Identification of Key Genes Involved in the Pathogenesis of Recurrent Pelvic Organ Prolapse Using Bioinformatics Analysis

Ling Zhu
Hangzhou Ninth People's Hospital
Xiaoling Ni
Hangzhou Women's Hospital (Hangzhou Maternity and Child Health Care Hospital)
Shanshan Tang
Hangzhou Women's Hospital (Hangzhou Maternity and Child Health Care Hospital)
Wenhua Liu (945368820@qq.com)
Hangzhou Women's Hospital (Hangzhou Maternity and Child Health Care Hospital)

Research Article

Keywords: Differentially expressed gene, Recurrent pelvic organ prolapse, Bioinformatics analysis

Posted Date: November 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1107398/v1

Abstract

Background: The causes of the recurrence of pelvic organ prolapse (POP) are sufficiently understood. However, few studies investigate the key genes of recurrence POP. The present study aimed to screen the hub genes of recurrence POP. Microarray data of 4 recurrent POP and 4 primary POP uterosacral ligaments in the GSE28660 gene expression dataset were used as research objects. we used the online Gene Expression Omnibus (GEO) microarray expression profiling dataset to identify differentially expressed genes (DEGs). Also, functional enrichment and protein-protein interaction (PPI) network analyses were performed, and the key modules were identified. Then, we investigated the differential immune cell infiltration between recurrent POP and primary POP tissues using the CIBERSORT algorithm.

Results: In total, 84 upregulated and 32 downregulated genes were identified in the differential expression analysis.

Conclusion: This human genome DNA microarrays analysis identified a recurrence POP signature of 116 genes, and 2 hub genes, including cell death-inducing DFFA-like effector (CIDEA) and hemoglobin subunit delta (HBD) may participate in the pathogenesis of recurrence POP, giving them a certain diagnostic and therapeutic value.

Introduction

Pelvic Organ Prolapse (POP) is a common gynecological condition related to pelvic floor dysfunction in women, is estimated to have a prevalence of 30-50% among women, aged 50 and over. The prevalence of POP is expected to reach 46% by 2050[1]. However, the recurrence POP rate is high, which causes a huge economic burden. A retrospective cohort study involving 1,811 patients who underwent primary surgery for POP over almost 20 years showed a re-operation incidence of 5.1 per 1,000 women-years, with a cumulative incidence of 5.6 %[2]. Parity, vaginal delivery, age, and Body Mass Index (BMI) are risk factors for POP, and the preoperative stage is a risk factor for POP recurrence[3]. Nevertheless, the molecular mechanism of recurrence POP remains unclear, and there is a lack of suitable prevention and treatment measures in clinical practice. Therefore, it is imperative to explore the molecular mechanism of recurrence POP for the benefit of both women and society.

The uterosacral ligaments (USLs) are condensations of endopelvic fascia and contribute to primary uterine support, the disruption of structural components of these ligaments may lead to a loss of support and eventually cause POP[4,5]. A decrease of the connective tissue extracellular matrix (ECM) proteins, such as collagen and elastic in USLs of POP in the form of quality and quantity, has been discovered. Bioinformatics analysis has become an important tool in the analysis of biomedical literature[6,7]. Protein-Protein Interaction (PPI) network analysis is one such tool, which identifies gene module characteristics and hub genes to connect the gene modules and sample characteristics[8]. A new algorithm for enumeration of immune cell subsets, CIBERSORT (https://cibersort.stanford.edu/), provides the possibility to identify immune biomarkers for diagnosis and prognosis[9]. Our study found hub genes and

immune cells highly related to recurrent POP by analyzing expression spectrum data in public databases using (PPI) network analysis and the CIBERSORT algorithm, providing novel ideas and methods for the treatment of recurrence POP.

Results

Differentially Expressed Genes

We downloaded the microarray expression dataset GSE28660 from the GEO database and analyzed the DEGs between primary POP and recurrent POP patients using the online analysis tool GEO2R. In total, 84 upregulated and 32 downregulated genes were identified in the differential expression analysis. The results were shown in Table 1. The heatmap and volcano map showed the differentially expression expressed genes in primary POP and recurrent POP groups. The top 20 genes-PLIN1, PPP1R1A, LPL, TNMD, HBM, RBP4, CIDEA, LEP, LOC101927531, LINC01618, JUNB, EGR1, DUSP1, HSD17B8, KLF4, RGCC, INHBB, RBP7, BTNL9, HBD were analyzed in a heatmap. The results were shown in Fig. 1,2.

Functional and Pathway Enrichment of DEGs

In our study, 17 enriched functional category terms, 30 enriched GO terms, and 3 KEGG pathways were identified. The enriched GO terms with $P < 0.05$ are presented in Table 2; they included positive regulation of transcription from RNA (P=0.01951), negative regulation of transcription from RNA polymerase II promoter (P=0.046058), intracellular signal transduction (P=0.024445), inflammatory response (P=0.018727), lipid metabolic process (P=0.011286), chemotaxis (P=0.004709)in the CC category. In addition, 6 enriched UP_KEYWORDS terms and 3 enriched UP_SEQ_FEATURE terms with $P < 0.05$ were identified, including Heme (P=5.33E-04), NAD (P=1.16E-02), Disulfide bond (P=1.16E-02), Iron (P=2.35E-02), Calcium (P =3.17E-02), DNA-binding region:Basic motif (P =1.41E-03), Disulfide bond (P=1.74E-02), and Binding site:Substrate (P=1.81E-02). The number of genes and P-values of the 15 enriched functional terms are displayed in Fig.3.

PPI Network Analysis of DEGs

A PPI network with 34 nodes and 53 edges was obtained; the network had an interaction score >0.4 according to the STRING online database (Fig.4A). The nodes correspond to genes, and the edges represent the links between genes. Red represents upregulated genes, and green represents downregulated genes. We used Mcode in Cytoscape to perform network gene clustering to identify the key PPI network modules. As shown in Fig. 4B, C, two key modules with two upregulated genes (CIDEA and HBD) were identified. Furthermore, functional enrichment analysis indicated that these two genes were mainly involved in activates apoptosis, immune response, and inflammation.

