



Bioinformatics Analysis Identifies Molecular Markers Regulating Development and Progression of Endometriosis and Potential Therapeutic Drugs

Ying Peng, Cheng Peng, Zheng Fang and Gang Chen*

Department of Obstetrics and Gynecology, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China

OPEN ACCESS

Edited by:

Peng-Hui Peter Wang,
Taipei Veterans General Hospital,
Taiwan

Reviewed by:

Md. Rakibul Islam,
Daffodil International University,
Bangladesh
Esra Gov,
Adana Alparslan Turkes Science
and Technology University, Turkey

*Correspondence:

Gang Chen
medchen@163.com

Specialty section:

This article was submitted to
Systems Biology Archive,
a section of the journal
Frontiers in Genetics

Received: 30 October 2020

Accepted: 07 July 2021

Published: 04 August 2021

Citation:

Peng Y, Peng C, Fang Z and
Chen G (2021) Bioinformatics
Analysis Identifies Molecular Markers
Regulating Development
and Progression of Endometriosis
and Potential Therapeutic Drugs.
Front. Genet. 12:622683.
doi: 10.3389/fgene.2021.622683

Endometriosis, a common disease that presents as polymorphism, invasiveness, and extensiveness, with clinical manifestations including dysmenorrhea, infertility, and menstrual abnormalities, seriously affects quality of life in women. To date, its underlying etiological mechanism of action and the associated regulatory genes remain unclear. This study aimed to identify molecular markers and elucidate mechanisms underlying the development and progression of endometriosis. Specifically, we downloaded five microarray expression datasets, namely, GSE11691, GSE23339, GSE25628, GSE7305, and GSE105764, from the Gene Expression Omnibus (GEO) database. These datasets, obtained from endometriosis tissues, alongside normal controls, were subjected to in-depth bioinformatics analysis for identification of differentially expressed genes (DEGs), followed by analysis of their function and pathways *via* gene ontology (GO) and KEGG pathway enrichment analyses. Moreover, we constructed a protein–protein interaction (PPI) network to explore the hub genes and modules, and then applied machine learning algorithms support vector machine-recursive feature elimination and least absolute shrinkage and selection operator (LASSO) analysis to identify key genes. Furthermore, we adopted the CIBERSORTx algorithm to estimate levels of immune cell infiltration while the connective map (CMAP) database was used to identify potential therapeutic drugs in endometriosis. As a result, a total of 423 DEGs, namely, 233 and 190 upregulated and downregulated, were identified. On the other hand, a total of 1,733 PPIs were obtained from the PPI network. The DEGs were mainly enriched in immune-related mechanisms. Furthermore, machine learning and LASSO algorithms identified three key genes, namely, apelin receptor (*APLNR*), C–C motif chemokine ligand 21 (*CCL21*), and Fc fragment of IgG receptor IIa (*FCGR2A*). Furthermore, 16 small molecular compounds associated with endometriosis treatment were identified, and their mechanism of action was also revealed. Taken together, the findings of this study provide new insights into the molecular factors regulating

occurrence and progression of endometriosis and its underlying mechanism of action. The identified therapeutic drugs and molecular markers may have clinical significance in early diagnosis of endometriosis.

Keywords: endometriosis, bioinformatics analysis, differentially expressed genes, immune mechanism, molecular markers, potential drugs

INTRODUCTION

Endometriosis (EM), in which active endometrium is implanted in any site outside the uterine cavity, is a common gynecological disease characterized by chronic pelvic pain, dysmenorrhea, and infertility (Aghajanova and Giudice, 2010). Approximately 1% of all EM cases exhibit malignant transformation (Guo et al., 2012), which greatly reduces the quality of life of women at a childbearing age. Treatment of EM is challenging due to a high postoperative recurrence rate and the inhibitory effect of hormone drugs on ovarian function (Wu et al., 2017). Currently, surgical resection and hormone suppression represent the gold standard treatment for EM, although these are limited by high occurrence of side effects (Vercellini et al., 2013), recurrence rates, and treatment costs (Simoens et al., 2012). Therefore, identification of novel molecular factors, coupled with unraveling of the underlying mechanisms of EM will aid in development of effective treatment therapies.

Numerous hypotheses have been proposed to describe the underlying molecular mechanisms of EM, including the menstrual reflux theory (Kruitwagen et al., 1991), ectopic implantation, epigenetic as well as immune and inflammatory factors, immunodeficiency (Ahn et al., 2015), eutopic endometrium determinism, and stem cell factors (Burney and Giudice, 2012). To date, however, no single theory can explain all the causes of EM, although reflux and ectopic EM implantation represent the most commonly accepted theories. Notably, only a handful of women have EM (Hill et al., 2020). Advancement in molecular biology and pathophysiology technologies has revealed that pathogenesis of EM is closely associated with the body's immune imbalance (Burney and Giudice, 2012). Previous studies have shown that cell-mediated and humoral immune changes contribute to disease progression in women with EM. Notably, functional changes in immune components, such as monocytes/macrophages, natural killer (NK) cells, T lymphocytes, B cells, and cytokines, in peritoneal fluid of women with EM have been described (Zou et al., 2021). Although numerous and brisk EMaICI, comprising several types of immune cells in all EM forms, have revealed acute immunological reactions within the microenvironment of EM lesions (Herington et al., 2011), the role of these changes in disease development remains unclear. Previous studies have demonstrated that overexpression and downregulation of antiapoptotic and proapoptotic factors, respectively, may interfere with peritoneal homeostasis (Erekat et al., 2018), while the mitogen-activated protein kinase (MAPK) signaling pathway has been shown to play a key role in disease occurrence (Lotfaliansaremi et al., 2020). Moreover, apoptosis pathways of Fas-FasL and tumor necrosis factor (TNF)- α seem to play a key role in immune monitoring of the peritoneal microenvironment (Yamada et al., 2017). Overall,

these changes may prevent clearance of endometrial cells that reach the peritoneal cavity, thereby allowing their implantation and development. The current consensus is that changes in immune-related cell function play a crucial role in occurrence and progression of EM (Ahn et al., 2015).

Although numerous studies have reported occurrence of immune abnormalities, the role of the immune system in EM is not fully understood, although some evidence has demonstrated that impaired immune homeostasis is associated with increased implantation, proliferation, and angiogenesis of ectopic endometrium (Holtan et al., 2009). Although EM is a benign disease, EM cells exhibit many characteristics similar to those observed in malignant tumors, including migration and aggressiveness. Therefore, exploring immune dysfunction in EM may help elucidate its role in pathogenesis of the disease and generate relevant insights to guide development of new treatment strategies. In addition, with the rapid development of high-throughput sequencing technologies, more and more potential biomarkers associated with EM were identified. Liu et al. (2015) found that the EGF, IL- β and AGTR1 play an important role in the pathogenesis of EM by using the microarray dataset. Cui et al. (2020) employed three microarray datasets, constructed a network, and revealed that YAP/TAZ was upregulated, and MOB1, pMOB1, SAV1, LATS1, and LATS2 were downregulated in endometrium. Dai et al. (2019) using four microarray datasets (GSE11691, GSE23339, GSE25628, and GSE78851) identified a number of differentially expressed genes (DEGs). However, these studies lack in-depth analysis and have no external validation dataset to evaluate their expression level. Therefore, a comprehensive analysis of the potential biomarkers based on a number of microarray datasets is in urgent demand.

In the present study, we downloaded five microarray datasets (GSE11691, GSE23339, GSE25628, GSE105764, and GSE7305) from the Gene Expression Omnibus (GEO) database to explore the potential biomarkers. We employed in-depth bioinformatics analysis to identify candidate genes and key signaling pathways regulating occurrence and progression of EM, as well as possible therapeutic drugs for its treatment. Overall, these findings provide new insights into the underlying molecular mechanisms and potential candidate drugs for treatment of EM.

MATERIALS AND METHODS

Data Collection

We employed the "GEOquery" package, implemented in R software version 3.6.1¹ to download GSE11691, GSE23339, GSE25628, GSE105764, and GSE7305 datasets from the GEO

¹<http://r-project.org/>

database² using “endometriosis” as the keyword (Davis and Meltzer, 2007). A total of 87 samples were obtained: GSE11691 contained 9 healthy and 9 diseased samples (Hull et al., 2008), GSE105764 included 8 EM samples and 8 normal samples (Zhao et al., 2018), GSE23339 comprised 9 and 10 healthy and diseased samples (Hawkins et al., 2011), GSE7305 consisted of 10 diseased samples and 10 healthy samples (Hever et al., 2007), while GSE25628 contained 6 healthy and 8 diseased samples (Heydari et al., 2020). Among them, GSE11691, GSE23339, and GSE25628 served as a training dataset, while GSE7305 and GSE105764 were used as an external dataset.

Data Preprocessing and Differential Gene Expression Analysis

The batch effect in the training dataset was removed using the “sva” R package.³ The effect of the correction between samples was calculated by two-dimensional principal component analysis (PCA) clustering. DEGs were screened using “limma” R package and the volcano plots of the DEGs were generated using “ggplot2” to show differential expression patterns (Ritchie et al., 2015). Additionally, the genes with adjusted p -value < 0.05 and $|\log_2FC| > 1$ were considered statistically significant.

Gene Set Enrichment Analysis

Functional and KEGG pathway enrichment analyses of the DEGs were performed using “clusterProfiler” R package (Yu, 2018). The significant pathways were identified with the adjusted p -value < 0.05 .

