Significance of hub genes and immune cell infiltration identified by bioinformatics analysis in pelvic organ prolapse

Ying Zhao [1], **Zhijun Xia** [Corresp., 1], **Te Lin** [1], **Yitong Yin** [1]

[1] Department of Obstetrics and Gynecology, Pelvic Floor Disease Diagnosis and Treatment Center, Shengjing Hospital of China Medical University, Shenyang, China

Corresponding Author: Zhijun Xia
Email address: xzj18270916865@163.com

Objective Pelvic organ prolapse (POP) refers to the decline of pelvic organ position and dysfunction caused by weak pelvic floor support. The aim of the present study was to screen the hub genes and immune cell infiltration related to POP disease. **Methods** Microarray data of 34 POP tissues in the GSE12852 gene expression dataset were used as research objects. Weighted gene co-expression network analysis (WGCNA) was performed to elucidate the hub module and hub genes related to POP occurrence. Gene function annotation was performed using the DAVID tool. Differential analysis based on the GSE12852 dataset was carried out to explore the expression of the selected hub genes in POP and non-POP tissues, and RT-qPCR was used to validate the results. The differential immune cell infiltration between POP and non-POP tissues was investigated using the CIBERSORT algorithm. **Results** WGCNA revealed the module that possessed the highest correlation with POP occurrence. Functional annotation indicated that the genes in this module were mainly involved in immunity. *ZNF331*, *THBS1*, *IFRD1*, *FLJ20533*, *CXCR4*, *GEM*, *SOD2*, and *SAT* were identified as the hub genes. Differential analysis and RT-qPCR demonstrated that the selected hub genes were overexpressed in POP tissues as compared with non-POP tissues. The CIBERSORT algorithm was employed to evaluate the infiltration of 22 immune cell types in POP tissues and non-POP tissues. We found greater infiltration of activated mast cells and neutrophils in POP tissues than non-POP tissues, while the infiltration of resting mast cells was lower in POP tissues. Moreover, we investigated the relationship between the type of immune cell infiltration and hub genes by Pearson correlation analysis. The results indicate that activated mast cells and neutrophils had a positive correlation with the hub genes, while resting mast cells had a negative correlation with the hub genes. **Conclusions** Our research identified 8 hub genes and the infiltration of 3 types of immune cells related to POP occurrence. These hub genes may participate in the pathogenesis of POP through the immune system, giving them a certain diagnostic and therapeutic value.

1 Significance of hub genes and immune cell infiltration identified by bioinformatics
2 analysis in pelvic organ prolapse

3 Ying Zhao[1], Zhijun Xia*[1], Te Lin[1], Yitong Yin[1]

4 [1]Department of Obstetrics and Gynecology, Pelvic Floor Disease Diagnosis and Treatment Center,

5 Shengjing Hospital of China Medical University, Shenyang, China, 110004.

6 ***Correspondence:**

7 Corresponding Author: Zhijun Xia, Department of Obstetrics and Gynecology, Shengjing Hospital

8 of China Medical University, Shenyang, China, 110004

9 Email: xzj18270916865@163.com, xiazj@sj-hospital.org

10 **Acknowledgments**

11 We acknowledge the authors who provided information for the GEO public database.

12 **Funding statement:** This article was funded by Shengjing Hospital of China Medical University.

13 **Abstract**

14 **Objective** Pelvic organ prolapse (POP) refers to the decline of pelvic organ position and

15 dysfunction caused by weak pelvic floor support. The aim of the present study was to screen the

16 hub genes and immune cell infiltration related to POP disease.

17 **Methods** Microarray data of 34 POP tissues in the GSE12852 gene expression dataset were used

18 as research objects. Weighted gene co-expression network analysis (WGCNA) was performed to

19 elucidate the hub module and hub genes related to POP occurrence. Gene function annotation was

20 performed using the DAVID tool. Differential analysis based on the GSE12852 dataset was carried

21 out to explore the expression of the selected hub genes in POP and non-POP tissues, and RT-qPCR

22 was used to validate the results. The differential immune cell infiltration between POP and non-

23 POP tissues was investigated using the CIBERSORT algorithm.

24 **Results** WGCNA revealed the module that possessed the highest correlation with POP occurrence.

25 Functional annotation indicated that the genes in this module were mainly involved in immunity.

26 *ZNF331*, *THBS1*, *IFRD1*, *FLJ20533*, *CXCR4*, *GEM*, *SOD2*, and *SAT* were identified as the hub

27   genes. Differential analysis and RT-qPCR demonstrated that the selected hub genes were

28   overexpressed in POP tissues as compared with non-POP tissues. The CIBERSORT algorithm was

29   employed to evaluate the infiltration of 22 immune cell types in POP tissues and non-POP tissues.

30   We found greater infiltration of activated mast cells and neutrophils in POP tissues than non-POP

31   tissues, while the infiltration of resting mast cells was lower in POP tissues. Moreover, we

32   investigated the relationship between the type of immune cell infiltration and hub genes by Pearson

33   correlation analysis. The results indicate that activated mast cells and neutrophils had a positive

34   correlation with the hub genes, while resting mast cells had a negative correlation with the hub

35   genes.

36   **Conclusions** Our research identified 8 hub genes and the infiltration of 3 types of immune cells

37   related to POP occurrence. These hub genes may participate in the pathogenesis of POP through

38   the immune system, giving them a certain diagnostic and therapeutic value.

39   **Keywords:** Pelvic organ prolapse, WGCNA, CIBERSORT, immune cell infiltration, GEO

40   database

41

1 Introduction

43   Pelvic organ prolapse (POP) is caused by dysfunction of the pelvic floor supporting structures

44   and affects the quality of life of many women (Nygaard et al. 2014). The prevalence of POP is

45   expected to reach 46% by 2050 (Wu et al. 2009). Currently, only approximately 20% of women

46   have POP surgery during their lifetime (Haya et al. 2018); however, the reoperation rate is high

47   (Friedman et al. 2018), which causes a huge economic burden. POP is a complex, multifaceted

48   disease resulting from the interaction between environmental and genetic factors. Pregnancy,

49   vaginal pull-up, time of delivery, age, and obesity have been identified as risk factors (Hendrix et

50   al. 2002); nevertheless, the molecular mechanism of POP remains unclear, and there is a lack of

51   suitable prevention and treatment measures in clinical practice. Therefore, it is imperative to

52   explore the molecular mechanism of POP occurrence for the benefit for both women and society.

53   Owing to the development of gene chips and new generation sequencing technology,

54   bioinformatics analysis plays an increasingly important role in biomedical research. Weighted

55   gene co-expression network analysis (WGCNA) can identify gene module characteristics and hub

56   genes to connect the gene modules and sample characteristics (Wang et al. 2020b). The

57   CIBERSORT algorithm can be employed to evaluate immune cell infiltration in tissues based on

58   gene expression datasets. Recently,abundant researches have used this algorithm to explore the

59   function of immune cells in diseases(Lin et al. 2020), breast ductal and lobular carcinoma (Zhang

60   et al. 2019), osteoarthritis (Cai et al. 2020), and high-grade serous ovarian cancer (Liu et al. 2020).

61   Our study found hub genes and immune cells highly related to POP occurrence by analyzing POP

62   expression spectrum data in public databases using WGCNA and the CIBERSORT algorithm,

63   providing novel ideas and methods for the treatment of POP. Finally, the differential immune cells

64   was investigated between POP and non-POP tissues.

65   **2   Materials and Methods**

66   **Obtaining the training and validation POP datasets**

67   The workflow diagram is presented in Figure 1. The training dataset was obtained from the Gene

68   Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/) in NCBI based on the

69   keywords "pelvic organ prolapse" and "gene expression profiles" and "*Homo sapiens*" (Brizzolara

70   et al. 2009). In addition, we selected a dataset with a sample size greater than 30, which is the

71   minimum sample size required to construct a WGCNA network (POP > 15, non-POP > 15); the

72   GSE12852 dataset, containing 16 POP and 18 non-POP patients, was the only dataset to meet this

73   condition.   The gene expression profiles and corresponding clinical information including age,

74   menopausal status, race, and prolapse stage were acquired from the GSE12852 dataset.

75   Subsequently, log-scale robust multi-array analysis was used to perform background correction

76   and normalization of the datasets. A total of 12 POP patients enrolled at the Shengjing Hospital of

77   China Medical University from 2017 to 2018 were selected as the validation dataset. The Ethics

78   Committee of Shengjing Hospital of China Medical University approved the study protocol in

79   accordance with the guidelines of the Declaration of Helsinki (No. 2018Ps68K), and our study was

80   exempt from consent. Prolapse tissues were selected as the experimental group and non-prolapse

81   tissues were selected as the control group. All patients were married, without estrogen-related

82   disorders, and had received no hormone treatment within the previous 3 months. Basic information

83   of the patients in the validation dataset is presented in Supplementary Table 1.

84   **Co-expression network construction based on the training dataset**

85   We used the "WGCNA" package of the R software to establish the network based on the 34

86   samples in the training dataset according to a previously described method (Chen et al. 2018b;

87   Zhou et al. 2018). Firstly, POP and normal samples were analyzed using the "flash cluster"

88   package of the R software to remove outlier samples. The correlation matrix of gene co-expression

89   consists of the correlation coefficient between two genes. The average linkage matrix and Pearson

90   correlation method were performed to construct the correlation matrix between genes.

91   Subsequently, the formula $a_{mn} = |c_{mn}|^{\beta}$ was used to convert the correlation matrix to the

92   adjacency matrix (a_{mn} represents the correlation coefficient between gene m and gene n; c_{mn}

93   represents the connection coefficient between gene m and gene n; and β is a soft threshold that

94   can strengthen strong links between genes and weaken weak links). Finally, we converted the

95   adjacency matrix to the topological overlap matrix (TOM) based on an appropriate soft threshold,

96   and placed the similar genes into the same module.