Identification of the key genes

We used Mcode in Cytoscape to perform network gene clustering to identify the key PPI network modules. As shown in Fig.4B,C, two key modules with two upregulated genes (CIDEA and HBD) were identified.

Furthermore, functional enrichment analysis indicated that these two genes were mainly involved in activates apoptosis, oxygen transport.

Immune cell infiltration analysis

The CIBERSORT algorithm was employed to select samples with an output $p < 0.05$. A total of 5 samples including primary POP and recurrent POP tissues were obtained. A bar plot was generated to show the proportion of 22 immune cells in the 5 samples. We found that there was no significant difference between the 22 immune cells in the 5 samples (Fig.5).

Discussion

Current advances in gene microarray technology and bioinformatics analysis offer new opportunities to discover the

functional genes for certain diseases. In the present study, 4 primary POP and 4 recurrent POP tissues were obtained from the GEO GSE28660 data set. A total of 116 DEGs were identified, including 84 upregulated and 32 downregulated genes. To obtain a more in-depth understanding of these DEGs, we performed GO function and KEGG pathway analyses of the DEGs. The results showed that the DEGs were mainly enriched for the following pathways: response to linoleic acid, negative regulation of fibroblast growth factor production, response to stilbenoid, negative regulation of cartilage development, positive regulation of follicle-stimulating hormone secretion, positive regulation of cholesterol storage, positive regulation of sequestering of triglyceride, low-density lipoprotein particle clearance,

negative regulation of heterotypic cell-cell adhesion, cellular response to leptin stimulus, oxygen transport, response to dietary excess negative regulation of hormone secretion, central nervous system neuron development, negative regulation of DNA biosynthetic process, PPAR signaling pathway, adipocytokine signaling pathway, Cytokine-cytokine receptor interaction. Moreover, genes with high degrees of differential expression were obtained by using a PPI network and modules analysis; CIDEA and HBD were identified as hub nodes. The two most significant submodules of DEGs were extracted from the PPI network. Module analysis of the DEGs suggested that activates apoptosis, oxygen transport might be involved in POP development.

As the most significant hub gene, Cell death-inducing DNA fragmentation factor-like effector A (CIDEA), initially identified as an apoptotic gene, was discovered to interfere with the energy metabolic process [10]. A previous study indicated that upregulated CIDEA could lead to in several organs [11]. Moreover, evidence shows that CIDEA is also known to induce apoptosis [12]. PPARγ regulated CIDEA affects pro-apoptotic responses[13]. Thus, the functional features of CIDEA may partially explain the weakening of pelvic floor muscle tissue in POP. A previous study found that HBD was closely associated with inflammation, and upregulation of HBD was observed during infection and inflammation [14]. Recent studies have revealed that immunoregulatory and inflammatory processes play a crucial role in repairing the extracellular

matrix after pelvic floor muscle injuries resulting from pregnancy and childbirth[15] Thus, this study shows that cell apoptosis and inflammation are important factors in the development and progression of recurrent POP.

Recent studies have revealed that immunoregulatory and inflammatory processes play a crucial role in repairing the extracellular matrix after pelvic floor muscle injuries resulting from pregnancy and childbirth[16]. In view of the close relationship between hub genes and immunity, we investigated immune cell infiltration in recurrent POP and primary POP tissues using the CIBERSORT algorithm. Unfortunately, We found that there was no significant difference between the 22 immune cells in recurrent POP and primary POP tissues. It's possible that the sample size we studied was too small. Zhao, et al. identified eight hub genes and three immune cell types that may be related to POP occurrence[17]. Li, et al. confirmed that both non-immune and immune cell types were involved in extracellular matrix (ECM) dysregulation and immune reactions involvement[18]. Therefore, the relationship between immune cell

Infiltration and recurrent pelvic floor organ prolapse needs to be further confirmed.

There were several limitations to the current study. First, the data used in this study were obtained from GEO, rather than directly from POP patients. Second, the results were only analyzed using bioinformatics; experimental verification is required to better confirm the findings of the identified genes and pathways in our investigation. Third, the sample size for the microarray analysis was relatively small. Only 8 samples from 8 patients (4 recurrent POP and 4 primary POP) were obtained from the GSE28660 data set for bioinformatics analysis. It will be necessary to recruit more subjects in the future to get more accurate correlation results. A series of verification experiments must be performed based on a larger sample size to confirm our results.

Conclusion

In summary, we identified several new key genes and pathways closely associated with recurrent POP using a series of bioinformatics analyses on DEGs between 4 recurrent POP samples and 4 primary POP samples. The key genes CIDEA and HBD might play important roles in the development and progression of POP. Furthermore, PPAR signaling pathway, adipocytokine signaling pathway, Cytokine-cytokine receptor interaction potentially contributed to POP development. This study shows that cell apoptosis and inflammation are important factors in the development and progression of recurrent POP. These identified genes and pathways may provide a more detailed molecular mechanism underlying recurrent POP development and progression, and hold promise as potential biomarkers and therapeutic targets. However, further studies are required to confirm the present results.

Methods

Microarray Data

The microarray expression profiling dataset GSE28660, deposited by Eyster et al., was downloaded from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). The dataset was based on the GPL2895 GE Healthcare/Amersham Biosciences CodeLink Human Whole Genome Bioarray platform. The GSE28660 dataset, containing 4 primary POP and 4 recurrent POP patients, was the only dataset to meet this condition.

Differential Expression Analysis

Differential expression analysis was performed using the online analysis tool GEO2R; the expression profiles of primary POP and recurrent POP patients were compared to identify the DEGs. P-values and adjusted P-values were calculated using t-tests. Significant DEGs were selected using an adjusted P value of <0.05 and |log2 fold change|>1. We selected the most significant genes when the DEGs were duplicated. The heatmap and Volcano map for the DEGs was created using were drawn using the hiplot tool (https://hiplot.com.cn).