Construction of a Protein–Protein Interaction Network

To construct a protein–protein interaction (PPI) network, we uploaded all the DEGs to the STRING protein database.⁴ The gene interaction score > 0.7 was used to construct a network. We then used the MCODE algorithm and cytoscape software to identify and visualize the hub networks and hub genes (Bader and Hogue, 2003; Shannon et al., 2003). As a result, the top two hub networks that contained 22 hub genes were subsequently used for the downstream analysis.

Screening and Verification of Key Genes

The key genes were identified using least absolute shrinkage and selection operator (LASSO) regression analysis and support vector machine (SVM) analysis based on the genes of the top two subnetworks, respectively. The LASSO is a dimension-reduction algorithm and showed a higher superiority compared to regression analysis (Tibshirani, 1996). The support vector machine-recursive feature elimination (SVM-RFE) was a machine learning algorithm that was applied to identify the best variables *via* eliminating feature vectors generated by SVM and further demonstrated the biomarkers in EM through the “e1071” R package (Sanz et al., 2018). To further evaluate the expression

profile, we used the GSE7305 and GSE105764 dataset as an external dataset. The p -value was calculated using the Wilcoxon rank test analysis. In addition, we also used the “pROC”⁵ to estimate the performance between EM and normal tissue.

Functional Identification of Key Genes

To further investigate functions of the key genes, we performed gene set enrichment analysis (GSEA) analysis based on each key gene’s expression profile and then classified them into high and low expression groups based on median gene expression levels. The significant pathways were identified based on the p -value < 0.05 .

Correlation Between Gene Expression and Immunity

We uploaded the expression data for all training dataset samples to the CIBERSORTx⁶ and calculated the 22 immune cells’ infiltration level (Newman et al., 2019). We then performed correlation analysis between the 22 immune cells and the key genes.

Identification of Potential Drugs

To identify potential therapeutic drugs for EM treatment, we uploaded the DEGs to the connective map (CMAP) database⁷ and then applied enrichment analysis to pull out significantly enriched drugs (Musa et al., 2017). Generally, a negative enrichment value indicates that the drug is more likely to treat the disease, with a greater magnitude of the enrichment value implying more relevance of the drug to the disease. Therefore, we further screened for drugs with an enrichment of less than -0.5 .

RESULTS

DEGs in EM

To comprehensively investigate gene expression in EM, we first integrated three datasets, namely GSE11691, GSE23339, and GSE25628, into a complete dataset after removing batch effects using the “sva” R package. We then used PCA analysis to determine the distribution of the samples before and after correction. **Figure 1A** shows that the distribution of the three datasets before batch effect was removed, and **Figure 1B** confirms that all confounding factors were successfully removed from the corrected samples (**Figures 1A,B**). In addition, The Differential expression analysis revealed a total of 423 DEGs across the combined dataset, namely, 233 and 190 upregulated and downregulated genes (**Figures 2A,B**).

Function and Pathway Enrichment of the DEGs

To investigate the molecular function and pathways that are involved in the DEGs, we then conducted an enrichment analysis

²<https://www.ncbi.nlm.nih.gov/geo/>

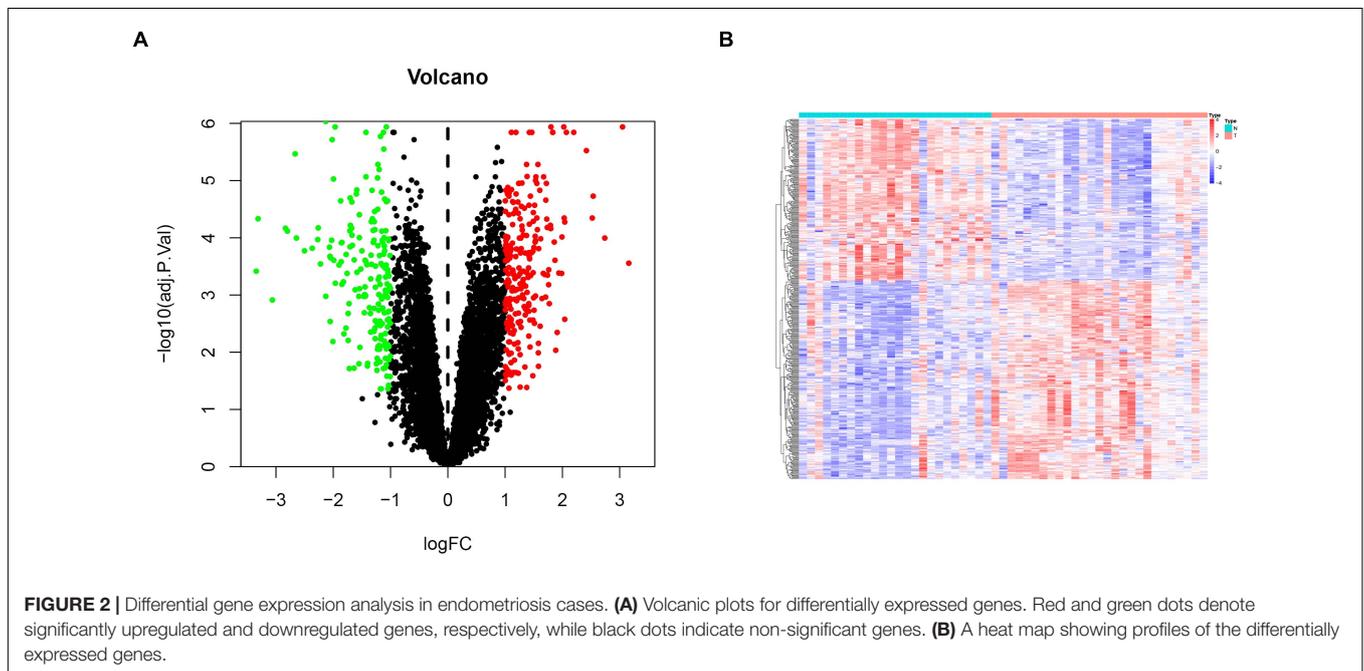
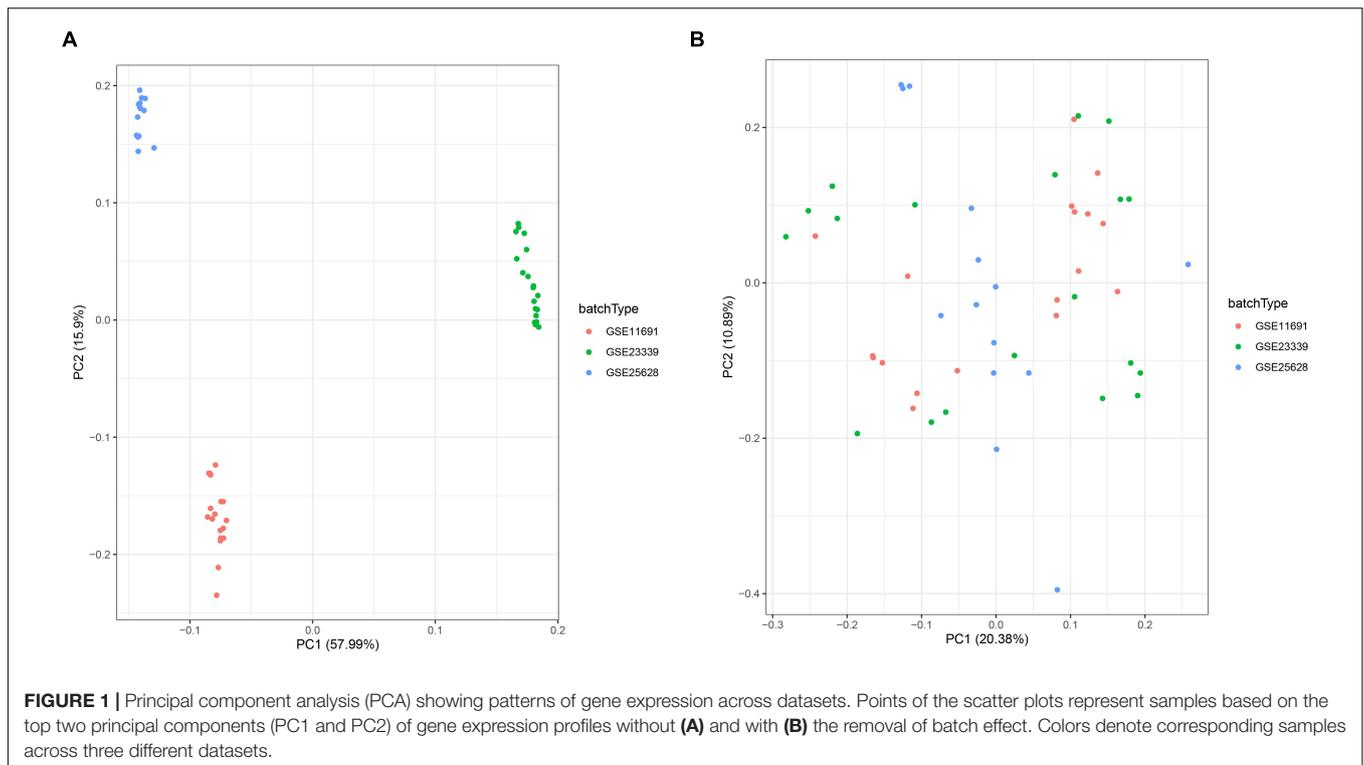
³<https://bioconductor.org/packages/release/bioc/html/sva.html>