97   **Identification of hub gene modules and genes**

98   Similar genes were placed into the same module based on the WGCNA co-expression network.

99   The main purpose of our research was to combine clinical information (POP or non-POP samples)

100   with gene modules to analyze gene significance (GS) and modular membership (MM). MM means

101   the correlation between gene expression profile level and module eigengenes (ME). ME are

102   considered the correlation between modules and clinical information. GS represents the degree of

103   correlation between gene expression profiles and clinical information. The average value of all

104   gene GS in the module represents the module significance (MS). We defined the correlation

105   between gene and disease as GS, and obtained the correlation between this module and disease as

106   MS. The hub genes characterized by a high MM and a high GS are described as having the closest

107   relationship with disease. In our research, the module with the highest MS was selected as the hub

108   module. The genes in the hub module with | MM | > 0.8 and | GS | > 0.47 were considered hub

109   genes.

110   **Functional enrichment analysis**

111   We used the online GO enrichment analysis and KEGG pathway analysis tools on the DAVID

112   website to annotate the genes in the module identified by WGCNA, and attempted to elucidate the

113   enrichment pathway and functions of the target genes. The   Human Genome U133 Plus 2 Array

114   was used as the background data (The FDR < 0.01).

115   **Immune cell infiltration in POP tissues**

116   CIBERSORT is an algorithm to characterize the proportion of 22 immune cells (Supplementary

117   Table 2) in tissues using 547 barcode gene expression values. The CIBERSORT algorithm was

118   employed to elucidate the proportion of 22 immune cells in POP tissues. The samples with p value

119   < 0.05 were significant (Chen et al. 2018a). Pearson correlation analysis was implemented to

120   obtain the related coefficient between the 22 immune cells. Then, we investigated the differential

121   immune cell infiltration between POP and non-POP tissues. Finally, we calculated the related

122   coefficient between infiltration of different immune cells and hub genes by Pearson correlation

123   analysis.

124   **RNA extraction and quantitative real-time PCR based on the validation dataset**

125   TRIzol® (1 mL) was used to isolate total RNA from POP tissues (200 mg) in the validation dataset,

126   and reverse transcriptase from the avian myeloblastoma virus and random primers were used to

127   create complementary DNA (cDNA) according to the instructions from TAKARA. SYBR Premix

128   Ex Taq II (Takara)   was used to amplify the cDNA. According to the samples from three

129 independent experiments, the $2^{-\triangle\triangle CT}$ value was used to analyze the data. Primers for the genes are

130 displayed in Supplementary Table 3.

Statistical analysis

132 The GraphPad Prism 7.0 and R 3.6.1 software were used for statistical analysis and image

133 generation. A t-test was employed to analyze the differences between two groups. The $p < 0.05$

134 was statistically significant.

135 **3 Results**

136 **Construction of WGCNA and identification of hub modules**

137 The "flash cluster" software package in R was used to cluster the 34 samples in the GSE12852

138 dataset to remove the outlier samples; 4 outlier samples were removed (Figure 2). There were few

139 differences among the remaining samples, which is conducive to the accuracy of the results. The

140 correlation matrix between genes was calculated according to the average linkage matrix and

141 Pearson correlation method. The formula amn = | cmn | $^\beta$ (β=7)was used to transform the

142 correlation matrix into the adjacency matrix. In order to better construct a scale-free network

143 distribution, the "picks of threshold" function of "WGCNA" package calculated the value of

144 parameter β. In POP and non-POP samples, 1−20 thresholds were selected and the correlation

145 coefficient, mean connectivity, and average correlation degree between log (k) and log (P (k)) were

146 calculated for each threshold. At this time, the average network connectivity corresponding to the

147 threshold is close to zero, indicating that the network connectivity is very low, which is similar to

148 the scale-free network (Figure 3). According to the corresponding steps of WGCNA, the gene

149 network was built after the hierarchical clustering tree. In this experiment, the dynamic pruning

150 tree method merged with similar genes into one gene module. The minimum number of genes in

151 a module was 30, and 11 modules were obtained (Figure 4A). According to the thermogram of

152 correlation between modules and clinical information, the highest correlation existed in the blue

153 module and sample type, with a correlation coefficient of 0.47 (Figure 4B, $p = 0.009$). Therefore,

154 the blue module was considered the most study-worthy. Protein−protein interactions (PPIs) are

155   widely involved in the process of vital movement. http://omnipathdb.org/ was used to construct

156   the PPI network for genes in the blue module (Figure 5)(Türei et al. 2016).

157   **Functional enrichment analysis**

158   To further study the biological function of the genes in the blue module, GO and KEGG

159   enrichment analyses were used. GO enrichment analysis showed that these genes mainly

160   participated in the immune regulation (Figure 6A, Supplementary Table 4). KEGG enrichment

161   analysis identified that these genes participated in the regulation of immune-related pathways, such

162   as the IL-17 and TNF signaling pathways (Figure 6B, Supplementary Table 5).

163   **Identification of the hub genes**

164   According to the screening criteria | MM | > 0.8 and | GS | > 0.47, 8 genes (*ZNF331*, *THBS1*,

165   *TMEM70*, *CXCR4*, *GEM*, *SOD2*, and *SAT*) were identified in the blue module as hub genes (Figure

166   7A). We found that the hub genes were highly expressed in POP tissues as compared with non-

167   POP tissues by differential analysis based on the GSE12852 dataset (Figure 7B). From the

168   heatmap, it can be seen that hub genes are overexpressed in POP tissues (Figure 7C). In addition,

169   the correlation coefficient between the hub genes was calculated, which showed that there was a

170   strong co-expression relationship between these genes (Figure 7D).

171   **Validation of the hub genes**

172   To verify the accuracy of the prediction results, RT-qPCR was used to detect the expression of the

173   hub genes in 12 pairs of POP and non-POP tissues. The results show that the hub genes were

174   overexpressed in POP tissues, which is consistent with the prediction results (Figure 8A−H).

175   **Immune cell infiltration analysis**

176   The CIBERSORT algorithm was employed to select samples with a output $p < 0.05$.   12 samples

177   including 4 non-POP and 8 POP tissues were obtained. A bar plot was generated to show the

178   proportion of 22 immune cells in the 12 samples (Figure 9A). We found that macrophages account

179   for the largest proportion among the immune cells in the samples. Figure 9B indicates that M1

180   macrophages had the strongest positive correlation with resting mast cells (correlation coefficient,

181   0.58), whereas resting mast cells had the strongest negative correlation with activated mast cells

182   (correlation coefficient, 0.83). We found a higher infiltration of activated mast cells and

183   neutrophils in POP tissues than in non-POP tissues, while the infiltration of resting mast cells was

184   lower in POP tissues (Figure 10, $p < 0.05$). Finally, Pearson correlation analysis was used to

185   calculate the related coefficient between the infiltration of different immune cells and hub genes.

186   The results reveal that activated mast cells and neutrophils had a positive correlation with the hub

187   genes, while resting mast cells had a negative correlation with the hub genes (Table 1).

188   **4   Discussion**

189   Pelvic organ prolapse (POP) is caused by weakening of the supporting structures of the pelvic

190   floor, resulting in the position of pelvic floor organs moving downward and the function becoming

191   abnormal. POP seriously affects the quality of life of middle-aged and elderly women. In recent

192   years, with the rapid development of gene sequencing and bioinformatics technologies, further

193   analysis and utilization of sequencing data have become possible. To explore the molecular

194   mechanism related to the development of POP, 8 hub genes (*ZNF331*, *THBS1*, *IFRD1*, *FLJ20533*,

195   *CXCR4*, *GEM*, *SOD2*, and SAT) related to POP occurrence were screened out based on WGCNA

196   and their biological functions were explored.

197   Thrombospondins (THBSs) are a group of glycoproteins that bind to collagen and tissue,

198   participating in the interaction between cells and the extracellular matrix during the process of

199   tissue development and repair. *THBS1* is the first member of the *THBS* gene family, which plays

200   a significant role in many biological processes related to the occurrence and progression of

201   cardiovascular diseases, such as angiogenesis, inflammation, and tissue remodeling (Zhao et al.

202   2018). In addition, *THBS1* can also affect tumor cell adhesion, invasion, migration, proliferation,

203   apoptosis, and immune evasion (Huang et al. 2017). The overexpression of *THBS1* in POP tissues

204   as compared with non-POP tissues has been predicted and verified (Brizzolara et al. 2009), which

205   is consistent with the results of our secondary analysis. *ZNF331* is located on chromosome 19q13,

206   a recently cloned gene encoding a zinc-finger protein involved in thyroid tumorigenesis (Babinger

207   et al. 2007). *ZNF331*, as a tumor suppressor gene, has also been reported to have low expression

208   in colorectal cancer (Wang et al. 2017), esophageal cancer (Jiang et al. 2015), gastric cancer (Yu

209   et al. 2013), and liver cancer (Wang et al. 2013), and its low expression is related to

210   hypermethylation of its promoter region. However, the research on *ZNF331* in other diseases was

211   rare. *Interferon-related developmental regulator 1* (*IFRD1*) is located on chromosome 7q22-q31

212   and acts a significant role in the development and differentiation of embryonic muscle cells (Kraus

213   et al. 2001; Lincoln et al. 2004). *IFRD1* has been proved to be a modified gene of cystic fibrosis

214   lung disease, which can regulate the effector function of neutrophil(Gu et al. 2009a).

215   Transmembrane protein 70 (TMEM70), also named FLJ20533, is a mitochondrial membrane

216   protein that acts a role in the biosynthesis of mitochondrial ATPase (Hejzlarova et al. 2011). At

217   present, research related to TMEM70 mainly focuses on cardiomyopathy and pulmonary

218   hypertension. CXCR4, also known as CD184, is a highly conserved receptor for the chemokine

219   CXCL12 (Kashyap et al. 2017). CXCR4 belongs to the G protein-coupled receptor superfamily

220   and is expressed in a cortical protein-dependent manner on the cell surface (Teicher & Fricker

221   2010). CXCR4 has been reported to be overexpressed in a variety of tumor cells and is involved

222   in tumor proliferation, invasion, metastasis, and worse prognosis (Ottaiano et al. 2020; Wang et

223   al. 2020a; Zhu et al. 2020). GTP-binding protein overexpressed in skeletal muscle (GEM) is a

224   GTPase originally identified in mitogen-stimulated T lymphocytes and v-Abl-transformed pre-B

225   cells and is highly expressed in the spleen, thymus, and kidneys (Cohen et al. 1994; Huang et al.

