	C3,PLN
PW:0000138	Vitamin B6 metabolic	2.58E-03	1	AOX1
1270236	Tight junction interactions	2.84E-03	2	CLDN11,CLDN3
M39341	Spinal Cord Injury	3.47E-03	3	PLA2G2A,CCL2,CXCL2
M11355	Tight junction	4.64E-03	3	CLDN11,CLDN3,MYH11

Figure 1

Figure 1. Differentially expressed genes (DEGs) between endometriosis tissues and normal endometrium.

(A) Heatmap of DEGs. (B) Volcano Plot of DEGs.

A

B

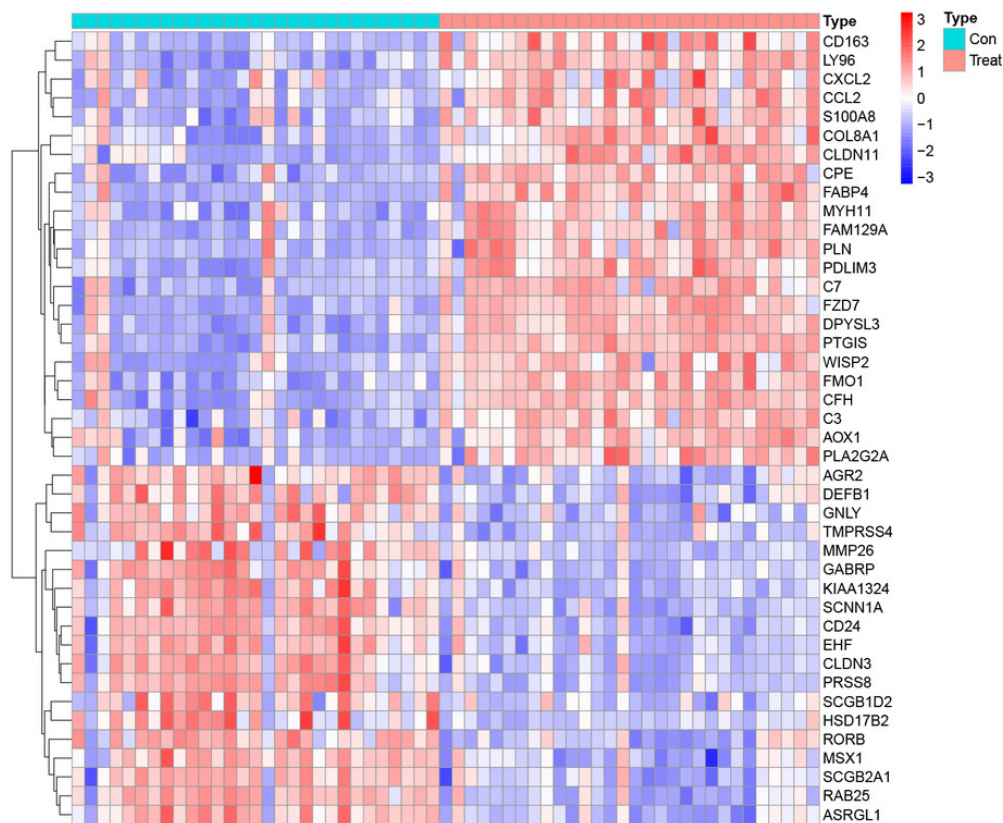



Figure 2

Figure 2. Functional enrichment analyses to identify potential biological processes.

(A□B) Gene ontology enrichment analysis of DEGs between endometriosis and control samples. (C□D) Enrichment analyses via gene set enrichment analysis.