⁴<https://string-db.org/>

⁵<https://cran.r-project.org/web/packages/pROC/index.html>

⁶<https://cibersortx.stanford.edu/>

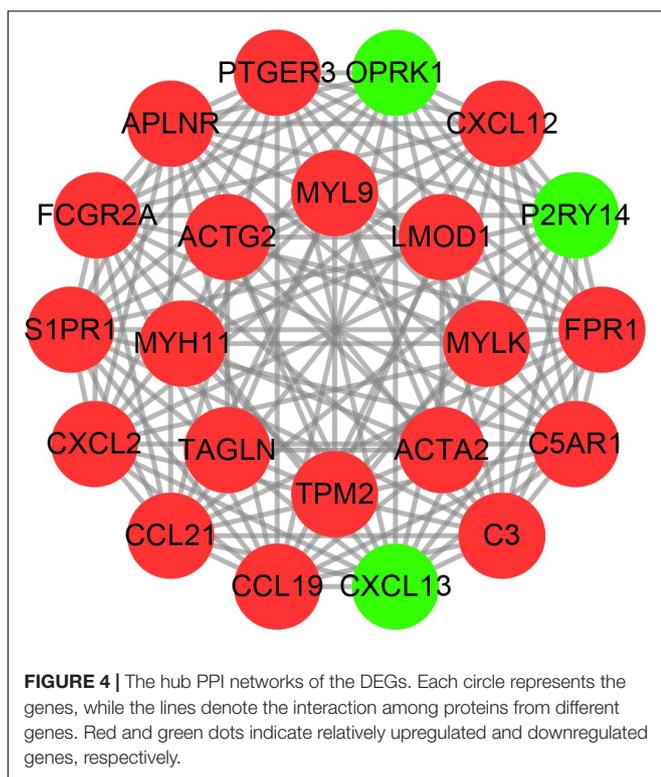
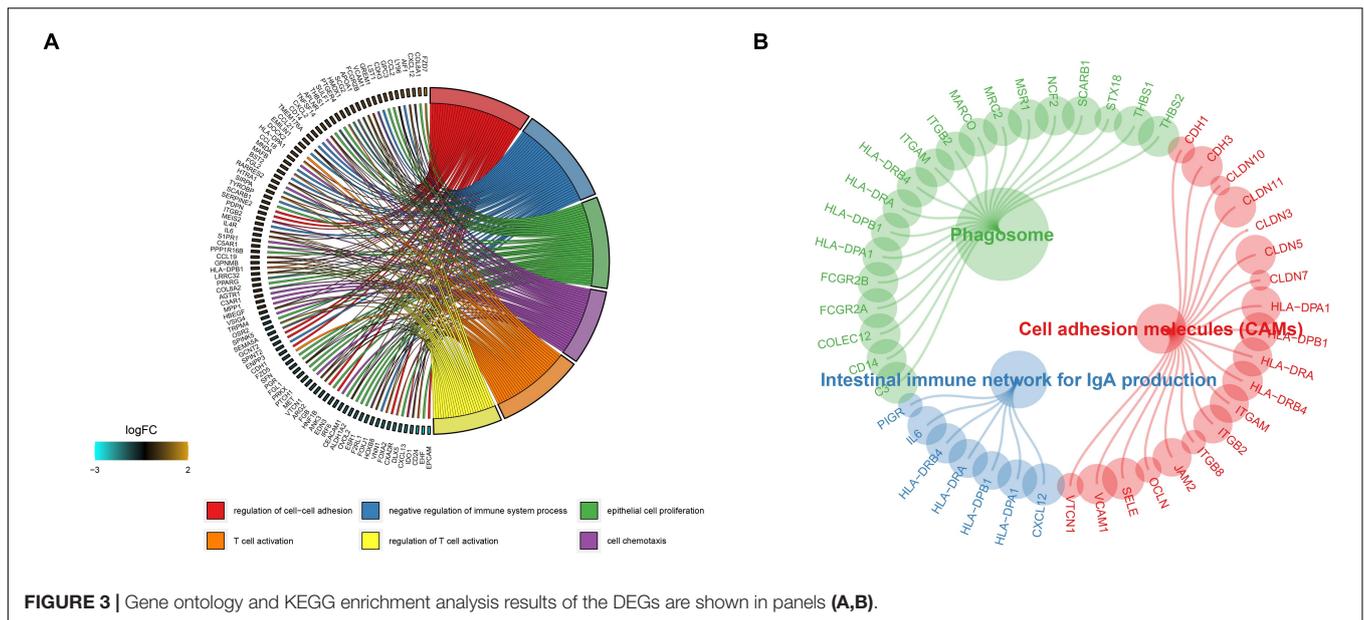
⁷<https://portals.broadinstitute.org/cmap/>



using the clusterProfiler R package. Based on the screening criterion, we discovered that the genes are mainly involved in T cell activation, regulation T cell activation, and cell chemotaxis in gene ontology (GO) results (**Figure 3A**), while phagosome, cell adhesion molecules (CAMs), and the intestinal immune network for IgA production pathway were enriched in KEGG results (**Figure 3B**).

Construction of PPI Networks

We used the string online database to explore the interaction network of all DEGs and further visualized it using the cytoscape software. We then applied the MCODE algorithm to explore the hub networks and genes. As a result, the top two subnetworks with 22 hub genes were identified, namely, CXCL13, CCL19, APLNR, PTGER3, P2RY14, OPRK1, CCL21, FPR1, CXCL2,

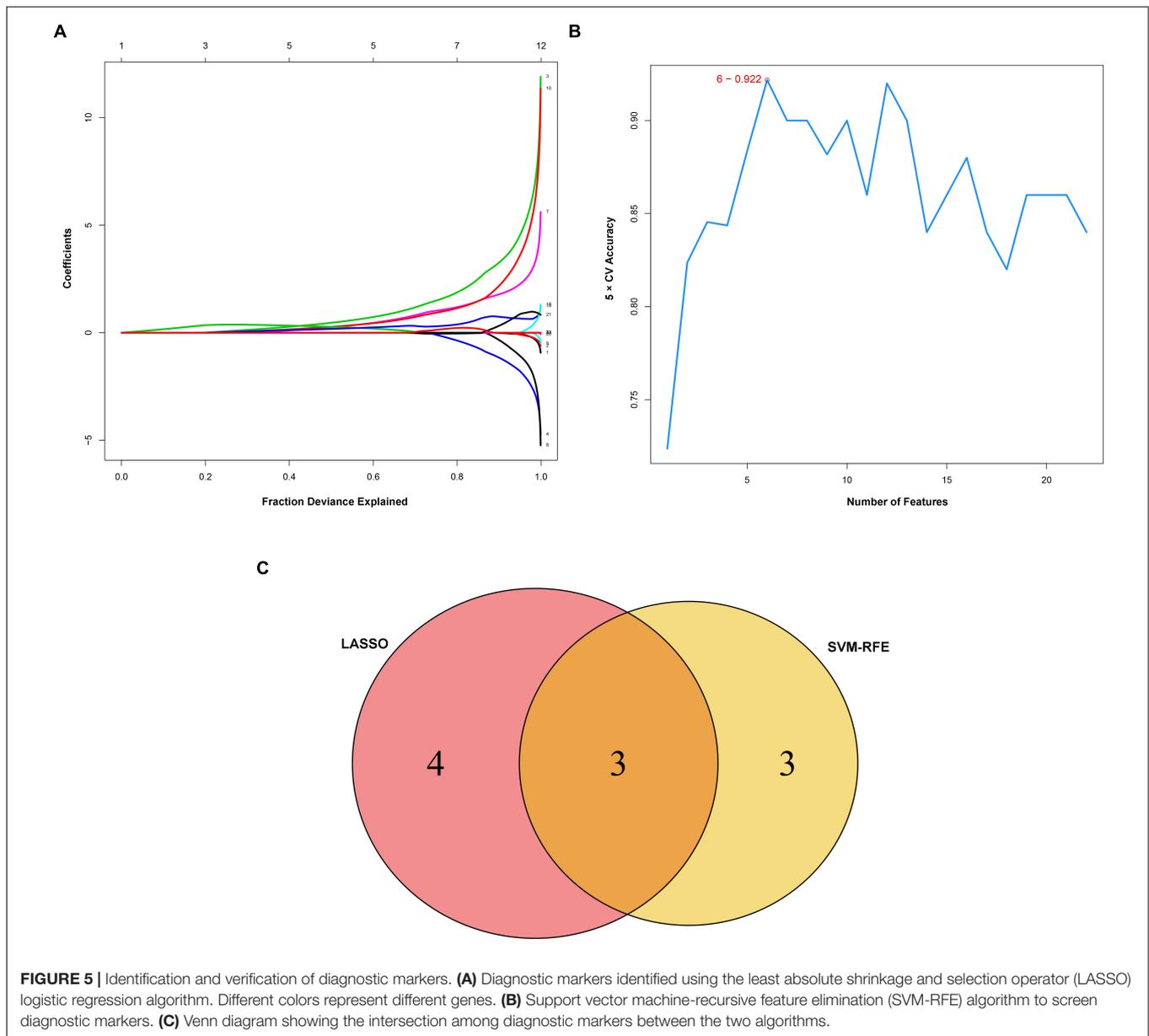


FCGR2A, S1PR1, C3, CXCL12, C5AR1, TAGLN, LMOD1, TPM2, ACTG2, MYL9, MYH11, MYLK, and ACTA2 (Figure 4).