226   2014). Mitochondrial superoxide dismutase (SOD2) is an antioxidant enzyme that reduces the

227   damage caused by oxidative stress to protect mitochondria (Koltai et al. 2018).

228   Spermidine/spermine N1-acetyltransferase (SAT) is the rate-limiting enzyme in polyamine

229   catabolism and functions through the acetylation of spermidine and spermidine to affect cell

230   growth, proliferation, and apoptosis (Pegg 2008; Pegg 2016). Despite a detailed literature review,

231   we found no reports of the involvement *ZNF331*, *IFRD1*, *FLJ20533*, *CXCR4*, *GEM*, *SOD2*, or

232 *SAT* in POP. To explore the expression of the selected hub genes in POP and non-POP tissues,

233 differential analysis of the GSE12852 dataset was performed by RT-qPCR. The results indicate

234 that the selected hub genes were overexpressed in POP tissues as compared with non-POP tissues,

235 suggesting that the selected hub genes may be related to POP occurrence.

236   GO and KEGG functional enrichment analyses were performed to study the function of the

237 hub genes using the "R" software. GO functional annotation showed that the hub genes were

238 mainly participated in the immune response. KEGG enrichment analysis revealed that the hub

239 genes participated in the regulation of immune-related pathways such as the IL-17 signaling

240 pathway. IL-17 is a member of the inflammatory cytokine family and is mainly produced by Th17

241 cells. IL-17 signaling has been reported to be associated with immunopathology and autoimmune

242 diseases (Amatya et al. 2017). Previous research has confirmed that the extracellular matrix

243 components in connective tissue can control the physical strength of the pelvic floor (Dietz et al.

244 2002). The immune system can maintain the homeostasis of POP tissues by adjusting extracellular

245 matrix components (Yu et al. 2010); thus, it is reasonable to suggest that regulation of the immune

246 system is closely related to the occurrence of POP. At present, there is no direct evidence indicating

247 that IL-17 is related to POP occurrence; however, we have reason to believe that the IL-17

248 signaling pathway plays a significant role in the development of POP given its role in immune

249 regulation. In view of the close relationship between hub genes and immunity, we investigated

250 immune cell infiltration in POP and non-POP tissues using the CIBERSORT algorithm. We found

251 a higher infiltration of activated mast cells and neutrophils in POP tissues than in non-POP tissues,

252 while infiltration of resting mast cells was lower in POP tissues. Activated mast cells and

253 neutrophils had a positive correlation with the hub genes, while resting mast cells had a negative

254 correlation. According to the results, it is reasonable to suggest that the hub genes may be related

255 to the development of POP by regulating the levels of activated/resting mast cells and neutrophils.

256 Mast cells are important antigen-presenting cells that can release histamine and cytokines through

257 degranulation and act a significant role in the occurrence and development of various inflammatory

258 diseases (Grabauskas et al. 2020; Novruzov 2008; Sajay-Asbaghi et al. 2020). Very recently,

259 THBS1 was demonstrated to promote the inflammatory response of mast cells in CIU (chronic

260 idiopathic urticaria) and the permeability of HDMECs (human dermal microvascular endothelial

261 cells) by regulating the TGF-β/SMAD pathway, the effects of which can be inhibited by miR-194

262 (Qu et al. 2020). Several studies have shown that CXCR4 can promote mast cell chemotaxis to

263 inflammatory sites (Limón-Flores et al. 2009; Lv et al. 2019; Patadia et al. 2010), and IFRD1 may

264 involve in neutrophilic inflammation in cystic fibrosis (Blanchard et al. 2011; Gu et al. 2009b;

265 Hector et al. 2013). According to the current literature, there is no direct evidence to prove the

266 accuracy of our prediction results; however, the relationship between the hub genes and infiltration

267 of different immune cells suggests they are correct.

268   Some limitations of this research need to be discussed. Firstly, the research remained at the

269 prediction stage and there is insufficient experimental evidence to verify our prediction results. In

270 the future, verification *in vitro* and *in vivo* should strengthen our observations. Secondly, the

271 sample size for this study was not large enough due to the limited datasets in the database, which

272 may have caused some bias. Thirdly, the samples of our research were coming from the

273 ligamentum. The ligamentum are full of fibroblasts and it is not sure that they capture many white

274 blood cells. We could not guarantee the accuracy of the CIBERSORT algorithms completely due

275 to the blood specificity of the algorithms.

276 **Conclusions**

277 Our research identified 8 hub genes and 3 immune cell types that may be related to POP

278 occurrence. These hub genes may participate in the pathogenesis of POP by regulating the immune

279 environment, giving them certain diagnostic and therapeutic value in POP.

280 **Acknowledgments**

281 We thank the authors who provided the GEO public datasets.

282 **Conflict of Interest Statement**

283 The authors declare no competing interests.

284

285

286

287 **FIGURE 1.** Workflow diagram. Data collection, analysis and validation.

288

289 **FIGURE 2.** Clustering dendrogram of 34 POP samples in GSE12852. POP samples were assigned

290 as 1; non-POP samples were assigned as 0. Pre-menopause status samples were assigned as 0;

291 menopause status samples were assigned as 1. Red color intensity is proportional to POP samples,

292 pre-menopause status, and higher age.

293 **FIGURE 3. Identification of the soft threshold in the scale-free network.**

294 (A) Different soft-threshold and corresponding scale free topology model(B) Different soft-

295 thresholding powers and corresponding mean connectivity. (C) The distribution of the

296 connectedness by histogram. (D) Inspection of the scale free topology.

297 **FIGURE 4. Identification of the hub module.**

298 (A) Dendrogram showed the genes with similar function classed into the same module according

299 to the dissimilarity measure. (B) Heatmap revealed the correlation between modular eigengenes

300 (ME) and clinical information of POP.

301 **FIGURE 5. PPI network analysis.**

302 PPI network analysis of the genes in the blue module based on the OmniPath online database.

303 Nodes represent the hub genes in the blue module identified by WGCNA. Lines represent

304 interactions between hub genes.

305 **FIGURE 6. GO functional and KEGG pathway enrichment analysis.**

306 (A) Biological process (BP), (B)cellular component (CC) and (C) molecular function (MF) of GO

307 analyzed by the "org.Hs.eg.db" package in the R software. (D) KEGG pathway enrichment

308 analysis using the "org.Hs.eg.db" package in the R software.

309 **FIGURE 7. Selected hub gene.**

310 (A) The relationship between gene MM and GS. (B) Differential expression of the hub genes. (C)

311 Heatmap of the hub genes in GSE12852. (D) Correlation analysis of the hub genes in GSE12852.

312 **FIGURE 8. Hub gene validation.**

313 (A−H) RT-qPCR detection of the expression of hub genes in 12 pairs of POP and non-POP tissues.

314 **FIGURE 9. The landscape of immune cell infiltration in GSE12852 (CIBERSORT *p* value <**

315 **0.05).**

316 (A) Proportion of the 22 immune cell types in GSE12852. (B) Correlation matrix between the 22

317 immune cell types.

318 **FIGURE 10. The differential immune cell infiltration between POP and non-POP tissues.**

319 (A) Resting mast cells; (B) activated mast cells; (C) neutrophils. Group 1: non-POP; Group 2:

320 POP.

321 **TABLE 1.** The correlation between hub genes and immune cell infiltration

322 **SUPPLEMENTARY TABLE 1.** Basic information of the patients in the validation dataset

323 **SUPPLEMENTARY TABLE 2.** The 22 immune cells

324 **SUPPLEMENTARY TABLE 3.**The primers for the hub genes

325 **SUPPLEMENTARY TABLE 4.** GO functional enrichment analysis

326 **SUPPLEMENTARY TABLE 5.** KEGG pathway enrichment analysis

327

328 **Reference**

329

330 Amatya N, Garg AV, and Gaffen SL. 2017. IL-17 Signaling: The Yin and the Yang.
331   *Trends Immunol* 38:310-322. 10.1016/j.it.2017.01.006
332 Babinger P, Volkl R, Cakstina I, Maftei A, and Schmitt R. 2007. Maintenance DNA
333   methyltransferase (Met1) and silencing of CpG-methylated foreign DNA in
334   Volvox carteri. *Plant Mol Biol* 63:325-336. 10.1007/s11103-006-9091-1
335 Blanchard E, Marie S, Riffault L, Bonora M, Tabary O, Clement A, and Jacquot J.
336   2011. Reduced expression of Tis7/IFRD1 protein in murine and human cystic
337   fibrosis airway epithelial cell models homozygous for the F508del-CFTR
338   mutation. *Biochemical and biophysical research communications* 411:471-

339   476. 10.1016/j.bbrc.2011.06.104

340   Brizzolara SS, Killeen J, and Urschitz J. 2009. Gene expression profile in
341   pelvic organ prolapse. *Mol Hum Reprod* 15:59-67. 10.1093/molehr/gan074

342   Cai W, Li H, Zhang Y, and Han G. 2020. Identification of key biomarkers and
343   immune infiltration in the synovial tissue of osteoarthritis by
344   bioinformatics analysis. *PeerJ* 8:e8390-e8390. 10.7717/peerj.8390

345   Chen B, Khodadoust MS, Liu CL, Newman AM, and Alizadeh AA. 2018a. Profiling
346   Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol Biol* 1711:243-
347   259. 10.1007/978-1-4939-7493-1_12

348   Chen Y, Bi F, An Y, and Yang Q. 2018b. Coexpression network analysis identified
349   Kruppel-like factor 6 (KLF6) association with chemosensitivity in ovarian
350   cancer. *J Cell Biochem*. 10.1002/jcb.27567