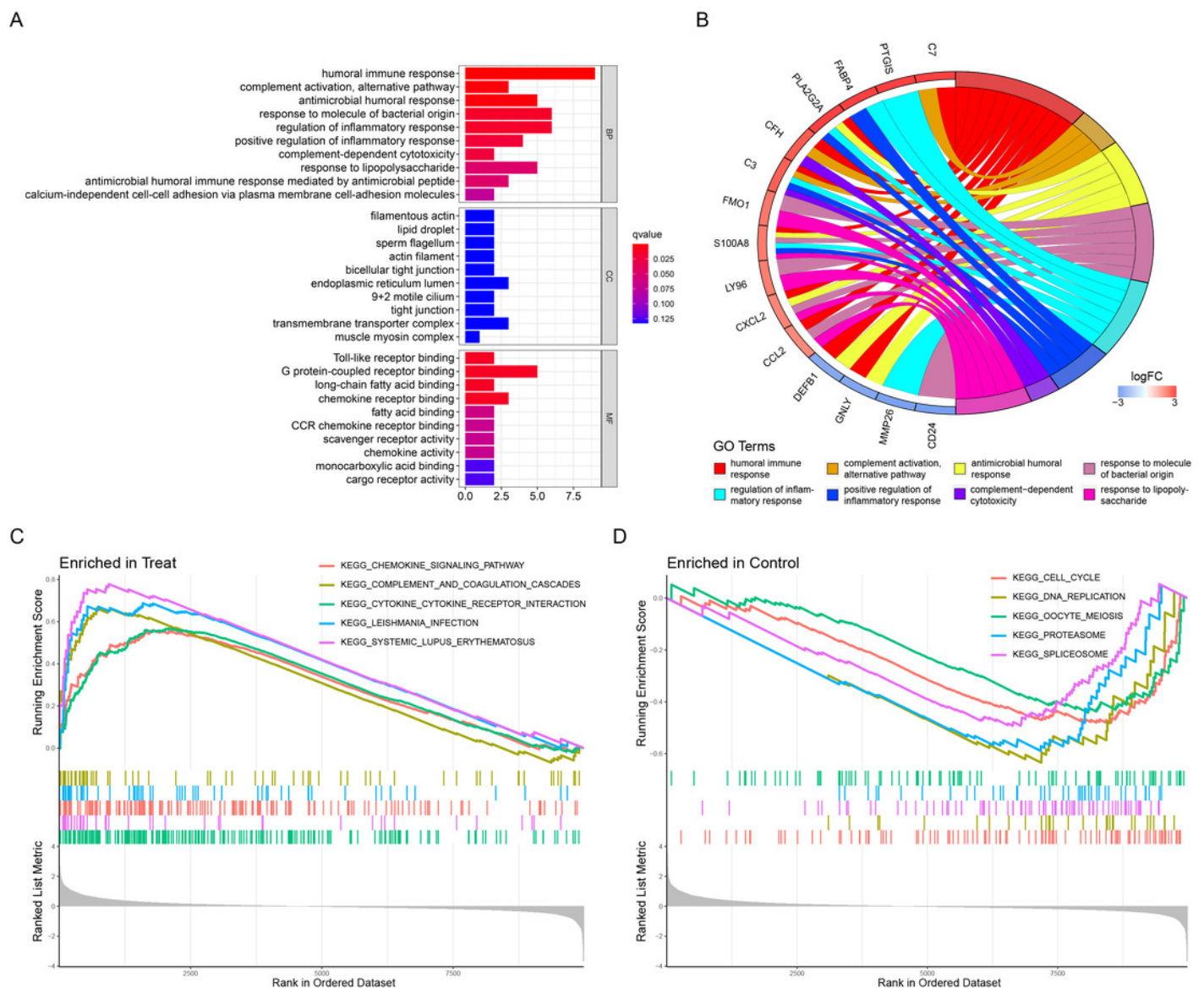



Figure 3

Figure 3. Screening diagnostic biomarker candidates for endometriosis.

(A) A PPI network of DEGs. (B) A plot of biomarkers selection by LASSO regression algorithm. (C) The receiver operating characteristic (ROC) curve of FMO1, PDLIM3, FAM129A, CLDN11, C3, TMPRSS4 and DEFB1 in the metadata cohort. (D) ROC curve of FMO1, PDLIM3, FAM129A, CLDN11, C3, TMPRSS4 and DEFB1 in the GSE120103 dataset.

A

B

C

D

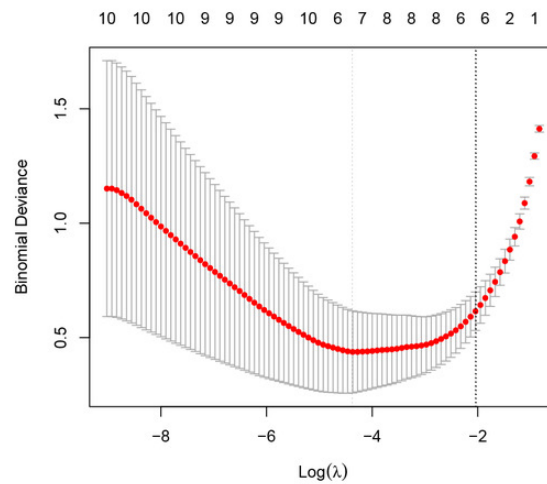



Figure 4

Figure 4. Validation of the expression of diagnostic biomarkers.

(A) The expression level of FMO1, PDLIM3, FAM129A, CLDN11, C3, TMPRSS4 and DEFB1 in the GSE120103 dataset. (B) The mRNA expression level of PDLIM3 in endometriosis and control samples. (C) The protein expression level of PDLIM3 in endometriosis and control samples. Mean±SD, ***P < 0.001, **P < 0.01, and *P < 0.05.