Key Genes and Enriched Pathways in EM

Considering too much hub genes were identified from the hub networks, we then used the LASSO regression analysis and SVM-RFE analysis to make a feature selection. We firstly used

the LASSO analysis to screen the key genes from the 22 hub genes, and 7 genes, namely, CXCL13, APLNR, CCL21, FCGR2A, CXCL12, TAGLN and ACTA2, were identified (Figure 5A). Three genes including APLNR, CCL21, and FCGR2A were identified from 22 hub genes by the SVM-RFE analysis (Figure 5B). The combined results from LASSO regression and SVM-RFE analyses revealed three key genes in ME, namely, apelin receptor (APLNR), C-C motif chemokine ligand 21 (CCL21), and Fc fragment of IgG receptor IIa (FCGR2A) (Figure 5C). GSEA analysis of expression of the three genes in APLNR revealed significant enrichment of focal adhesion and ECM-receptor interaction in the high expression group, and none in the low expression group (Figure 6A). In the CCL21 group, genes in the high expression group were significantly enriched in the CAMs pathway, whereas those in the low expression group were significantly enriched in cell cycle, mismatch repair, P53 signaling pathway, and ubiquitin-mediated proteolysis (Figures 6B,C). In the FCGR2A group, the high expression group was significantly enriched in B cell receptor and chemokine signaling pathways, cytokine-cytokine receptor interaction, and Jak Stat and NOD-like receptor signaling pathways. On the other hand, genes in the low expression group were significantly enriched in cell cycle, DNA replication, nucleotide excision repair, oocyte meiosis, spliceosome, and ubiquitin-mediated proteolysis pathways (Figures 6D,E). In addition, we also adopted the CIBERSORTx algorithm to explore levels of 22 immune cells' infiltration, and found that T follicular helper cells, gamma delta T cells, resting NK cells, activated NK cells, and M2 macrophages have a significant difference between normal and EM tissues (Figure 7). Since these three genes were significantly involved in immune-related pathways, we further evaluated their association with immune cells. Results showed that APLNR was most positively and negatively correlated with M2 macrophages and activated NK cells, respectively. On the other hand, CCL21 was most positively and negatively correlated



with gamma delta T cells and plasma cells, respectively. Similarly, FCGR2A was most positively and negatively correlated with M2 macrophages and follicular helper T cells, respectively (Figure 8).

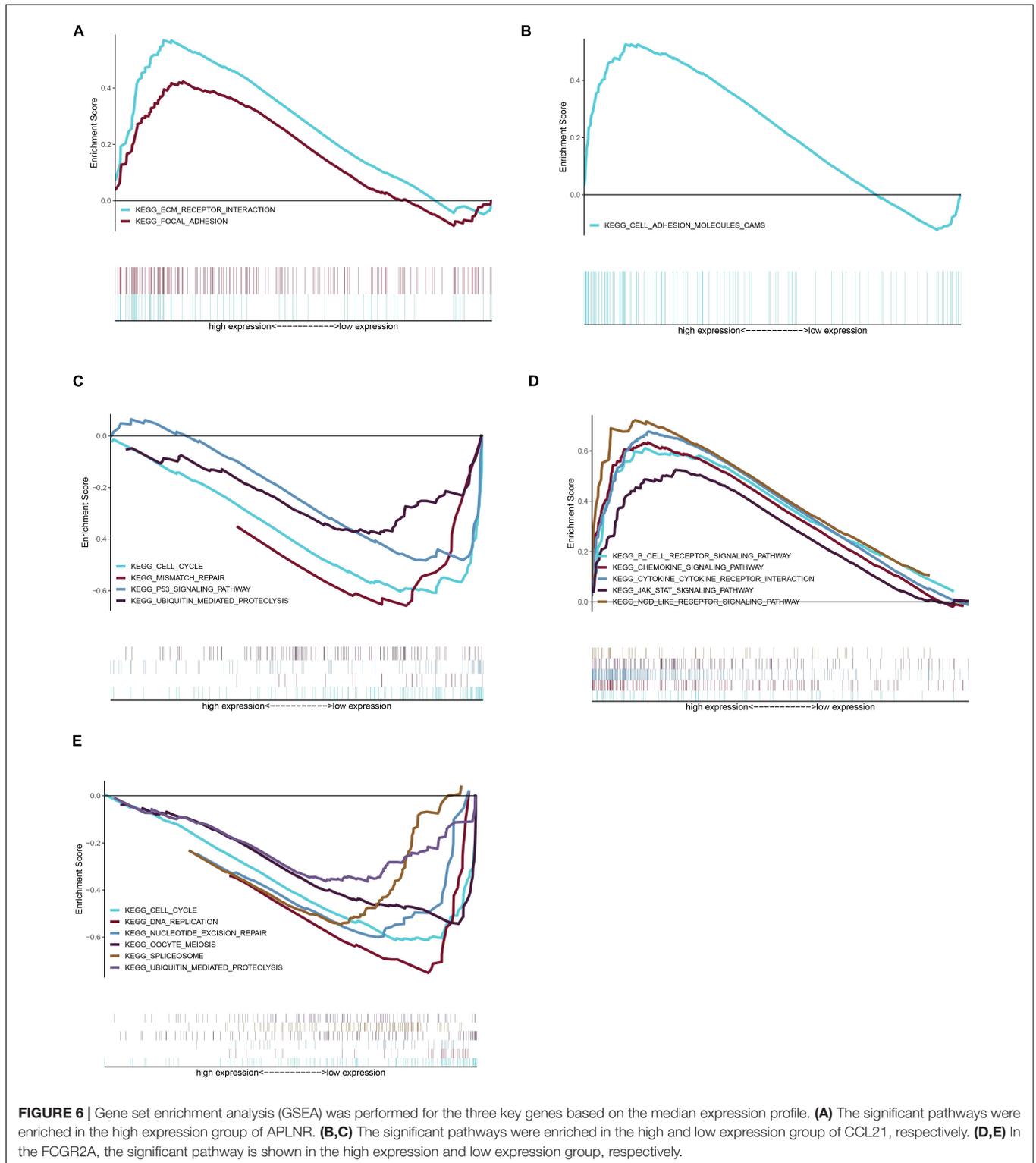
Validation of the Key Genes

To ensure reliable and stability of the results, we further validated expression of the three key genes using two external datasets (GSE7305 and GSE105764). Results revealed that CCL21 and FCGR2A were significantly upregulated in diseased relative to normal tissues. Conversely, APLNR was slightly, but not significantly, upregulated in diseased relative to normal tissues. Overall, these results corroborated the above findings, thereby affirming their reliability (Figures 9A–F). Moreover, to assess the discrimination power of the three genes in the EM and normal tissue, we also performed a ROC analysis on the two

external datasets. Interestingly, the three genes showed a high performance to distinguish the EM and normal tissue, indicating that the three genes have potential as a diagnostic biomarker (Supplementary Figure 1).

Identification of Small Molecular Compounds for Treatment of EM

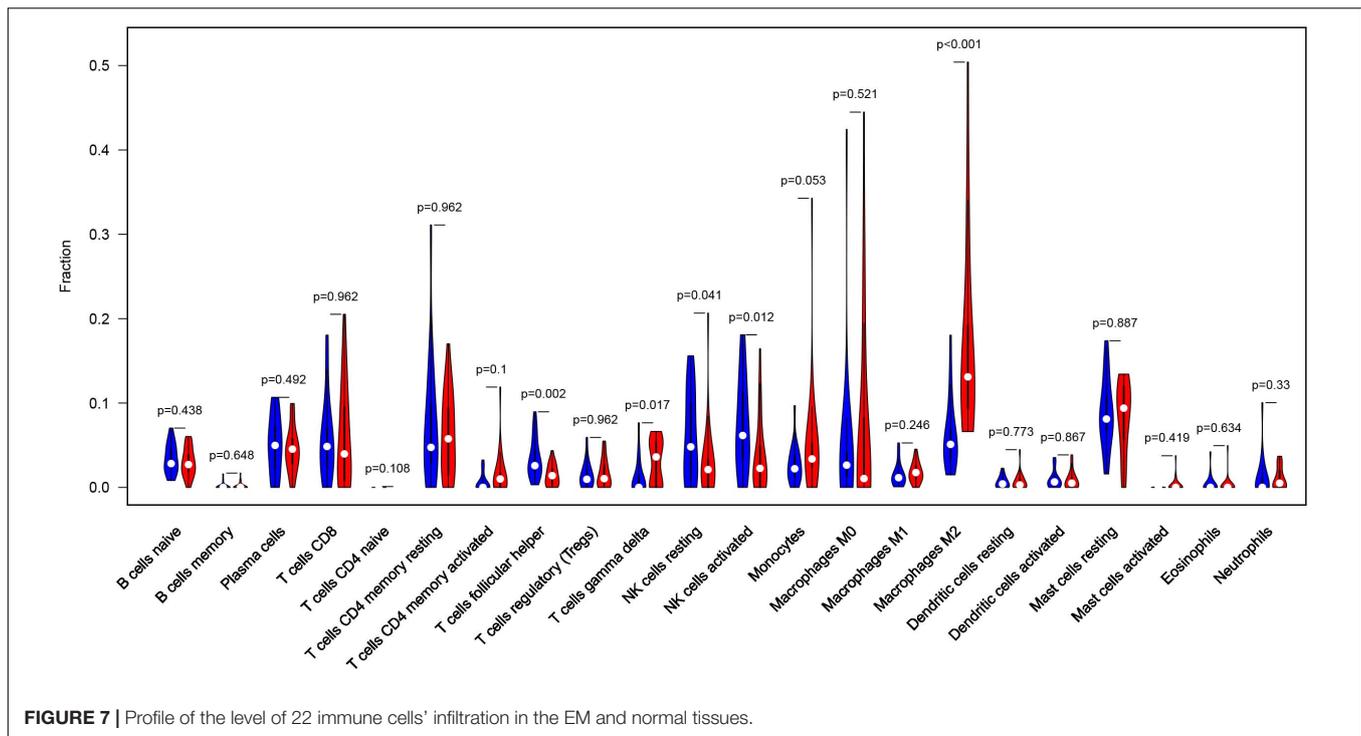
We further identified potential therapeutic drugs for EM treatment by screening the identified DEGs against the CMAP database. Results revealed a total of 47 potential small molecular compounds. Analysis of the mechanism of action of each drug revealed that 21 drugs, namely, W-13, phenazone, metixene, cinchonine, tetracycline, dihydroergocristine, quinpirole, lovastatin, memantine,



metrizamide, naringenin, pentoxifylline, lobeline, zaprinast, chlortalidone, SC-560, oxybenzone, fasudil, fluoxetine, fludrocortisone, and vanoxerine, have the capacity for treating EM (**Figure 10**).