Functional Enrichment Analysis of DEGs

The functional enrichment analysis of DEGs was performed with the DAVID Bioinformatics Tool (version 6.8, https://david-d.ncifcrf.gov/). This analysis included the functional categories, Gene Ontology (GO) terms, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The GO analysis included 3 categories, namely, biological process (BP), cellular component (CC), and molecular function (MF), which were used to predict protein functions. KEGG pathway analysis was used to assign sets of DEGs to specific pathways to enable the construction of the molecular interaction, reaction, and relationship networks. Benjamini-adjusted P < 0.05 and an enriched gene count >5 were chosen as the criteria for significance.

Protein-Protein Interaction (PPI) Network Analysis

The PPI network analysis was conducted using STRING (https://string-db.org/), which is an online database of known and predicted protein-protein interactions. These interactions include physical and functional associations, and the data are mainly derived from computational predictions, high-throughput experiments, automated text mining, and co-expression networks. We mapped the DEGs onto the PPI network and set an interaction score of >0.4 as the threshold value. In addition, Cytoscape v3.6.0 software was used to visualize and construct the PPI network. Nodes with the greatest numbers of interactions with neighboring nodes were considered hub nodes. To identify the key PPI network modules, the app Mcode from the Cytoscape software suite was used to perform the gene network clustering analysis. P <0.05 was set as the significance threshold for identifying key modules.

Immune cell infiltration in POP tissues

The CIBERSORT algorithm was employed to elucidate the proportion of 22 immune cells in POP tissues. The samples with p-value < 0.05 were significant[8]. Pearson correlation analysis was implemented to

obtain the related coefficient between the 22 immune cells. Then, we investigated the differential immune cell infiltration between primary POP and recurrent POP tissues.

Abbreviations

POP:pelvic organ prolapse; CIDEA: cell death-inducing DFFA-like effector); HBD: hemoglobin subunit delta; BMI:Body Mass Index; USLs:uterosacral ligaments; ECM: extracellular matrix; PPI:Protein-Protein Interaction.

Declarations

Author Contributions

ZHU.L,NI.X ,TANG.S contributed to the conception, design, data collection, statistical analysis, and drafting of the manuscript. LIU.W contributed to the preliminary review. All authors have seen and approved the final manuscript.

Acknowledgments

We thank the author who provided the GEO public datasets.

Consent for publication

All authors provided their consent to publish.

Competing interests

The authors declare that they have no competing interests associated with this publication.

Availability of data and materials

All data generated or analyzed during this study are included in this article.

Ethics approval and consent to participate

The study protocol has been approved by the research institute's committee on ethics (approval number 2021-A-03-08), following the guidelines of the Helsinki Declaration.

Funding Sources

Not applicable.

References

[1] Pelvic Organ Prolapse: ACOG Practice Bulletin, Number 214. Obstet Gynecol. 2019 Nov;134(5):e126-e142. doi: 10.1097/AOG.0000000000003519.

[2] Dällenbach P, Jungo Nancoz C, Eperon I, et al. Incidence and risk factors for reoperation of surgically treated pelvic organ prolapse. Int Urogynecol J. 2012;23(1):35-41. doi:10.1007/s00192-011-1483-3.

[3] Friedman T, Eslick GD, Dietz HP. Risk factors for prolapse recurrence: systematic review and meta-analysis. Int Urogynecol J. 2018;29(1):13-21. doi:10.1007/s00192-017-3475-4.

[4] Ashton-Miller JA, DeLancey JO. Functional anatomy of the female pelvic floor. Ann N Y Acad Sci. 2007;1101:266-296. doi:10.1196/annals.1389.034

[5] Liang SQ, Chen CL, Liu P, et al. Three-dimensional MRI reconstruction research on the anatomical relationship among uterosacral ligament and ureter or rectum in pelvic organ prolapse patients. Zhonghua Fu Chan Ke Za Zhi. 2021;56(1):27-33. doi:10.3760/cma.j.cn112141-20200612-00500.

[6] Khadzhieva MB, Kolobkov DS, Kamoeva SV, Salnikova LE. Expression changes in pelvic organ prolapse: a systematic review and in silico study. Sci Rep. 2017;7(1):7668. doi:10.1038/s41598-017-08185-6.

[7] Dai W, Chang Q, Peng W, Zhong J, Li Y. Network Embedding the Protein-Protein Interaction Network for Human Essential Genes Identification. Genes (Basel). 2020;11(2):153. doi:10.3390/genes11020153.

[8] Yu H, Gu D, Yue C, et al. An Immune Cell-Based Signature Associating With EMT Phenotype Predicts Postoperative Overall Survival of ESCC. Front Oncol. 2021;11:636479. doi:10.3389/fonc.2021.636479.

[9] Chen B, Khodadoust MS, Liu CL, et al. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol. 2018;1711:243-259. doi: 10.1007/978-1-4939-7493-1_12.

[10] Gummesson A, Jernås M, Svensson PA, et al. Relations of adipose tissue CIDEA gene expression to basal metabolic rate, energy restriction, and obesity: population-based and dietary intervention studies. J Clin Endocrinol Metab. 2007;92(12):4759-65. doi: 10.1210/jc.2007-1136.

[11] Slayton M, Gupta A, Balakrishnan B, et al. CIDE Proteins in Human Health and Disease. Cells. 2019;8(3):238. doi:10.3390/cells8030238.

[12] Chatterjee A, Mondal P, Ghosh S, et al. PPARγ regulated CIDEA affects pro-apoptotic responses in glioblastoma. Cell Death Discov. 2015;1:15038. doi:10.1038/cddiscovery.2015.38.

[13] Chatterjee A, Mondal P, Ghosh S, Mehta VS, Sen E. PPARγ regulated CIDEA affects pro-apoptotic responses in glioblastoma. Cell Death Discov. 2015;1:15038. doi:10.1038/cddiscovery.2015.38.

[14] Özdemir M, Caglayan F, Bikker FJ, et al. Gingival tissue human beta-defensin levels in relation to infection and inflammation. J Clin Periodontol. 2020;47(3):309-318. doi:10.1111/jcpe.13227.

[15] Li Y, Zhang QY, Sun BF, et al. Single-cell transcriptome profiling of the vaginal wall in women with severe anterior vaginal prolapse. Nat Commun. 2021;12(1):87. doi:10.1038/s41467-020-20358-y.