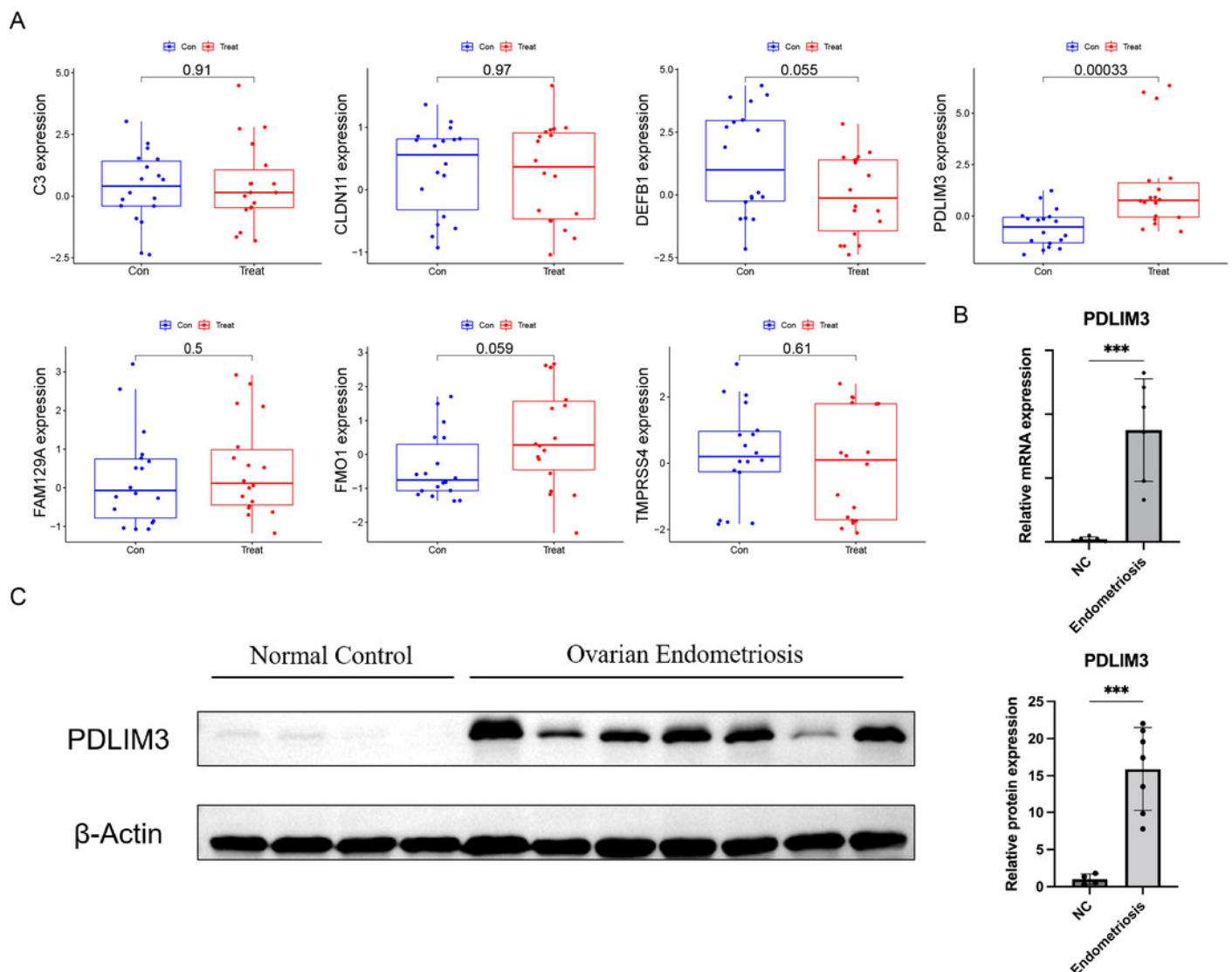
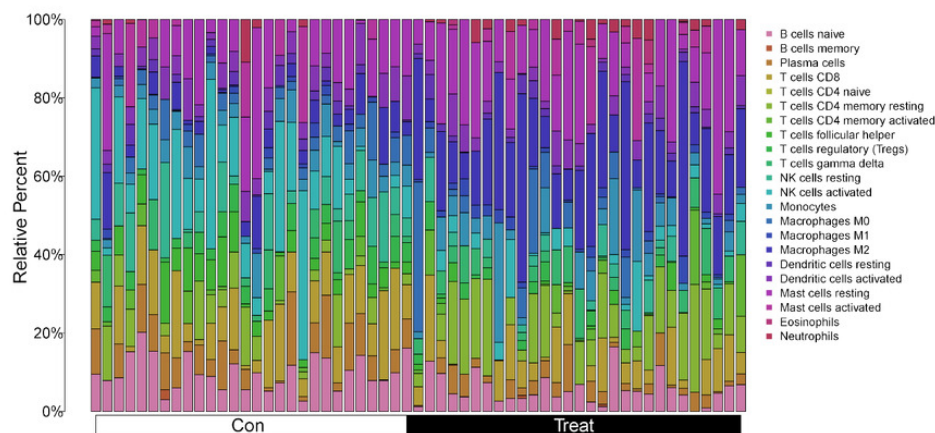
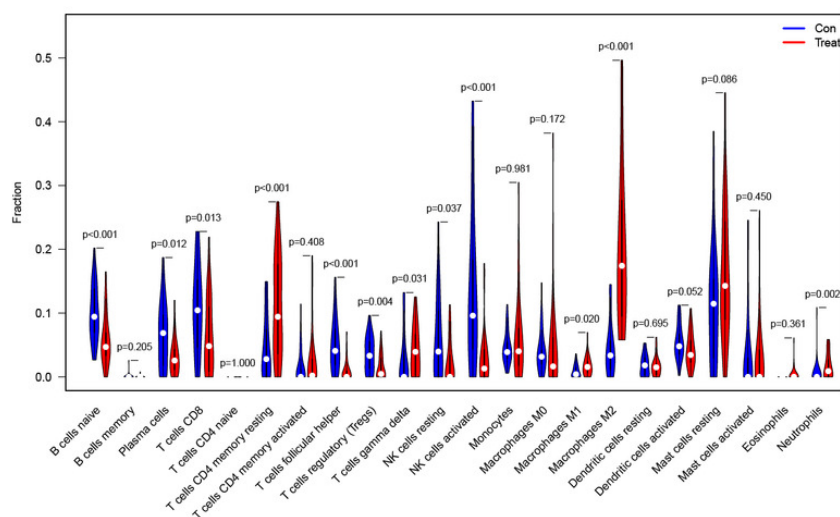