351   Cohen L, Mohr R, Chen YY, Huang M, Kato R, Dorin D, Tamanoi F, Goga A, Afar D,
352   Rosenberg N. 1994. Transcriptional activation of a ras-like gene (kir) by
353   oncogenic tyrosine kinases. *Proc Natl Acad Sci U S A* 91:12448-12452.
354   10.1073/pnas.91.26.12448

355   Dietz HP, Jarvis SK, and Vancaillie TG. 2002. The assessment of levator muscle
356   strength: a validation of three ultrasound techniques. *Int Urogynecol J*
357   *Pelvic Floor Dysfunct* 13:156-159; discussion 159. 10.1007/s192-002-8346-
358   5

359   Friedman T, Eslick GD, and Dietz HP. 2018. Risk factors for prolapse recurrence:
360   systematic review and meta-analysis. *Int Urogynecol J* 29:13-21.
361   10.1007/s00192-017-3475-4

362   Grabauskas G, Wu X, Gao J, Li J-Y, Turgeon DK, and Owyang C. 2020. Prostaglandin
363   E2, Produced by Mast Cells in Colon Tissues from Patients with Irritable
364   Bowel Syndrome, Contributes to Visceral Hypersensitivity in Mice.
365   *Gastroenterology*:S0016-5085(0020)30232-30238.
366   10.1053/j.gastro.2020.02.022

367   Gu Y, Harley IT, Henderson LB, Aronow BJ, Vietor I, Huber LA, Harley JB,
368   Kilpatrick JR, Langefeld CD, Williams AH, Jegga AG, Chen J, Wills-Karp M,
369   Arshad SH, Ewart SL, Thio CL, Flick LM, Filippi MD, Grimes HL, Drumm ML,
370   Cutting GR, Knowles MR, and Karp CL. 2009a. Identification of IFRD1 as a
371   modifier gene for cystic fibrosis lung disease. *Nature* 458:1039-1042.
372   10.1038/nature07811

373   Gu Y, Harley ITW, Henderson LB, Aronow BJ, Vietor I, Huber LA, Harley JB,
374   Kilpatrick JR, Langefeld CD, Williams AH, Jegga AG, Chen J, Wills-Karp M,
375   Arshad SH, Ewart SL, Thio CL, Flick LM, Filippi M-D, Grimes HL, Drumm ML,
376   Cutting GR, Knowles MR, and Karp CL. 2009b. Identification of IFRD1 as a
377   modifier gene for cystic fibrosis lung disease. *Nature* 458:1039-1042.

378   10.1038/nature07811

379 Haya N, Feiner B, Baessler K, Christmann-Schmid C, and Maher C. 2018.
380   Perioperative interventions in pelvic organ prolapse surgery. *Cochrane*
381   *Database Syst Rev* 8:Cd013105. 10.1002/14651858.Cd013105

382 Hector A, Kormann M, Kammermeier J, Burdi S, Marcos V, Rieber N, Mays L, Illig
383   T, Klopp N, Falkenstein F, Kappler M, Riethmueller J, Graepler-Mainka U,
384   Stern M, Eickmeier O, Serve F, Zielen S, Döring G, Griese M, and Hartl D.
385   2013. Expression and regulation of interferon-related development
386   regulator-1 in cystic fibrosis neutrophils. *American journal of*
387   *respiratory cell and molecular biology* 48:71-77. 10.1165/rcmb.2012-0061OC

388 Hejzlarova K, Tesarova M, Vrbacka-Cizkova A, Vrbacky M, Hartmannova H, Kaplanova
389   V, Noskova L, Kratochvilova H, Buzkova J, Havlickova V, Zeman J, Kmoch S,
390   and Houstek J. 2011. Expression and processing of the TMEM70 protein.
391   *Biochim Biophys Acta* 1807:144-149. 10.1016/j.bbabio.2010.10.005

392 Hendrix SL, Clark A, Nygaard I, Aragaki A, Barnabei V, and McTiernan A. 2002.
393   Pelvic organ prolapse in the Women's Health Initiative: gravity and
394   gravidity. *Am J Obstet Gynecol* 186:1160-1166. 10.1067/mob.2002.123819

395 Huang T, Sun L, Yuan X, and Qiu H. 2017. Thrombospondin-1 is a multifaceted
396   player in tumor progression. *Oncotarget* 8:84546-84558.
397   10.18632/oncotarget.19165

398 Huang X, Cong X, Yang D, Ji L, Liu Y, Cui X, Cai J, He S, Zhu C, Ni R, and Zhang
399   Y. 2014. Identification of Gem as a new candidate prognostic marker in
400   hepatocellular carcinoma. *Pathology, research and practice* 210:719-725.
401   10.1016/j.prp.2014.07.001

402 Jiang S, Linghu E, Zhan Q, Han W, and Guo M. 2015. Methylation of ZNF331 Promotes
403   Cell Invasion and Migration in Human Esophageal Cancer. *Curr Protein Pept*
404   *Sci* 16:322-328. 10.2174/1389203716041504291555255

405 Kashyap MK, Amaya-Chanaga CI, Kumar D, Simmons B, Huser N, Gu Y, Hallin M,
406   Lindquist K, Yafawi R, Choi MY, Amine A-A, Rassenti LZ, Zhang C, Liu S-
407   H, Smeal T, Fantin VR, Kipps TJ, Pernasetti F, and Castro JE. 2017.
408   Targeting the CXCR4 pathway using a novel anti-CXCR4 IgG1 antibody (PF-
409   06747143) in chronic lymphocytic leukemia. *J Hematol Oncol* 10:112-112.
410   10.1186/s13045-017-0435-x

411 Koltai E, Bori Z, Osvath P, Ihasz F, Peter S, Toth G, Degens H, Rittweger J,
412   Boldogh I, and Radak Z. 2018. Master athletes have higher miR-7, SIRT3
413   and SOD2 expression in skeletal muscle than age-matched sedentary
414   controls. *Redox biology* 19:46-51. 10.1016/j.redox.2018.07.022

415 Kraus F, Haenig B, and Kispert A. 2001. Cloning and expression analysis of the
416   mouse T-box gene tbx20. *Mech Dev* 100:87-91. 10.1016/s0925-4773(00)00499-

417   8
418 Limón-Flores AY, Chacón-Salinas R, Ramos G, and Ullrich SE. 2009. Mast cells
419   mediate the immune suppression induced by dermal exposure to JP-8 jet
420   fuel. *Toxicological sciences : an official journal of the Society of*
421   *Toxicology* 112:144-152. 10.1093/toxsci/kfp181
422 Lin J, Yu M, Xu X, Wang Y, Xing H, An J, Yang J, Tang C, Sun D, and Zhu Y. 2020.
423   Identification of biomarkers related to CD8(+) T cell infiltration with
424   gene co-expression network in clear cell renal cell carcinoma. *Aging*
425   *(Albany NY)* 12:10.18632/aging.102841. 10.18632/aging.102841
426 Lincoln J, Alfieri CM, and Yutzey KE. 2004. Development of heart valve leaflets
427   and supporting apparatus in chicken and mouse embryos. *Dev Dyn* 230:239-
428   250. 10.1002/dvdy.20051
429 Liu R, Hu R, Zeng Y, Zhang W, and Zhou H-H. 2020. Tumour immune cell infiltration
430   and survival after platinum-based chemotherapy in high-grade serous
431   ovarian cancer subtypes: A gene expression-based computational study.
432   *EBioMedicine* 51:102602-102602. 10.1016/j.ebiom.2019.102602
433 Lv Y, Zhao Y, Wang X, Chen N, Mao F, Teng Y, Wang T, Peng L, Zhang J, Cheng P,
434   Liu Y, Kong H, Chen W, Hao C, Han B, Ma Q, Zou Q, Chen J, and Zhuang Y.
435   2019. Increased intratumoral mast cells foster immune suppression and
436   gastric cancer progression through TNF-α-PD-L1 pathway. *Journal for*
437   *immunotherapy of cancer* 7:54-54. 10.1186/s40425-019-0530-3
438 Novruzov RM. 2008. Hemomicrocirculation channel and mastocyte population in
439   outer oblique aponeurosis stomach muscles and uterus ligament in women
440   with small pelvic organ prolapses. *Georgian medical news*:69-72.
441 Nygaard IE, Shaw JM, Bardsley T, and Egger MJ. 2014. Lifetime physical activity
442   and pelvic organ prolapse in middle-aged women. *Am J Obstet Gynecol*
443   210:477.e471-412. 10.1016/j.ajog.2014.01.035
444 Ottaiano A, Scala S, Normanno N, Botti G, Tatangelo F, Di Mauro A, Capozzi M,
445   Facchini S, Tafuto S, and Nasti G. 2020. Prognostic and Predictive Role
446   of CXC Chemokine Receptor 4 in Metastatic Colorectal Cancer Patients.
447   *Applied immunohistochemistry & molecular morphology :*
448   *AIMM*:10.1097/PAI.0000000000000828. 10.1097/PAI.0000000000000828
449 Patadia M, Dixon J, Conley D, Chandra R, Peters A, Suh LA, Kato A, Carter R,
450   Harris K, Grammer L, Kern R, and Schleimer R. 2010. Evaluation of the
451   presence of B-cell attractant chemokines in chronic rhinosinusitis.
452   *American journal of rhinology & allergy* 24:11-16.
453   10.2500/ajra.2010.24.3386
454 Pegg AE. 2008. Spermidine/spermine-N(1)-acetyltransferase: a key metabolic
455   regulator. *American journal of physiology Endocrinology and metabolism*