[16] Tyagi T, Alarab M, Leong Y, et al., Shynlova O. Local oestrogen therapy modulates extracellular matrix and immune response in the vaginal tissue of post-menopausal women with severe pelvic organ prolapse. J Cell Mol Med. 2019 Apr;23(4):2907-2919. doi: 10.1111/jcmm.14199.

[17] Zhao Y, Xia Z, Lin T, et al. Significance of hub genes and immune cell infiltration identified by bioinformatics analysis in pelvic organ prolapse. PeerJ. 2020;8:e9773. doi:10.7717/peerj.9773.

[18] Li Y, Zhang QY, Sun BF Lin T, et al.Single-cell transcriptome profiling of the vaginal wall in women with severe anterior vaginal prolapse. Nat Commun. 2021 Jan 4;12(1):87. doi: 10.1038/s41467-020-20358-y.

Tables

Table 1. Differentially expressed genes.

Gene symbol	Adjusted P-value	logFC	Gene title
Up regulated genes			
LEP	0.0228	7.554	leptin
GPD1	0.0387	6.924	glycerol-3-phosphate dehydrogenase 1
PLIN1	0.0174	6.855	perilipin 1
LPL	0.0174	6.277	lipoprotein lipase
ADIPOQ	0.0299	6.263	adiponectin, C1Q and collagen domain containing
CIDEA	0.0221	6.141	cell death-inducing DFFA-like effector a
RBP4	0.0228	5.527	retinol binding protein 4
PPP1R1A	0.0228	5.312	protein phosphatase 1 regulatory inhibitor subunit 1A
HBD	0.038	4.899	hemoglobin subunit delta
LOC101930114	0.0256	4.815	uncharacterized LOC101930114
TNMD	0.0319	4.718	tenomodulin
FABP4	0.0221	4.316	fatty acid binding protein 4
LINC01485	0.0256	4.311	long intergenic non-protein coding RNA 1485
HBM	0.0354	4.176	hemoglobin subunit mu
AZGP1	0.0354	3.949	alpha-2-glycoprotein 1, zinc-binding
PROK2	0.0221	3.948	prokineticin 2
MESP1	0.0493	3.846	mesoderm posterior bHLH transcription factor 1
LINC01618	0.0346	3.713	long intergenic non-protein coding RNA 1618
SLC19A3	0.0354	3.683	solute carrier family 19 member 3
G0S2	0.0238	3.66	G0/G1 switch 2
LOC338667	0.0279	3.459	V-set and immunoglobulin domain-containing protein 10-like
SOCS3	0.0354	3.404	suppressor of cytokine signaling 3
CLC	0.0342	3.388	Charcot-Leyden crystal galectin
CD36	0.0453	3.285	CD36 molecule

CXCL8	0.0384	3.048	C-X-C motif chemokine ligand 8
S100A12	0.0403	3.023	S100 calcium binding protein A12
BTNL9	0.0238	2.958	butyrophilin like 9
S100P	0.0238	2.931	S100 calcium binding protein P
NEU3	0.0493	2.91	neuraminidase 3
EGR1	0.0354	2.867	early growth response 1
FAM120A	0.04	2.818	family with sequence similarity 120A
KIAA1324	0.0462	2.812	KIAA1324
ZFP36	0.0346	2.805	ZFP36 ring finger protein
CMTM2	0.0162	2.737	CKLF like MARVEL transmembrane domain containing 2
LOC101927531	0.0354	2.662	uncharacterized LOC101927531
HBQ1	0.0465	2.645	hemoglobin subunit theta 1
IFRD2	0.046	2.61	interferon related developmental regulator 2
ALAS2	0.0399	2.588	5'-aminolevulinate synthase 2
TPO	0.0354	2.548	thyroid peroxidase
RBP7	0.0354	2.539	retinol binding protein 7
FPR1	0.0238	2.522	formyl peptide receptor 1
DUSP1	0.044	2.474	dual specificity phosphatase 1
JUNB	0.0238	2.377	JunB proto-oncogene, AP-1 transcription factor subunit
INHBB	0.0346	2.371	inhibin beta B subunit
ASGR2	0.044	2.301	asialoglycoprotein receptor 2
CCR3	0.0476	2.3	C-C motif chemokine receptor 3
HSD11B1	0.0354	2.161	hydroxysteroid 11-beta dehydrogenase 1
NFE2	0.0354	2.156	nuclear factor, erythroid 2
PLA2G16	0.0492	2.129	phospholipase A2 group XVI
PADI4	0.0254	2.019	peptidyl arginine deiminase 4
RGCC	0.044	2.018	regulator of cell cycle

LOC106146153	0.044	2.01	uncharacterized lncRNA LOC106146153
LOC101928284	0.0228	2.009	uncharacterized LOC101928284
MGST1	0.0384	1.922	microsomal glutathione S-transferase 1
AQP9	0.0354	1.809	aquaporin 9
HSD17B8	0.0369	1.799	hydroxysteroid 17-beta dehydrogenase 8
LINC01272	0.0476	1.781	long intergenic non-protein coding RNA 1272
WNT11	0.0384	1.781	Wnt family member 11
MIR193BHG	0.0462	1.753	MIR193B host gene
CSF3R	0.0238	1.72	colony stimulating factor 3 receptor
KLF4	0.0279	1.693	Kruppel like factor 4
AHSP	0.038	1.637	alpha hemoglobin stabilizing protein
MBP	0.0403	1.609	myelin basic protein
OSM	0.0346	1.59	oncostatin M
NR4A1	0.0403	1.585	nuclear receptor subfamily 4 group A member 1
RAC3	0.0378	1.457	ras-related C3 botulinum toxin substrate 3
Table 1. Continued			
Gene symbol	Adjusted P-value	logFC	Gene title
SPTBN1	0.0354	1.431	spectrin beta, non-erythrocytic 1
VNN2	0.0256	1.426	vanin 2
ZZEF1	0.0447	1.395	zinc finger ZZ-type and EF-hand domain containing 1
STEAP1	0.0476	1.372	six transmembrane epithelial antigen of the prostate 1
MRO	0.0262	1.3	maestro
FUT7	0.0377	1.213	fucosyltransferase 7
HADH	0.0384	1.157	hydroxyacyl-CoA dehydrogenase
AMOTL2	0.0493	1.156	angiomotin like 2
TSHZ2	0.0344	1.141	teashirt zinc finger homeobox 2