DISCUSSION

In the present study, we identified DEGs across four datasets obtained from EM and normal tissues. These DEGs included

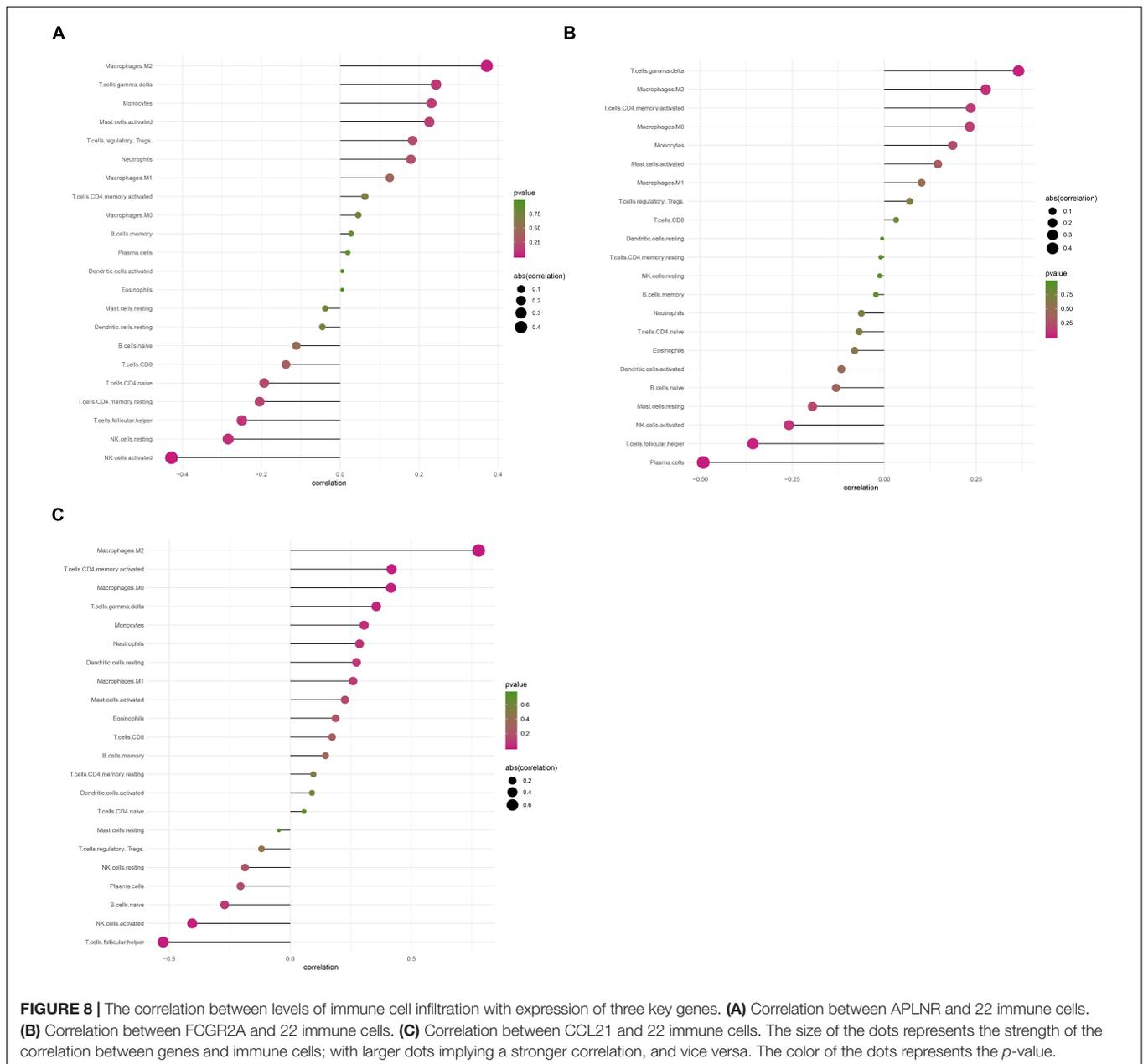


233 upregulated and 190 downregulated genes, respectively. GO and KEGG analyses revealed that these DEGs were involved in regulating immune-related signaling pathways, such as T-cell activation and regulation, as well as cell chemotaxis. Results from LASSO analysis revealed three key genes, namely, APLNR, CCL21, and FCGR2A. Among them, APLNR was firstly identified as a class A G protein-coupled receptor in 1993, which contained 380 amino acids. Ho et al. (2017) demonstrated that knocking out APLNR in pregnant mice resulted in PE-like symptoms due to defective placental angiogenesis, and this was accompanied by poor placental angiogenesis, elevated apoptosis, and decreased proliferation. Results from analysis of developmental model organisms have suggested that the Apelin/APJ pathway plays an important role in embryo angiogenesis (Kidoya et al., 2008). For example, dysregulated APLNR was associated with growth and migration, as well as cell cycle progression, in ovarian clear cell carcinoma (OCCC) cell lines (Xu and Shen, 2018). However, nothing is known regarding the role of APELA in EM.

C-C motif chemokine ligand 21, a homeostatic lymphoid chemokine that is widely expressed in fibroblasts, smooth muscles, as well as intravascular, T, and dendritic cells, plays an essential role in mediating activated DCs and T cells to lymph system niches and triggering subsequent immune response to foreign antigens (Lin et al., 2014). A previous study showed that poor differentiation was associated with larger tumor size in patients with CCL21-positive gallbladder cancer. Moreover, positive CBS and CCL21 expression are closely associated with clinical severity and poor prognosis in GBC, suggesting that it can be a marker for diagnosis of AC and SC/ASC type in GBC (Li et al., 2020). A CC chemokine receptor 7 (CCR7) and CCL21

are abundant in oral squamous cell carcinoma (OSCC) tissues, where they regulate EMT process and promote OSCC desiccation through activation of the JAK2/STAT3 signaling pathway. CCR7 expression has been associated with poor prognosis of OSCC, while CCL21/CCR7 may be an effective target for prevention and treatment of the condition (Chen et al., 2020). Previous evidence has shown that CCL21 enhances GC progression via the MALAT1/SRSF1/mTOR axis, thereby providing a novel therapeutic target for the treatment of GC. Moreover, CCL21-mediated activation of the MALAT1/SRSF1/mTOR axis underpins the development of gastric carcinoma. In fact, Fu et al. (2021) found that CCL21 is significantly upregulated in the EM relative to normal tissues, which is consistent with our results.

The human FCGR2A gene contains at least 800 single-nucleotide polymorphisms (SNPs), located on chromosome 1q23.3 (Velissari et al., 2015). Under physiological conditions, activated FCGR2A is constitutively expressed in platelets, neutrophils, macrophages, and dendritic cells, thereby providing an important link between cellular and humoral immunity through a cross-talk between IgG antibody-antigen complexes (Gillis et al., 2014). In fact, this phenomenon is well documented in a number of autoimmune diseases, including rheumatoid arthritis (Jiménez Morales et al., 2019), autoimmune thyroid disease (Mestiri et al., 2020), and idiopathic nephrotic syndrome in children (Rossi et al., 2017). Moreover, Lee et al. (2016) reported that FCGR2A (rs1801274) polymorphism was associated with better response to anti-TNF- α therapy in patients with spondyloarthritis, psoriasis, and CD. Previous studies have also demonstrated that FCGR2A is the susceptibility gene for Kawasaki disease (KD), while levels in methylation of its CpG site promoter could act as a crucial biomarker for