456   294:E995-E1010. 10.1152/ajpendo.90217.2008
457 Pegg AE. 2016. Functions of Polyamines in Mammals. *J Biol Chem* 291:14904-14912.
458   10.1074/jbc.R116.731661
459 Qu S, Yang L, and Liu Z. 2020. MicroRNA-194 reduces inflammatory response and
460   human dermal microvascular endothelial cells permeability through
461   suppression of TGF-β/SMAD pathway by inhibiting THBS1 in chronic
462   idiopathic urticaria. *J Cell Biochem* 121:111-124. 10.1002/jcb.28941
463 Sajay-Asbaghi M, Sadeghi-Shabestrai M, Monfaredan A, Seyfizadeh N, Razavi A,
464   and Kazemi T. 2020. Promoter region single nucleotide polymorphism of
465   siglec-8 gene associates with susceptibility to allergic asthma.
466   *Personalized medicine*:10.2217/pme-2018-0080. 10.2217/pme-2018-0080
467 Teicher BA, and Fricker SP. 2010. CXCL12 (SDF-1)/CXCR4 pathway in cancer. *Clin
468   Cancer Res* 16:2927-2931. 10.1158/1078-0432.CCR-09-2329
469 Türei D, Korcsmáros T, and Saez-Rodriguez J. 2016. OmniPath: guidelines and
470   gateway for literature-curated signaling pathway resources. *Nat Methods*
471   13:966-967. 10.1038/nmeth.4077
472 Wang D, Wang X, Si M, Yang J, Sun S, Wu H, Cui S, Qu X, and Yu X. 2020a. Exosome-
473   encapsulated miRNAs contribute to CXCL12/CXCR4-induced liver metastasis
474   of colorectal cancer by enhancing M2 polarization of macrophages. *Cancer
475   letters* 474:36-52. 10.1016/j.canlet.2020.01.005
476 Wang H, Chen G, Wang H, and Liu C. 2013. RITA inhibits growth of human
477   hepatocellular carcinoma through induction of apoptosis. *Oncol Res* 20:437-
478   445. 10.3727/096504013x13685487925059
479 Wang M, Wang J, Liu J, Zhu L, Ma H, Zou J, Wu W, and Wang K. 2020b. Systematic
480   prediction of key genes for ovarian cancer by co-expression network
481   analysis. *J Cell Mol Med*. 10.1111/jcmm.15271
482 Wang Y, He T, Herman JG, Linghu E, Yang Y, Fuks F, Zhou F, Song L, and Guo M.
483   2017. Methylation of ZNF331 is an independent prognostic marker of
484   colorectal cancer and promotes colorectal cancer growth. *Clin Epigenetics*
485   9:115. 10.1186/s13148-017-0417-4
486 Wu JM, Hundley AF, Fulton RG, and Myers ER. 2009. Forecasting the prevalence of
487   pelvic floor disorders in U.S. Women: 2010 to 2050. *Obstet Gynecol*
488   114:1278-1283. 10.1097/AOG.0b013e3181c2ce96
489 Yu J, Liang QY, Wang J, Cheng Y, Wang S, Poon TC, Go MY, Tao Q, Chang Z, and
490   Sung JJ. 2013. Zinc-finger protein 331, a novel putative tumor suppressor,
491   suppresses growth and invasiveness of gastric cancer. *Oncogene* 32:307-
492   317. 10.1038/onc.2012.54
493 Yu Q, Vazquez R, Zabadi S, Watson RR, and Larson DF. 2010. T-lymphocytes mediate
494   left ventricular fibrillar collagen cross-linking and diastolic

495   dysfunction in mice. *Matrix Biol* 29:511–518. 10.1016/j.matbio.2010.06.003

496 Zhang B, Wang Q, Fu C, Jiang C, and Ma S. 2019. Exploration of the immune-
497   related signature and immune infiltration analysis for breast ductal and
498   lobular carcinoma. *Annals of translational medicine* 7:730–730.
499   10.21037/atm.2019.11.117

500 Zhao C, Isenberg JS, and Popel AS. 2018. Human expression patterns: qualitative
501   and quantitative analysis of thrombospondin-1 under physiological and
502   pathological conditions. *J Cell Mol Med* 22:2086–2097. 10.1111/jcmm.13565

503 Zhou Z, Cheng Y, Jiang Y, Liu S, Zhang M, Liu J, and Zhao Q. 2018. Ten hub genes
504   associated with progression and prognosis of pancreatic carcinoma
505   identified by co-expression analysis. *Int J Biol Sci* 14:124–136.
506   10.7150/ijbs.22619

507 Zhu Q, Luo R, Gu J, Hou Y, Chen Z, Xu F, Wang L, Mao W, Lu C, and Ge D. 2020.
508   High CXCR4 Expression Predicts a Poor Prognosis in Resected Lung
509   Adenosquamous Carcinoma. *Journal of Cancer* 11:810–818. 10.7150/jca.36498

510

511

512

Figure 1

figure1

FIGURE 1. Workflow diagram. Data collection, analysis and validation.

Downloading data of GSE12852 from GEO database as training dataset

Remove outlier sample of GSE12852 dataset by "flash cluster" package of R software

Construct co-expression network and screen hub module

Construct PPI network using genes in hub module

GO and KEGG enrichment analysis for hub module

Screen hub genes and validate hub genes using the validation dataset

CIBERSORT algorithm characterized the proportion of 22 immune cells in samples from GSE12852 dataset

Investigation the differential immune cells infiltration between POP tissues and non-POP tissues

Figure 2

figure2

FIGURE 2. Clustering dendrogram of 34 POP samples in GSE12852. POP samples were assigned as 1; non-POP samples were assigned as 0. Pre-menopause status samples were assigned as 0; menopause status samples were assigned as 1. Red color intensity is proportional to POP samples, pre-menopause status, and higher age.
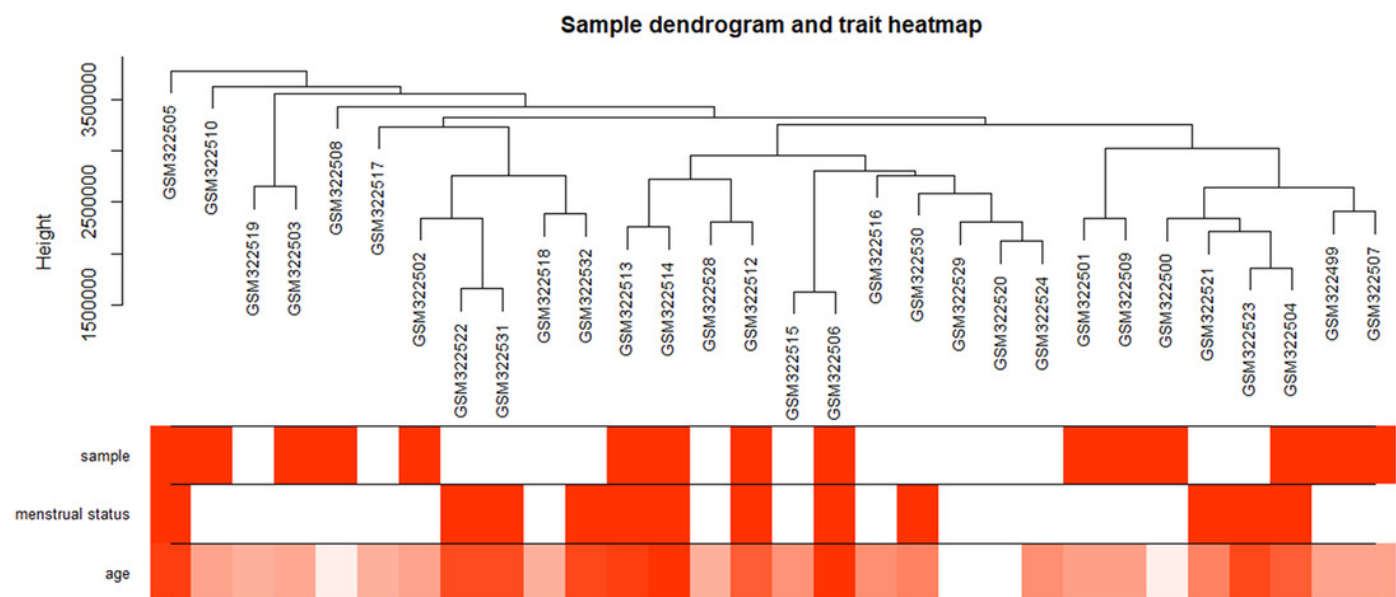
Figure 3

figure3

FIGURE 3. Identification of the soft threshold in the scale-free network. (A) Different soft-threshold and corresponding scale free topology model(B) Different soft-thresholding powers and corresponding mean connectivity. (C) The distribution of the connectedness by histogram. (D) Inspection of the scale free topology.