LOC105376997	0.0407	1.104	uncharacterized LOC105376997
GBA2	0.044	1.101	glucosylceramidase beta 2
DNAJC3	0.0384	1.087	DnaJ heat shock protein family (Hsp40) member C3
MMP25	0.0346	1.078	matrix metallopeptidase 25
MT1L	0.0417	1.077	metallothionein 1L (gene/pseudogene)
TCF15	0.0354	1.058	transcription factor 15 (basic helix-loop-helix)
RASL11A	0.0372	1.002	RAS like family 11 member A
RGS16	0.0354	1.001	regulator of G-protein signaling 16
Down regulated genes			
ZNF740	0.044	-1.045	zinc finger protein 740
AK5	0.0493	-1.062	adenylate kinase 5
U2SURP	0.0384	-1.065	U2 snRNP associated SURP domain containing
RPS6KA5	0.0346	-1.073	ribosomal protein S6 kinase A5
KRT18P59	0.0346	-1.079	keratin 18 pseudogene 59
GUCY1B3	0.0462	-1.088	guanylate cyclase 1 soluble subunit beta
SLF2	0.0332	-1.095	SMC5-SMC6 complex localization factor 2
FUS	0.0492	-1.109	FUS RNA binding protein
SNED1	0.0238	-1.115	sushi, nidogen and EGF like domains 1
C8orf44	0.0399	-1.117	chromosome 8 open reading frame 44
IDI1	0.044	-1.117	isopentenyl-diphosphate delta isomerase 1
ERP29	0.0354	-1.134	endoplasmic reticulum protein 29
NBEA	0.0465	-1.142	neurobeachin
HLF	0.0399	-1.149	HLF, PAR bZIP transcription factor
LOC644656	0.0238	-1.162	uncharacterized LOC644656
NAB2	0.0403	-1.175	NGFI-A binding protein 2
CATSPER2	0.044	-1.177	cation channel sperm associated 2
MAF	0.0434	-1.216	MAF bZIP transcription factor

LOC399900	0.0354	-1.264	uncharacterized LOC399900
C2orf50	0.0354	-1.293	chromosome 2 open reading frame 50
TMEM35A	0.044	-1.331	transmembrane protein 35A
CACNB3	0.0238	-1.373	calcium voltage-gated channel auxiliary subunit beta 3
TBL1X	0.0354	-1.388	transducin (beta)-like 1X-linked
DPY19L2P2	0.0354	-1.449	DPY19L2 pseudogene 2
CPA3	0.0476	-1.451	carboxypeptidase A3
HTR2A	0.0228	-1.561	5-hydroxytryptamine receptor 2A
SGIP1	0.0319	-1.658	SH3 domain GRB2 like endophilin interacting protein 1
TMEM252	0.0354	-1.706	transmembrane protein 252
ZNF680	0.0403	-1.734	zinc finger protein 680
TNKS	0.0493	-1.814	tankyrase
ITGA8	0.0233	-1.888	integrin subunit alpha 8
PRDM8	0.044	-2.144	PR/SET domain 8

Table 2. The enriched terms for DEGs.

Category	Term	Count	Genes	*P*-value
UP_KEYWORDS	Heme	6	GUCY1B3, HBM, TPO, HBD, HBQ1, STEAP1	5.33E-04
UP_SEQ_FEATURE	DNA-binding region:Basic motif	6	MESP1, NFE2, HLF, MAF, TCF15, JUNB	1.41E-03
UP_KEYWORDS	Lipid metabolism	9	HSD11B1, NEU3, PLA2G16, IDI1, LPL, PLIN1, HADH, HSD17B8, GBA2	1.50E-03
UP_KEYWORDS	NAD	5	TNKS, GPD1, STEAP1, HADH, HSD17B8	1.16E-02
UP_KEYWORDS	Disulfide bond	27	SNED1, CSF3R, CXCL8, FPR1, LPL, HTR2A, ASGR2, MMP25, TPO, FUT7, WNT11, KIAA1324, PROK2, CD36, CCR3, CPA3, BTNL9, PDIA2, ADIPOQ, OSM, INHBB, DNAJC3, AZGP1, RBP4, TNMD, LEP, ITGA8	1.73E-02
UP_SEQ_FEATURE	Disulfide bond	24	BTNL9, CPA3, SNED1, PDIA2, CSF3R, CXCL8, ADIPOQ, OSM, FPR1, LPL, INHBB, HTR2A, ASGR2, MMP25, TPO, AZGP1, RBP4, FUT7, LEP, KIAA1324, ITGA8, PROK2, CD36, CCR3	1.74E-02
UP_SEQ_FEATURE	Binding site:Substrate	6	HSD11B1, NEU3, IDI1, GPD1, PADI4, HSD17B8	1.81E-02
UP_KEYWORDS	Iron	6	GUCY1B3, HBM, TPO, HBD, HBQ1, STEAP1	2.35E-02
UP_KEYWORDS	Calcium	10	SNED1, CACNB3, MMP25, TPO, CATSPER2, ITGA8, S100A12, S100P, PADI4, ASGR2	3.17E-02
GOTERM_BP_DIRECT	positive regulation of transcription from RNA polymerase II promoter	12	EGR1, MESP1, NR4A1, RPS6KA5, RGCC, HLF, MAF, TNKS, OSM, TBL1X, KLF4, JUNB	0.01951
GOTERM_BP_DIRECT	negative regulation of transcription from RNA polymerase II promoter	9	EGR1, ZFP36, MAF, LEP, TSHZ2, CD36, TBL1X, KLF4, JUNB	0.046058
GOTERM_BP_DIRECT	intracellular signal transduction	7	GUCY1B3, ZFP36, RPS6KA5, CXCL8, PPP1R1A, DUSP1, RAC3	0.024445
GOTERM_BP_DIRECT	inflammatory	7	MMP25, RPS6KA5, CXCL8, FPR1,	0.018727

	response		PROK2, S100A12, CCR3	
GOTERM_BP_DIRECT	lipid metabolic process	5	LEP, CIDEA, LPL, PLIN1, CD36	0.011286
GOTERM_BP_DIRECT	chemotaxis	5	CXCL8, FPR1, PROK2, CMTM2, CCR3	0.004709