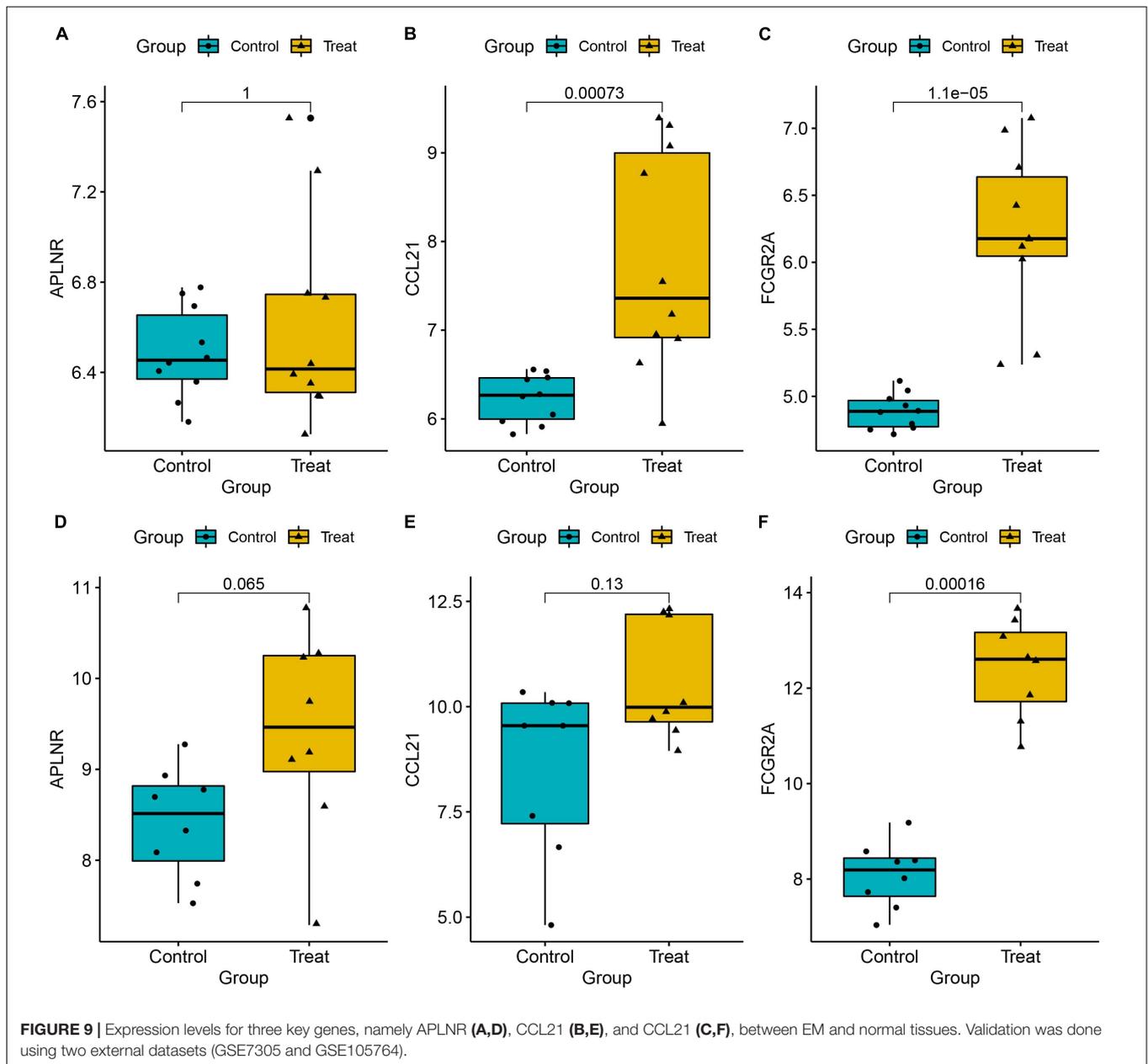


optimizing intravenous immunoglobulin (IVIG) therapy (Kuo et al., 2015). To date, however, the relationship between FCGR2A and EM remains unknown.

In addition, we also evaluated the landscape of 22 immune cells' infiltration level of EM and normal samples. We found that the T follicular helper cells, gamma delta T cells, resting NK cells, activated NK cells, and M2 macrophages showed a significant divergence between normal and EM tissues. Previous studies have shown that healthy women and EM patients exhibit the largest population of immune cells in their peritoneal fluid. For example, a transcriptome analysis revealed 10,280 and 7,250 single-cell transcripts in the peritoneal fluid of EM and control samples, respectively (Zou et al., 2021). The cells in the peritoneal

fluid play a role in immunity and are responsible for removing reflow menstrual debris as well as tissue defense. On the other hand, macrophages represent the largest immune population in the peritoneal fluid, followed by T and dendritic cells, which is consistent with previous studies (Guo et al., 2020). Pillarisetty et al. (2005) reported that other cell groups, including NK, mast, plasma, and epithelial cells, are an intermediate cell type, and named them NKDCs.

The macrophages, primary defense cells in the peritoneal cavity, mainly function to remove cell debris. Notably, EM cells in the peritoneal cavity are not targeted for clearance by phagocytes and NK cells; hence, they are able to escape and survive to invade the peritoneal cavity, *via* a mechanism called "immune escape"



(Vetvicka et al., 2016). Vallvé-Juanico et al. (2019) demonstrated that the number and activation of macrophages are high in EM samples, and this was attributed to phagocytosis that clears red blood cells, damaged tissue fragments, and cell debris, and also produces soluble mediators, such as cytokines, prostaglandins, complement components, and enzymes, to regulate the peritoneal environment. Secretion of these immune mediators causes macrophages to induce inflammation, tissue repair, and neovascularization, a phenomenon that may contribute to recruitment of fibroblasts and endothelial cells (Yang et al., 2017). Previous studies have shown that macrophage-derived cytokines stimulate activation of other immune cells, such as T and B lymphocytes. Notably, numerous studies have demonstrated that

macrophages are dysfunctional in EM despite an increase in their activation (Zou et al., 2021). Macrophages are functionally differentiated into M1 and M2 cell lines, in the presence of infectious microorganisms or host mediators (Thiruchelvam et al., 2012). Specifically, M1 macrophages are induced by expression of nitric oxide synthase (iNOS) and TNF to produce a large amount of nitric oxide, which is necessary for eliminating bacteria, viruses, and fungi infections. On the other hand, M2 macrophages, also called activated macrophages, respond to parasitic infection and tissue remodeling, and play an important role in angiogenesis and tumor progression. To date, the actual phenotypes of macrophages in the peritoneal fluid in EM patients remains controversial. Recent studies have shown that

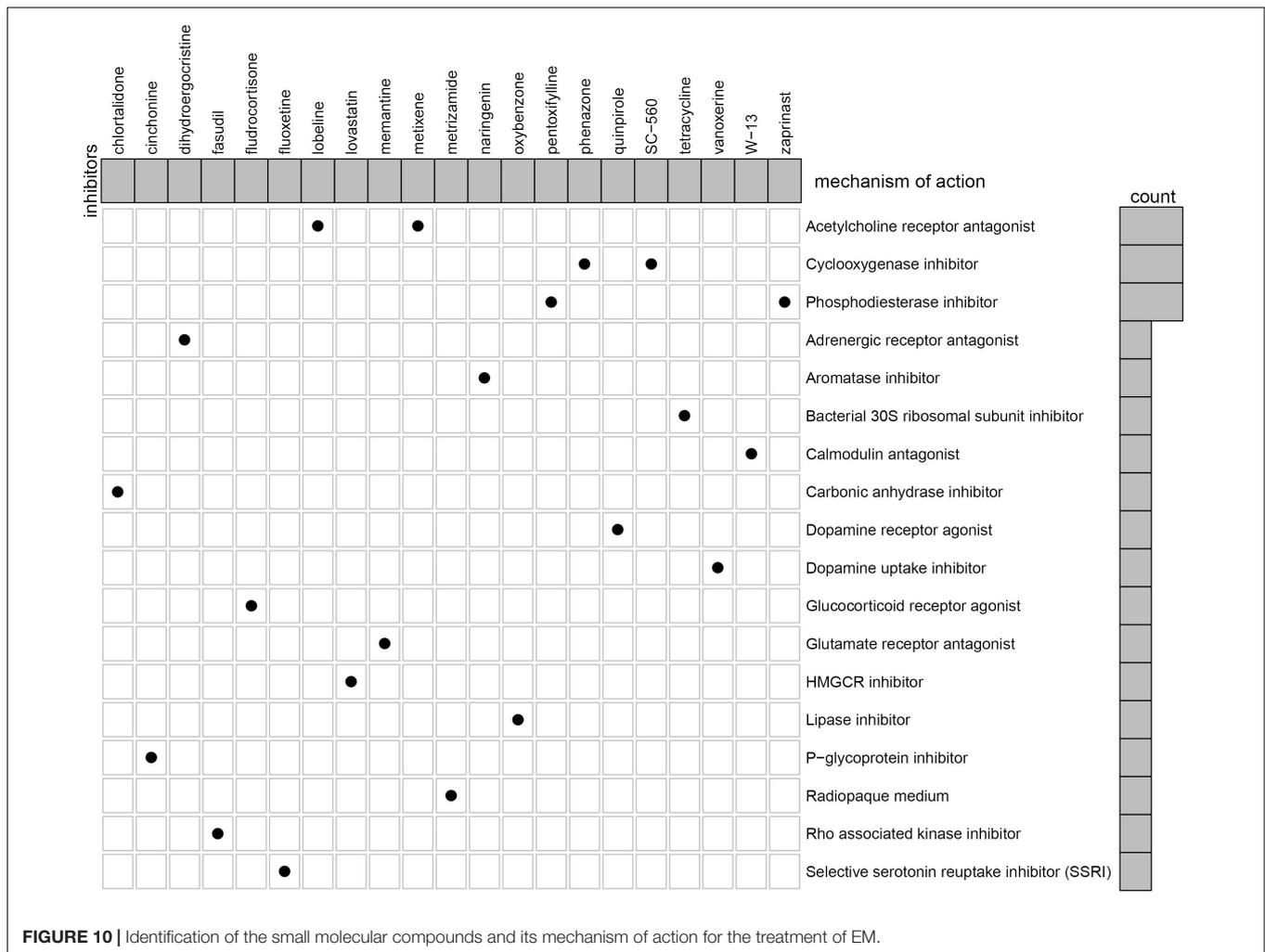


FIGURE 10 | Identification of the small molecular compounds and its mechanism of action for the treatment of EM.