A **Scale independence**

B **Mean connectivity**

C **Histogram of k**

D **Check scale free topology**
scale R^2= 0.98 , slope= -2.03
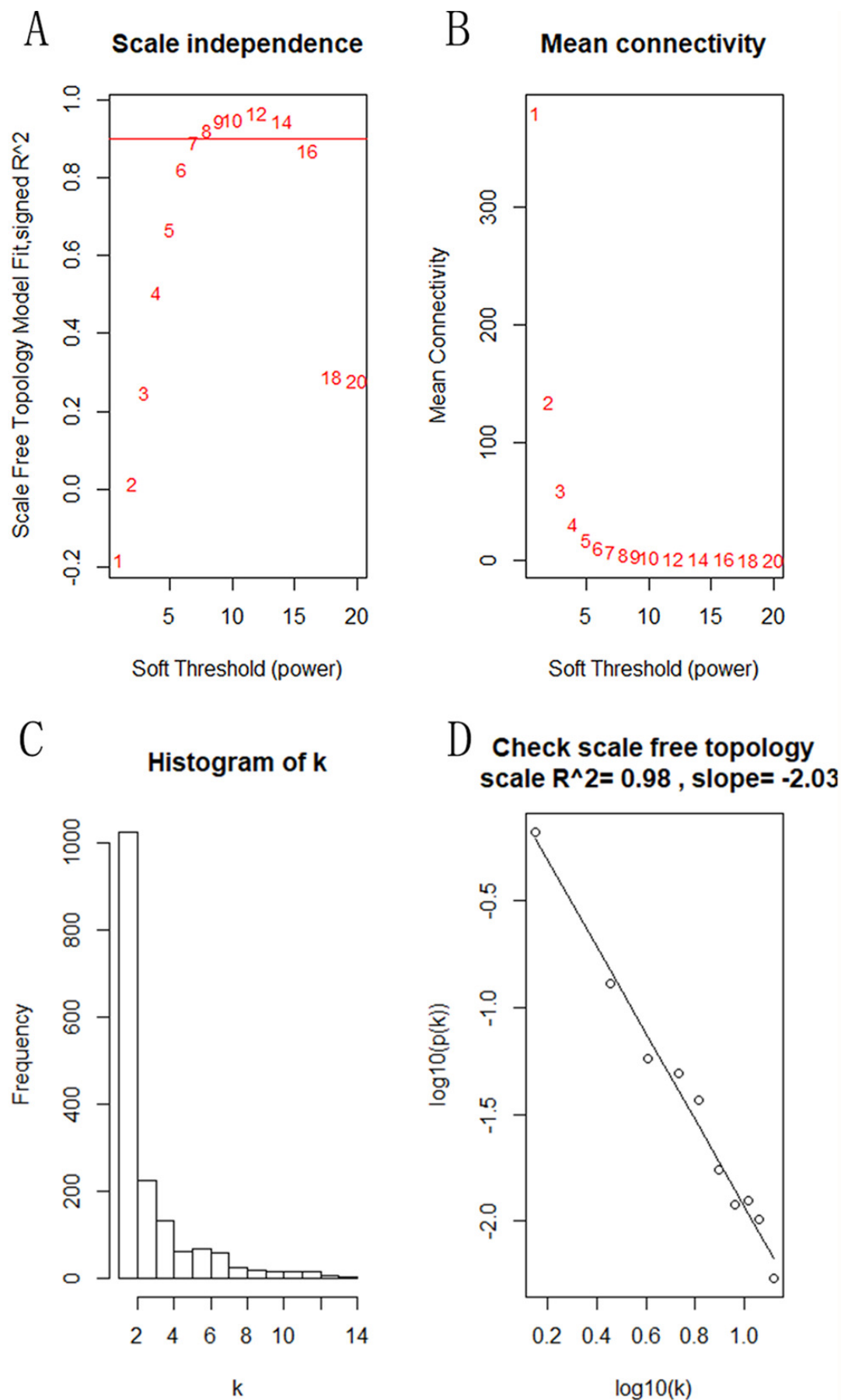
Figure 4

figure4

FIGURE 4. Identification of the hub module. (A) Dendrogram showed the genes with similar function classed into the same module according to the dissimilarity measure. (B) Heatmap revealed the correlation between modular eigengenes (ME) and clinical information of POP.

Cluster Dendrogram

A

B

Module−trait relationships
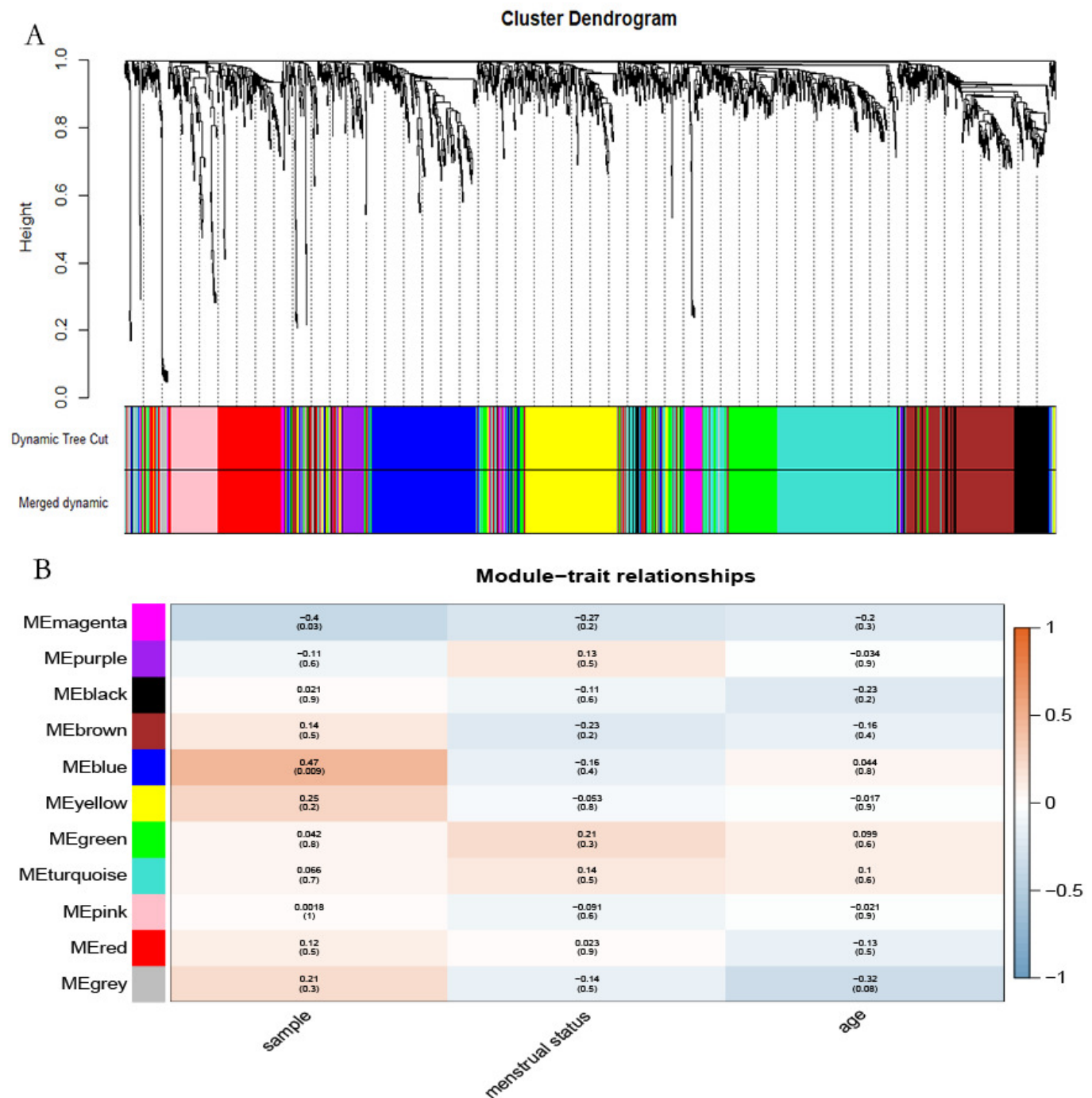
Figure 5

figure5

FIGURE 5. PPI network analysis. PPI network analysis of the genes in the blue module based on the OmniPath online database. Nodes represent the hub genes in the blue module identified by WGCNA. Lines represent interactions between hub genes.
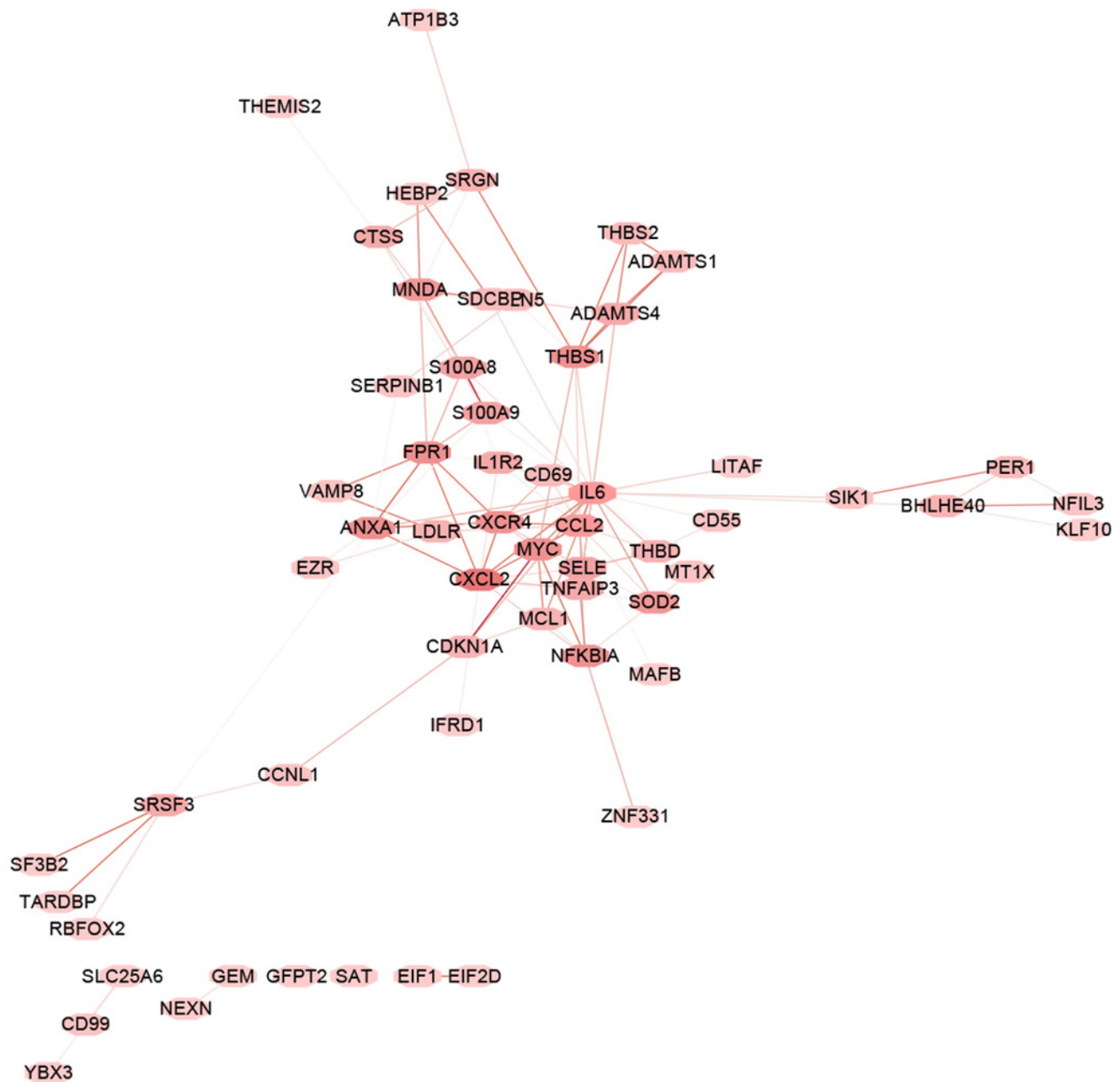
Figure 6