Figure 5

Figure 5. Distribution and visualization of immune cell infiltration.

(A) The composition of immune cells in endometriosis tissues and normal control tissues. (B) Comparison of 22 immune cell subtypes between endometriosis tissues and normal tissues. Blue and red colors represent normal and endometriosis samples, respectively. (C) Correlation matrix of all 22 immune cell subtype compositions. Both horizontal and vertical axes demonstrate immune cell subtypes. Immune cell subtype compositions (higher, lower, and same correlation levels are displayed in red, blue, and white, respectively).

A

B

C

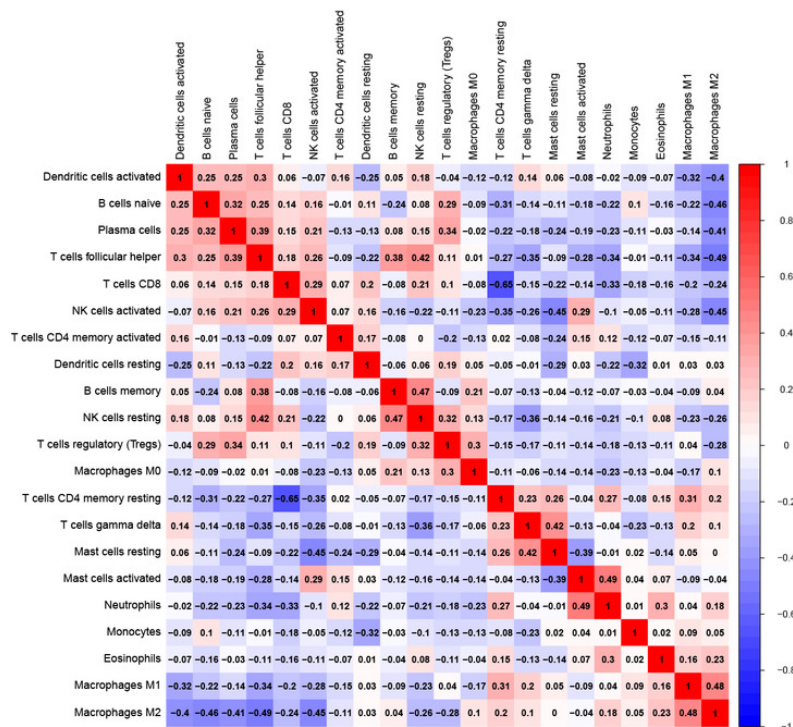



Figure 6

Figure 6. Correlation between PDLIM3 and infiltrating immune cells in endometriosis.

(A) Correlation between PDLIM3 and single type of immune cell. (B) A comprehensive correlation analysis of PDLIM3 and infiltrating immune cells in endometriosis.

A

B