EM pathogenesis is a complex and heterogeneous phenomenon (Guo et al., 2020). In our analysis, we analyzed the level of 22 immune cells' infiltration in EM using the CIBERSORTx algorithm and found that M2 macrophages were significantly elevated in EM, relative to normal tissues, indicating that the M2 macrophages might be involved in the pathogenesis of EM.

The NK cells are an important part of the innate immune system, where they play a crucial role in anti-tumor immunity. Some research evidence have shown that NK cytotoxicity increases in the peritoneal fluid of EM patients, with disease severity (Zou et al., 2021). Notably, these changes may interfere with the monitoring, recognition, and destruction of ectopic endometrium cells, thereby causing EM (Gogacz et al., 2017). In our result, we found that the level of resting NK cell and activated NK cell were significantly decreased from the healthy sample to the EM, which may increase the NK cytotoxicity level of EM.

Currently, a lack of non-invasive, clear, and consistent biomarkers for identification of EM have significantly affected accurate diagnosis of the disease. Notably, all pharmacological treatments for EM are inhibitory rather than radical, with

non-steroidal anti-inflammatory drugs (NSAIDs) used as the drug of choice for management of EM-related pain (Wattier, 2018). To date, selective progesterone receptor modulators (SPRMs) (Song et al., 2018), oral contraceptives (Grandi et al., 2019), mifepristone (RU486) (Song et al., 2018), Levonorgestrel system (LNGN-ius) (Song et al., 2018), and aromatase inhibitors (Suardika and Pelayun, 2018) have been employed in treating EM. However, their role in restoring fertility and efficacies remains unsatisfactory due to numerous side effects. Here, we investigated the potential drugs through the CMAP dataset to the treatment of EM. We analyzed the mechanism of action of each drug and revealed 21 drugs, namely, W-13, phenazone, metixene, cinchonine, tetracycline, dihydroergocristine, quinpirole, lovastatin, memantine, metrizamide, naringenin, pentoxifylline, lobeline, zaprinast, chlortalidone, SC-560, oxybenzone, fasudil, fluoxetine, fludrocortisone, and vanoxerine, with the potential for treating EM.

This study had a few limitations. Firstly, although we identified and validated potential DEGs, we did not experimentally evaluate their efficacy. There is a need for multi-center prospective

studies to evaluate their application. Secondly, we did not elucidate the underlying molecular mechanisms using *in vivo* and *in vitro* experiments. Finally, we did not obtain sufficient clinical information for the subjects herein; therefore, the clinical value of our findings needs further verification.

In conclusion, we identified three key genes between EM and healthy tissue, which may serve as potential biomarkers to the EM. We also evaluated the relationship between key genes and immune cell to explore the molecular mechanism of EM. Finally, we discovered molecular compounds that may contribute to the treatment of EM.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

REFERENCES

- Aghajanova, L., and Giudice, L. (2010). Molecular evidence for differences in endometriosis in severe versus mild endometriosis. *Reprod. Sci. (Thousand Oaks, Calif)* 18, 229–251. doi: 10.1177/1933719110386241
- Ahn, S. H., Monsanto, S., Miller, C., Singh, S., Thomas, R., and Tayade, C. (2015). Review article pathophysiology and immune dysfunction in endometriosis. *BioMed. Res. Int.* 2015:795976. doi: 10.1155/2015/795976
- Bader, G., and Hogue, C. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinform.* 4:2. doi: 10.1186/1471-2105-4-2
- Burney, R., and Giudice, L. (2012). Pathogenesis and pathophysiology of endometriosis. *Fertil. Steril.* 98, 511–519. doi: 10.1016/j.fertnstert.2012.06.029
- Chen, Y., Shao, Z., Jiang, E., Zhou, X., Wang, L., Wang, H., et al. (2020). CCL21/CCR7 interaction promotes EMT and enhances the stemness of OSCC via a JAK2/STAT3 signaling pathway. *J. Cell. Physiol.* 235, 5995–6009. doi: 10.1002/jcp.29525
- Cui, D., Liu, Y., Ma, J., Lin, K., Xu, K., and Lin, J. (2020). Identification of key genes and pathways in endometriosis by integrated expression profiles analysis. *PeerJ* 8:e10171. doi: 10.7717/peerj.10171
- Dai, F. F., Bao, A. Y., Luo, B., Zeng, Z. H., Pu, X. L., Wang, Y. Q., et al. (2019). Identification of differentially expressed genes and signaling pathways involved in endometriosis by integrated bioinformatics analysis. *Exp. Therapeutic Med.* 19, 264–272. doi: 10.3892/etm.2019.8214
- Davis, S., and Meltzer, P. (2007). GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics (Oxford, England)* 23, 1846–1847. doi: 10.1093/bioinformatics/btm254
- Erekat, N., Rababah, R., and Al-Jarrah, M. (2018). Overexpression of renal proapoptotic factors is attenuated subsequent to endurance exercise in Type I diabetes: an immunohistochemistry study. *J. Nat. Sci. Biol. Med.* 10:24. doi: 10.4103/jnsbm.JNSBM_60_18
- Fu, Q., Tan, X., Tang, H., and Liu, J. (2021). CCL21 activation of the MALAT1/SRSF1/mTOR axis underpins the development of gastric carcinoma. *J. Transl. Med.* 19:210. doi: 10.1186/s12967-021-02806-5
- Gillis, C., Aurélie, G.-C., Jönsson, F., and Bruhns, P. (2014). Contribution of human FcγRs to disease with evidence from human polymorphisms and transgenic animal studies. *Front. Immunol.* 5:254. doi: 10.3389/fimmu.2014.00254
- Gogacz, M., Gałczyński, K., Wojtaś, M., Winkler, I., Adamiak, A., Romanek-Piva, K., et al. (2017). Fas-Related apoptosis of peritoneal fluid macrophages in endometriosis patients: understanding the disease. *J. Immunol. Res.* 2017, 1–8. doi: 10.1155/2017/3175394
- Grandi, G., Barra, F., Ferrero, S., Sileo, F., Bertucci, E., Napolitano, A., et al. (2019). Hormonal contraception in women with endometriosis: a systematic review.

AUTHOR CONTRIBUTIONS

YP conceived and designed the experiments and revised and wrote the manuscript. CP carried out the investigation and collected materials. ZF analyzed the data. GC provided administrative, technical, and material support and provided a critical review of the intellectual content of the manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.622683/full#supplementary-material>

Supplementary Figure 1 | ROC analysis was performed to demonstrate the diagnostic power of the three genes in the two external validation datasets.

- Eur. J. Contraception Reprod. Health Care* 24, 1–10. doi: 10.1080/13625187.2018.1550576
- Guo, M., Bafligil, C., Tapmeier, T., Hubbard, C., Manek, S., Shang, C., et al. (2020). Mass cytometry analysis reveals a distinct immune environment in peritoneal fluid in endometriosis: a characterisation study. *BMC Med.* 18:3. doi: 10.1186/s12916-019-1470-y
- Guo, S.-W., Zilberberg, M., and Hummelshoj, L. (2012). Endometriosis and ovarian cancer. *Lancet Oncol.* 13, e189–e190. doi: 10.1016/S1470-2045(12)70199-7 author reply e90,
- Hawkins, S., Creighton, C., Han, D., Zariff, A., Anderson, M., Gunaratne, P., et al. (2011). Functional MicroRNA involved in endometriosis. *Mol. Endocrinol. (Baltimore, Md)* 25, 821–832. doi: 10.1210/me.2010-0371
- Herington, J., Bruner-Tran, K., Lucas, J., and Osteen, K. (2011). Immune interactions in endometriosis. *Expert Rev. Clin. Immunol.* 7, 611–626. doi: 10.1586/eci.11.53
- Hever, A., Roth, R., Hevezi, P., Marin, M., Acosta, J., Acosta, H., et al. (2007). Human endometriosis is associated with plasma cells and overexpression of B lymphocyte stimulator. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12451–12456. doi: 10.1073/pnas.0703451104
- Heydari, S., Kashani, L., and Noruzinia, M. (2020). Gene expression analysis signifies the association of inflammatory proteins with the development of endometriosis. *Gene Rep.* 20:100716. doi: 10.1016/j.genrep.2020.100716
- Hill, C., Fakhreldin, M., Maclean, A., Dobson, L., Nancarrow, L., Bradfield, A., et al. (2020). Endometriosis and the fallopian tubes: theories of origin and clinical implications. *J. Clin. Med.* 9:1905. doi: 10.3390/jcm9061905
- Ho, L., van Dijk, M., Sam, T., Messerschmidt, D., Chng, S., Ong, S., et al. (2017). ELABELA deficiency promotes preeclampsia and cardiovascular malformations in mice. *Science* 357, 707–713. doi: 10.1126/science.aam6607
- Holtan, S., Creedon, D., Haluska, P., and Markovic, S. (2009). Cancer and pregnancy: parallels in growth, invasion, and immune modulation and implications for cancer therapeutic agents. *Mayo Clin. Proc. Mayo Clin.* 84, 985–1000. doi: 10.1016/S0025-6196(11)60669-1
- Hull, M. L., Rangel, C., Godsland, J., Doig, J., Johnson, C., Phillips, S., et al. (2008). Endometrial-peritoneal interactions during endometriotic lesion establishment. *Am. J. Pathol.* 173, 700–715. doi: 10.2353/ajpath.2008.071128
- Jiménez Morales, A., Maldonado-Montoro, M., Martínez de la Plata, J. E., Pérez Ramírez, C., Daddaoua, A., Alarcón Payer, C., et al. (2019). Fcgr2a/fcgr3a gene polymorphisms and clinical variables as predictors of response to tocilizumab and rituximab in patients with rheumatoid arthritis. *J. Clin. Pharmacol.* 59, 517–531. doi: 10.1002/jcph.1341
- Kidoya, H., Ueno, M., Yamada, Y., Mochizuki, N., Nakata, M., Fujii, R., et al. (2008). Spatial and temporal role of the apelin/APJ system in the caliber size