figure6

FIGURE 6. GO functional and KEGG pathway enrichment analysis. (A) Biological process (BP), (B)cellular component (CC) and (C) molecular function (MF) of GO analyzed by the "org.Hs.eg.db" package in the R software. (D) KEGG pathway enrichment analysis using the "org.Hs.eg.db" package in the R software.
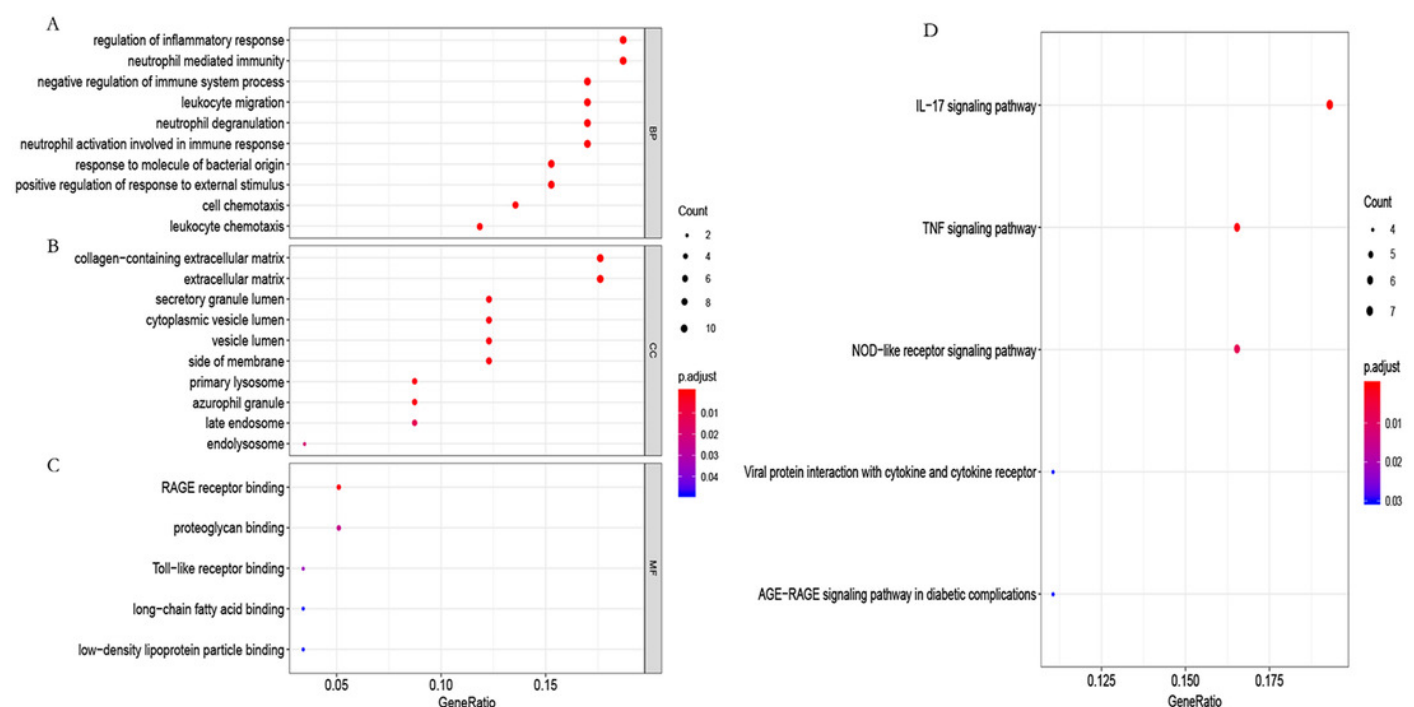
Figure 7

figure7

FIGURE 7. Selected hub gene. (A) The relationship between gene MM and GS. (B) Differential expression of the hub genes. (C) Heatmap of the hub genes in GSE12852. (D) Correlation analysis of the hub genes in GSE12852.
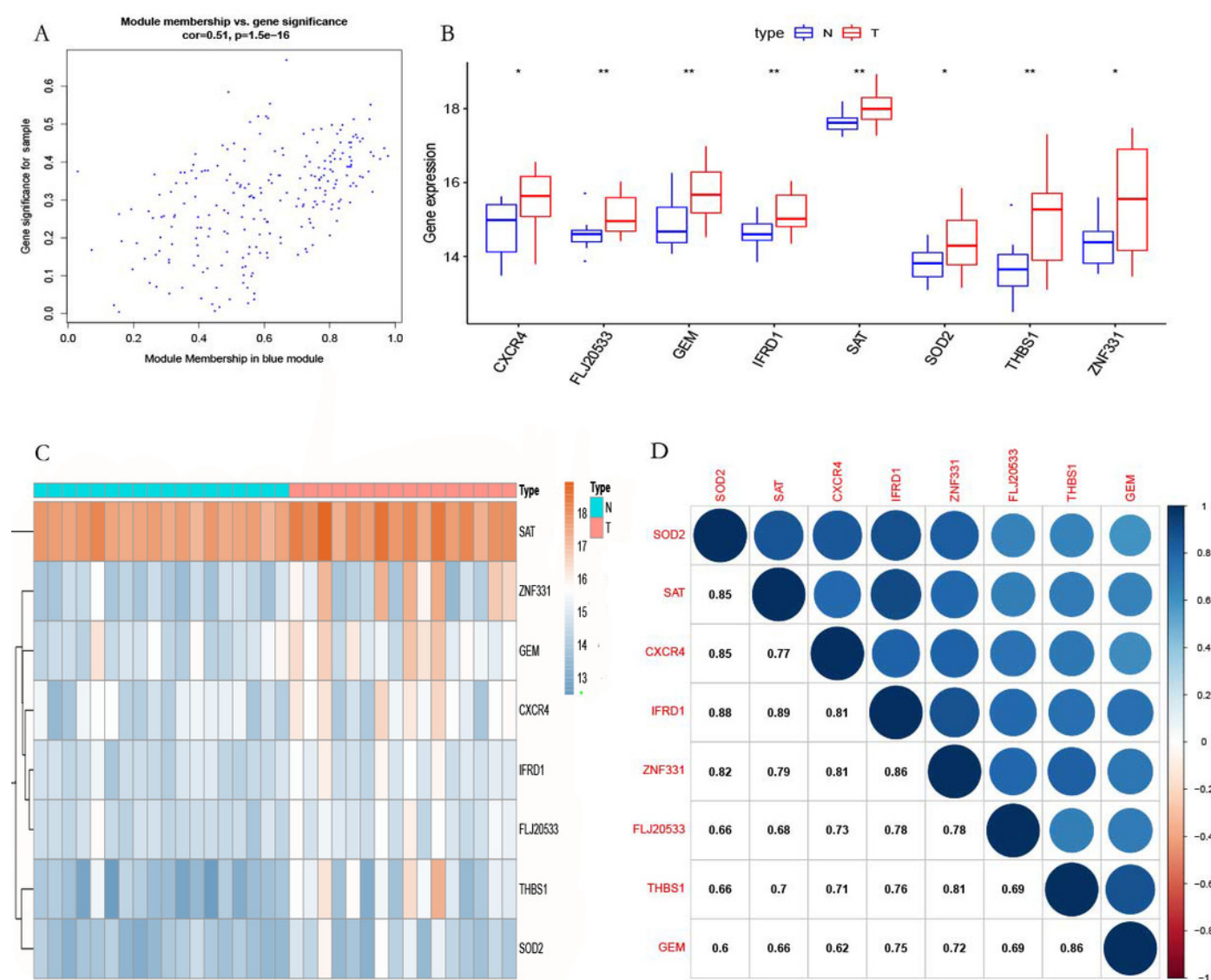
Figure 8

figure8

FIGURE 8. Hub gene validation. (A−H) RT-qPCR detection of the expression of hub genes in 12 pairs of POP and non-POP tissues.
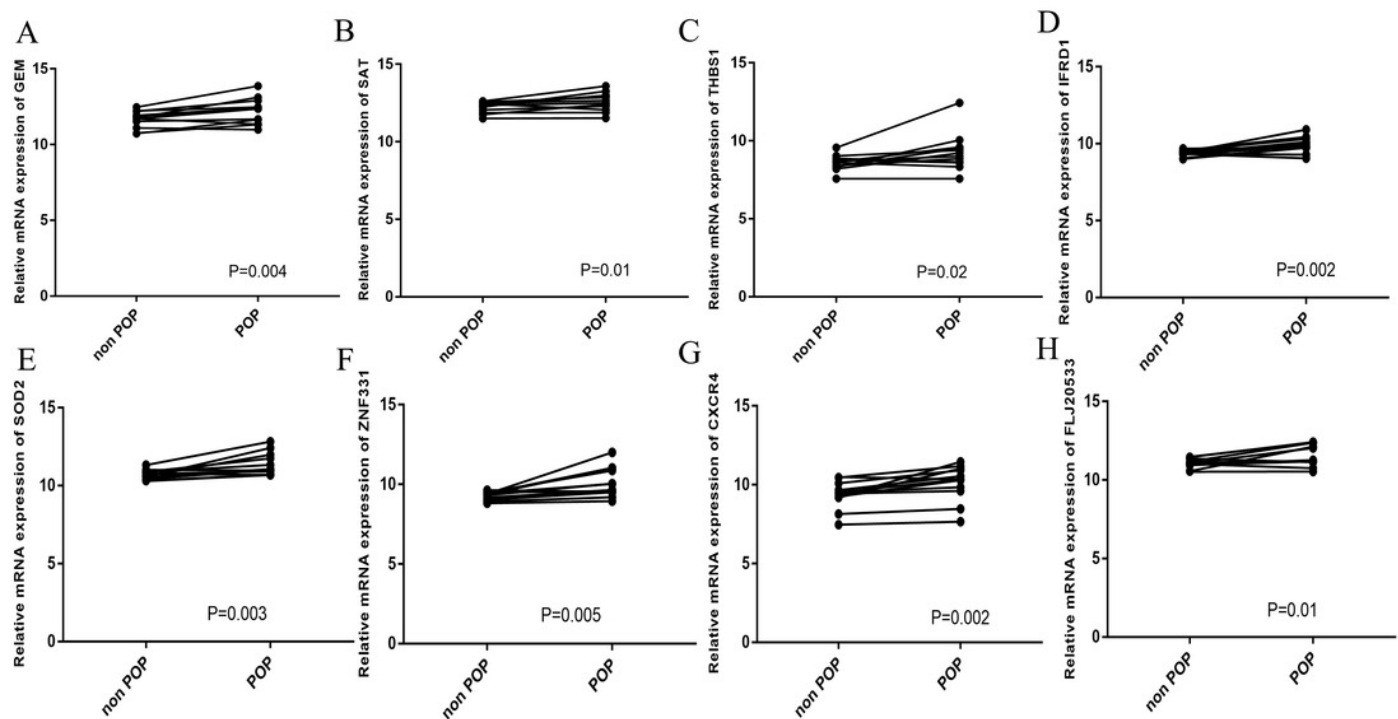
Figure 9