Figures

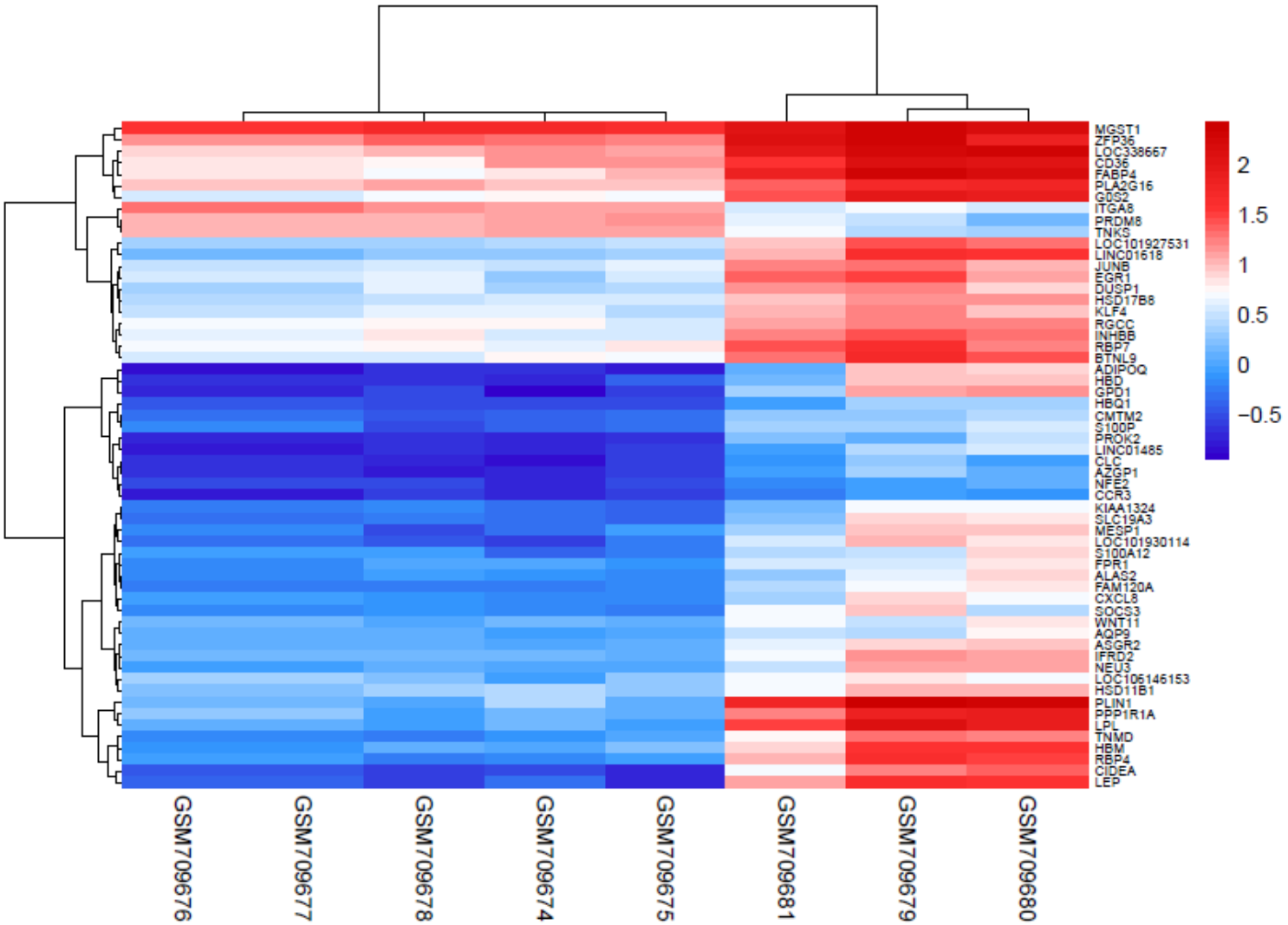

Figure 1

Heatmap of diferentially expressed gene (DEG) clustering. Blue represents downregulation and red represents upregulation.