- regulation of blood vessels during angiogenesis. *EMBO J.* 27, 522–534. doi: 10.1038/sj.emboj.7601982
- Kruitwagen, R., Poels, L., Willemsen, W., Ronde, I. J. Y., Jap, P., and Rolland, R. (1991). Endometrial epithelial cells in peritoneal fluid during the early follicular phase. *Fertil. Steril.* 55, 297–303.
- Kuo, H.-C., Chang, J.-C., Kuo, H.-C., Yu, H.-R., Wang, C.-L., Lee, C.-P., et al. (2015). Identification of an association between genomic hypomethylation of FCGR2A and susceptibility to Kawasaki disease and intravenous immunoglobulin resistance by DNA methylation array: FCGR2A methylation in Kawasaki disease. *Arthritis Rheumatol. (Hoboken, NJ)* 67, 828–836. doi: 10.1002/art.38976
- Lee, Y., Choi, S., Ji, J., and Song, G. (2016). Associations between functional FCGR2A R131H and FCGR3A F158V polymorphisms and responsiveness to TNF blockers in spondyloarthritis, psoriasis and Crohn's disease: a meta-analysis. *Pharmacogenomics* 17, 1465–1477. doi: 10.2217/pgs.16.27
- Li, D., Yang, Z., Liu, Z., Zou, Q., and Yuan, Y. (2020). Clinical significance of CBS and CCL21 in gallbladder adenocarcinomas and squamous cell/adenosquamous carcinomas. *Appl. Immunohistochem. Mol. Morphol.* 28, 103–110. doi: 10.1097/PAI.0000000000000705
- Lin, Y., Sharma, S., and John, M. (2014). CCL21 cancer immunotherapy. *Cancers* 6, 1098–1110. doi: 10.3390/cancers6021098
- Liu, F., Lv, X., Yu, H., Xu, P., Ma, R., and Zou, K. (2015). In search of key genes associated with endometriosis using bioinformatics approach. *Eur. J. Obstetrics Gynecol. Reprod. Biol.* 194, 119–124. doi: 10.1016/j.ejogrb.2015.08.028
- Lotfialiansaremi, S., Sabio, M., Comwell, S., and Tolia, P. (2020). The role of the mitogen-activated protein kinase (MAPK) signaling pathway in cancer. *Med. Res. Arch.* 8. doi: 10.18103/mra.v8i4.2086
- Mestiri, S., Zaaber, I., Nasr, I., and Marmouch, H. (2020). Implication of VDR Rs7975232 and FCGR2A Rs1801274 gene polymorphisms in the risk and the prognosis of autoimmune thyroid diseases in the tunisian population. *Balkan J. Med. Genet.* 23, 69–76. doi: 10.2478/bjmg-2020-0011
- Musa, A., Soltan Ghoraba, L., Zhang, S.-D., Glazko, G., Yli-Harja, O., Dehmer, M., et al. (2017). A review of connectivity map and computational approaches in pharmacogenomics. *Briefings Bioinform.* 18:903. doi: 10.1093/bib/bbx023
- Newman, A., Steen, C., Liu, C., Gentles, A., Chaudhuri, A., Scherer, F., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* 37, 773–782. doi: 10.1038/s41587-019-0114-2
- Pillarisetty, V., Katz, S., Bleier, J., Shah, A., and Dematteo, R. (2005). Natural killer dendritic cells have both antigen presenting and lytic function and in response to CpG produce IFN- γ via autocrine IL-12. *J. Immunol. (Baltimore, Md : 1950)* 174, 2612–2618. doi: 10.4049/jimmunol.174.5.2612
- Ritchie, M., Phipson, B., Wu, D., Hu, Y., Law, C., Shi, W., et al. (2015). LIMMA powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007
- Rossi, G. M., Bonatti, F., Adorni, A., Alberici, F., Bodria, M., Bonanni, A., et al. (2017). FCGR2A single nucleotide polymorphism confers susceptibility to childhood-onset idiopathic nephrotic syndrome. *Immunol. Lett.* 193, 11–13. doi: 10.1016/j.imlet.2017.11.002
- Sanz, H., Valim, C., Vegas, E., Oller, J., and Reverter, F. (2018). SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinform.* 19:432. doi: 10.1186/s12859-018-2451-4
- Shannon, P., Markiel, A., Ozier, O., Baliga, N., Wang, J., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Simoens, S., Dunselman, G., Dirksen, C., Hummelshoj, L., Bokor, A., Brandes, I., et al. (2012). The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Hum. Reprod. (Oxford, England)* 27, 1292–1299. doi: 10.1093/humrep/des073
- Song, J., Wang, Y., and Yu, L. (2018). Clinical comparison of mifepristone and gestrinone for laparoscopic endometriosis. *Pak. J. Pharm. Sci.* 31(5(Special)), 2197–2201.
- Suardika, A., and Pemayun, T. G. A. (2018). New insights on pathogenesis and novel non-surgical therapies for endometriosis. *J. Turkish-German Gynecol. Assoc.* 19, 158–164. doi: 10.4274/jtgga.2018.0090
- Thiruchelvam, U., Dransfield, I., Saunders, P., and Critchley, H. (2012). The importance of the macrophage within the human endometrium. *J. Leukocyte Biol.* 93, 217–225. doi: 10.1189/jlb.0712327
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc.: Series B (Methodological)* 58, 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x
- Vallvé-Juanico, J., Santamaria, X., Vo, K., Houshdaran, S., and Giudice, L. (2019). Macrophages display proinflammatory phenotypes in the eutopic endometrium of women with endometriosis with relevance to an infectious etiology of the disease. *Fertil. Steril.* 112, 1118–1128. doi: 10.1016/j.fertnstert.2019.08.060
- Velissari, A., Skalidakis, I., Oliveira, S., Koutsandrea, C., Kitsos, G., Petersen, M., et al. (2015). Novel association of FCGR2A polymorphism with age-related macular degeneration (AMD) and development of a novel CFH real-time genotyping method. *Clin. Chem. Lab. Med. : CCLM / FESCC* 53, 1521–1529. doi: 10.1515/cclm-2014-0920
- Vercellini, P., Viganò, P., Somigliana, E., and Fedele, L. (2013). Endometriosis: pathogenesis and treatment. *Nat. Rev. Endocrinol.* 10, 261–275. doi: 10.1038/nrendo.2013.255
- Vetvicka, V., Laganà, A. S., Salmeri, F., Triolo, O., Palmara, V., Vitale, S. G., et al. (2016). Regulation of apoptotic pathways during endometriosis: from the molecular basis to the future perspectives. *Arch. Gynecol. Obstetrics* 294, 897–904. doi: 10.1007/s00404-016-4195-6
- Wattier, J.-M. (2018). Antalgiques et alternatives thérapeutiques non médicamenteuses pluridisciplinaires, RPC Endométriose CNGOF-HAS. *Gynecol. Obstétrique Fertil. Senol.* 46, 248–255. doi: 10.1016/j.gofs.2018.02.002
- Wu, B., Yang, Z., Tobe, R., and Wang, Y. (2017). Medical therapy for preventing recurrent endometriosis after conservative surgery: a cost-effectiveness analysis. *BJOG: Int. J. Obstetrics Gynaecol.* 125, 469–477. doi: 10.1111/1471-0528.14786
- Xu, Y., and Shen, K. (2018). Identification of potential key genes associated with ovarian clear cell carcinoma. *Cancer Manag. Res.* 10, 5461–5470. doi: 10.2147/CMAR.S187156
- Yamada, A., Arakaki, R., Saito, M., Kudo, Y., and Ishimaru, N. (2017). Dual role of Fas/FasL-mediated signal in peripheral immune tolerance. *Front. Immunol.* 8:403. doi: 10.3389/fimmu.2017.00403
- Yang, H.-L., Zhou, W.-J., Kai, C., Mei, J., Huang, L.-Q., Wang, M.-Y., et al. (2017). The crosstalk between ESC and M ϕ impairs cytotoxicity of NK cell by secreting IL-10 and TGF- β in EMS. *Reproduction* 154:RE-17. doi: 10.1530/REP-17-0342
- Yu, G. (2018). ClusterProfiler: an universal enrichment tool for functional and comparative study. *bioRxiv* [preprint]. doi: 10.1101/256784
- Zhao, L., Gu, C., Ye, M., Zhang, Z., Li, L. A., Fan, W., et al. (2018). Integration analysis of microRNA and mRNA paired expression profiling identifies deregulated microRNA-transcription factor-gene regulatory networks in ovarian endometriosis. *Reprod. Biol. Endocrinol.* 16:4. doi: 10.1186/s12958-017-0319-5
- Zou, G., Wang, J., Xu, X., Xu, P., Zhu, L., Yu, Q., et al. (2021). Cell subtypes and immune dysfunction in peritoneal fluid of endometriosis revealed by single-cell RNA-sequencing. *Cell Biosci.* 11:98. doi: 10.1186/s13578-021-00613-5

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.

Copyright © 2021 Peng, Peng, Fang and Chen. 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.