figure9

FIGURE 9. The landscape of immune cell infiltration in GSE12852 (CIBERSORT *p* value < 0.05). (A) Proportion of the 22 immune cell types in GSE12852. (B) Correlation matrix between the 22 immune cell types.
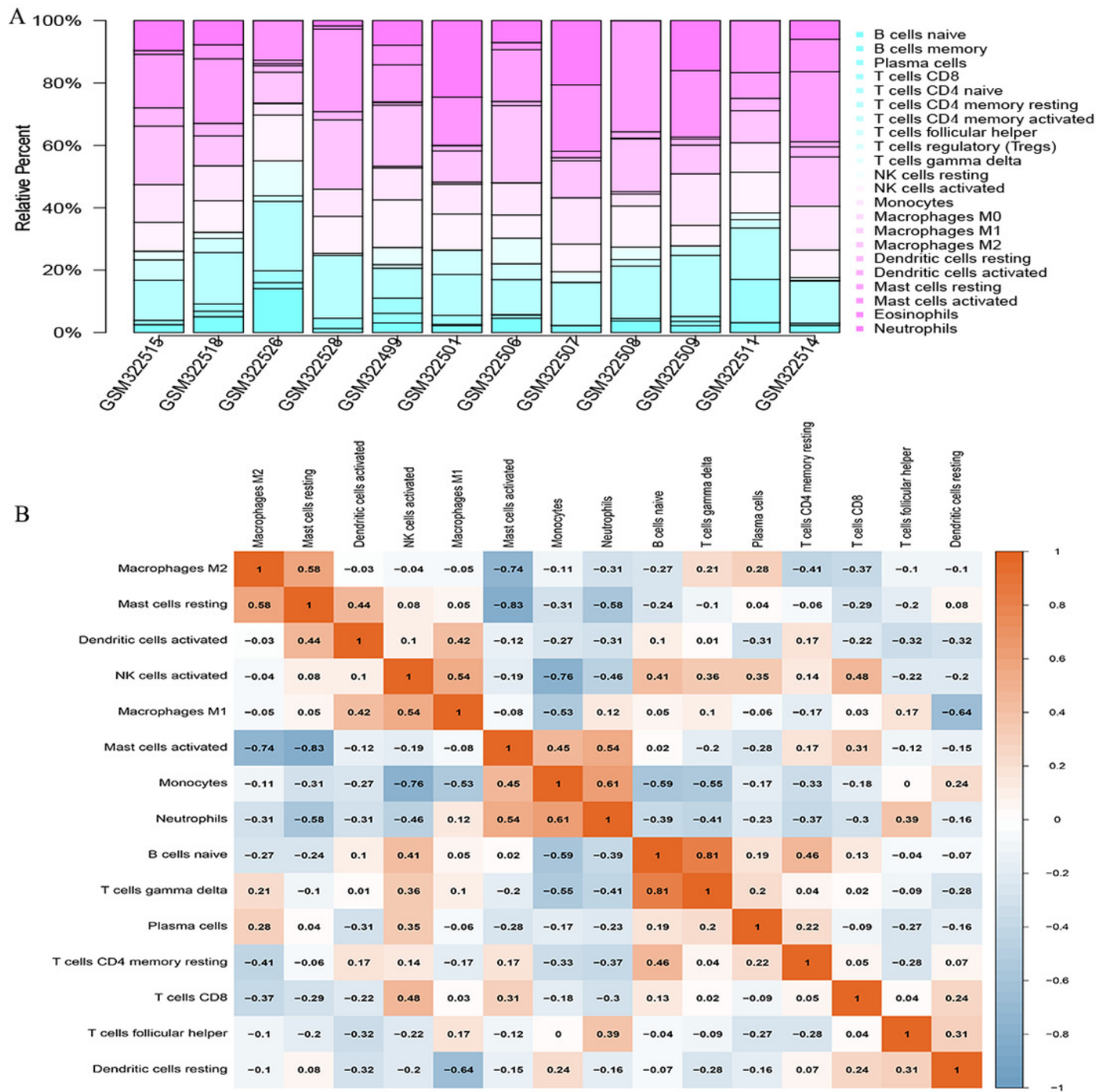
Figure 10

figure10

FIGURE 10. The differential immune cell infiltration between POP and non-POP tissues. (A) Resting mast cells; (B) activated mast cells; (C) neutrophils. Group 1: non-POP; Group 2: POP.
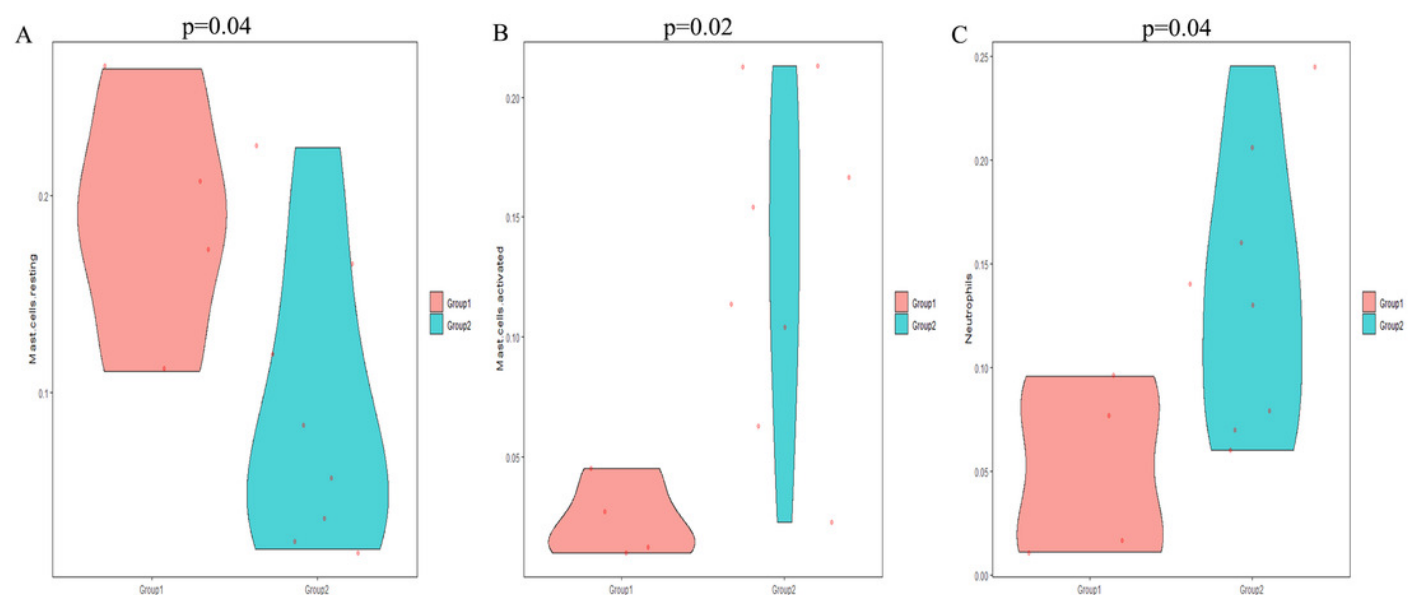
table1

TABLE 1. The correlation between hub genes and immune cell infiltration

GENE	IIMMUNE CELL	P	R
CXCR4	Mast cells activated	0.026025009	0.636596855
FLJ20533	Mast cells activated	0.037410947	0.604312579
GEM	Mast cells activated	0.00382779	0.763874376
IFRD1	Mast cells activated	0.010102021	0.707220887
SAT	Mast cells activated	0.026678483	0.634493564
SOD2	Mast cells activated	0.019559382	0.659816924
THBS1	Mast cells activated	0.023473376	0.645196019
ZNF331	Mast cells activated	0.005806179	0.741209161
GEM	Mast cells resting	0.044390877	-0.587895781
IFRD1	Mast cells resting	0.010475671	-0.704821204
SAT	Mast cells resting	0.029213165	-0.626671202
SOD2	Mast cells resting	0.002740399	-0.780472841
CXCR4	Neutrophils	0.002884121	0.778019124
FLJ20533	Neutrophils	0.009179588	0.713441726
GEM	Neutrophils	0.030016479	0.624294555
IFRD1	Neutrophils	0.001688177	0.802313109
SAT	Neutrophils	0.001387584	0.810472605
SOD2	Neutrophils	2.12E-04	0.872716498
THBS1	Neutrophils	0.014406502	0.682777143
ZNF331	Neutrophils	0.002414909	0.786415881