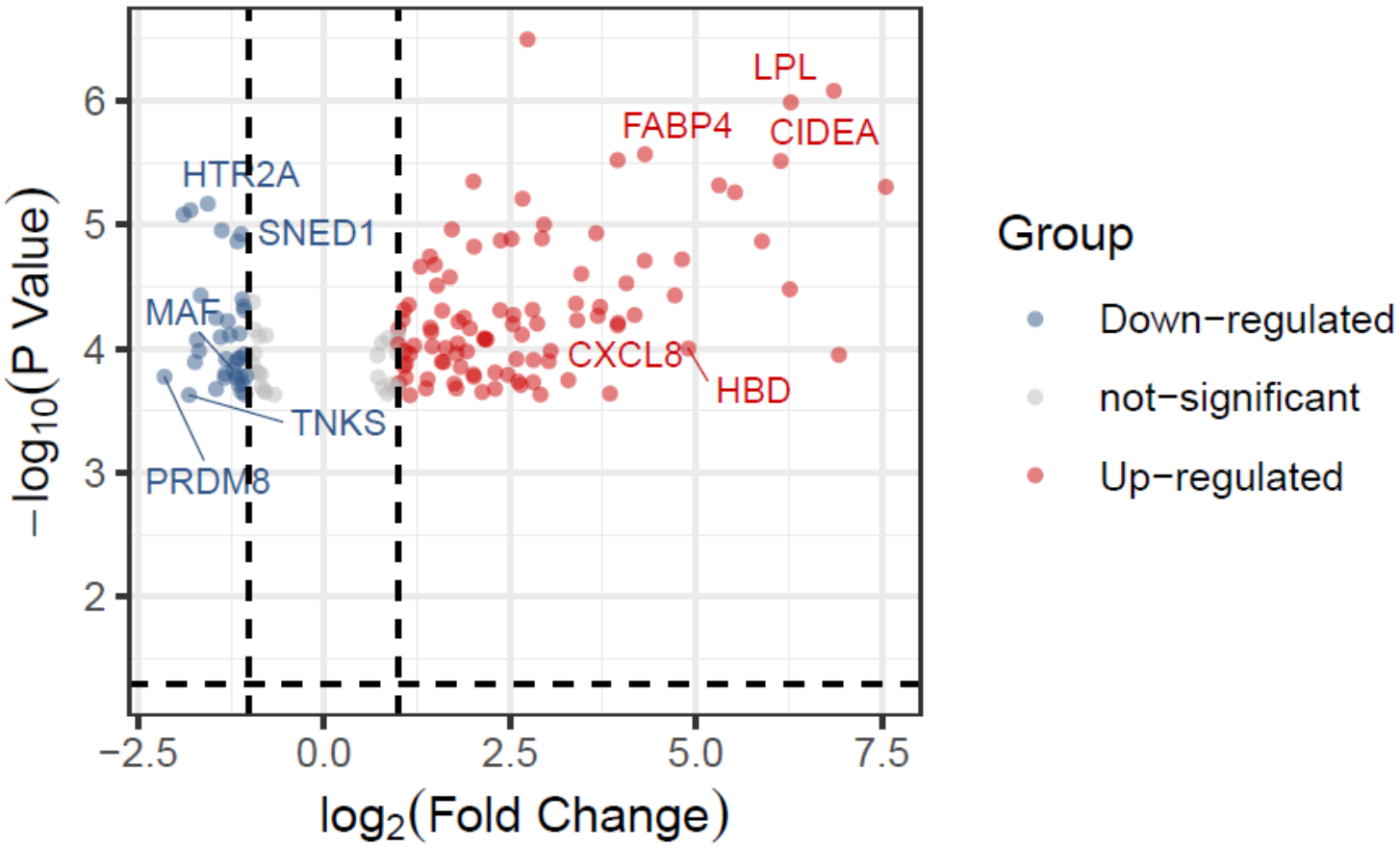

Figure 2

Volcano map of diferentially expressed gene (DEG) clustering. Blue represents downregulation and red represents upregulation.

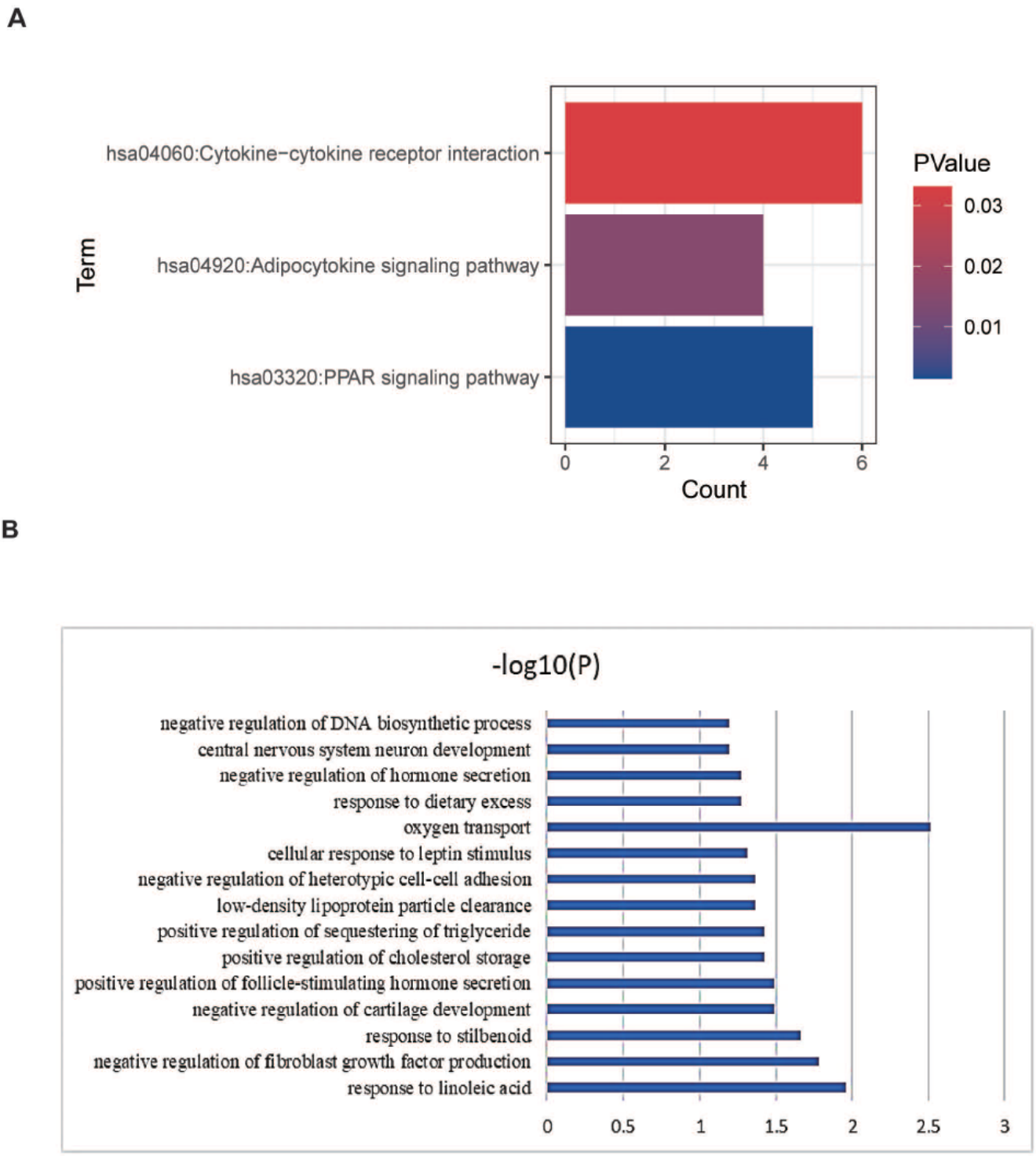

Figure 3

(A) Functional and Pathway Enrichment of DEG. Fig.3.(B) Bar graph of 15 representative enriched functional terms. The x-axis depicts the –log10 (P-value). The y-axis lists the enriched functional terms.

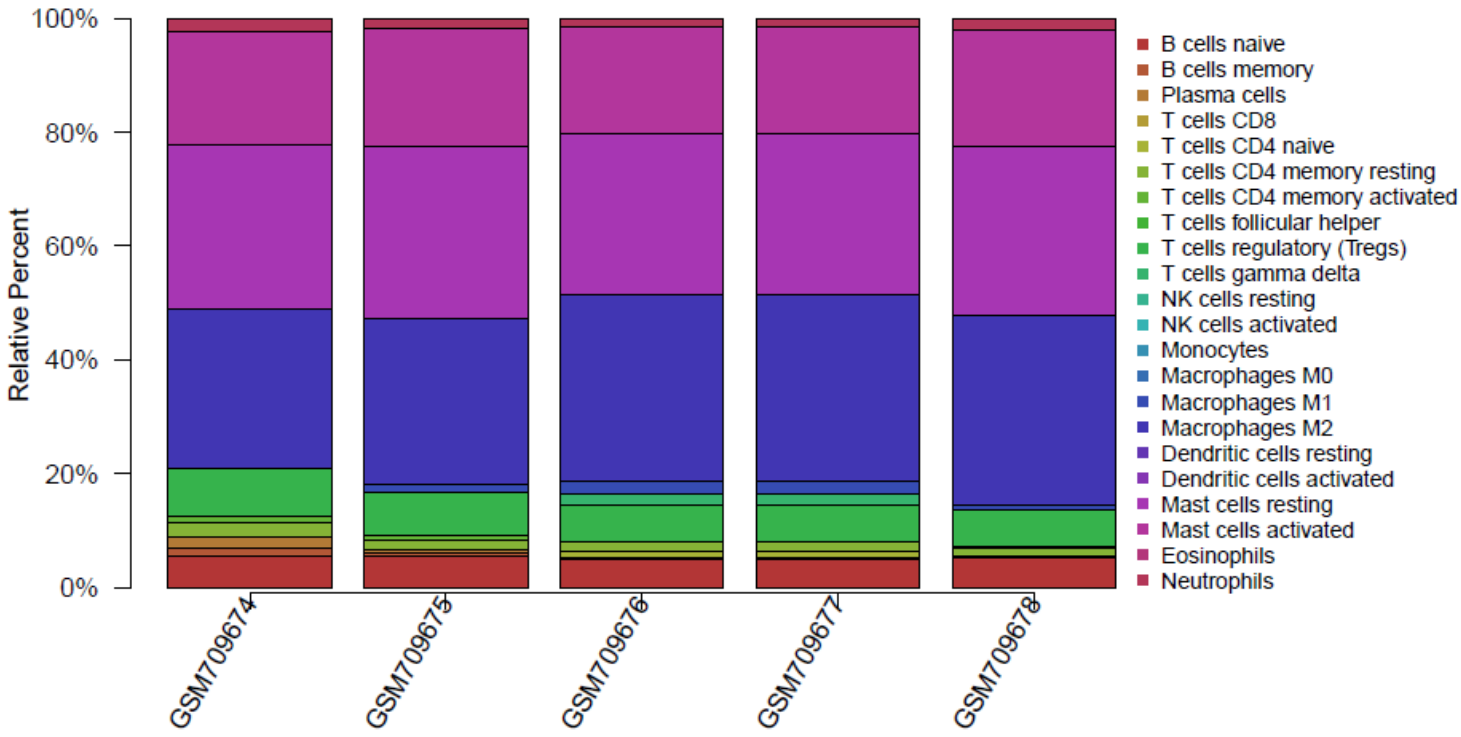

Figure 4

(A) Cytoscape network visualization of the 34 nodes and 53 edges that were obtained with interaction scores >0.4 according to the STRING online database. The nodes represent genes, and the edges represent links between genes. Red represents upregulated genes, and green represents downregulatedgenes. (B,C) Two key modules were identified by Mcode, which was used to identify network gene clustering.

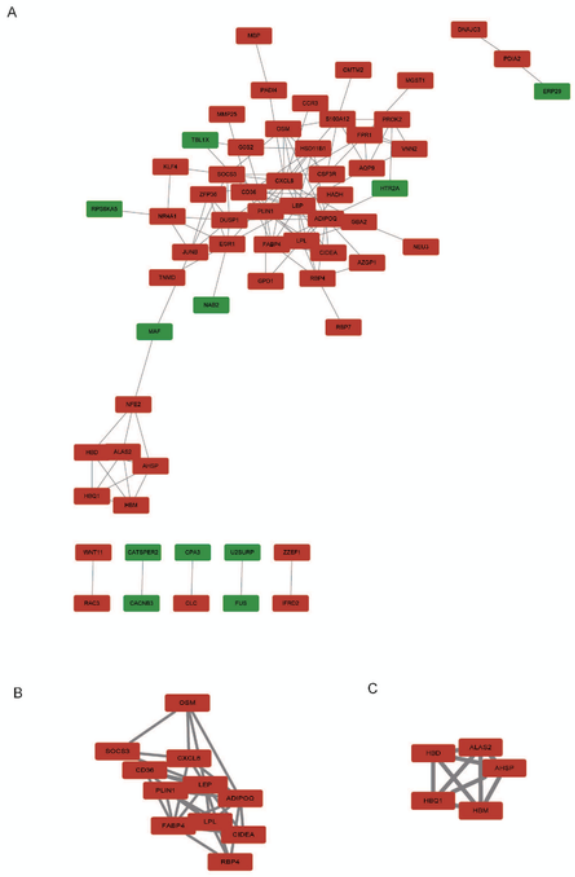

Figure 5

The landscape of immune cell infiltration in GSE28660 (CIBERSORT p value < 0.05). Proportion of the 22 immune cell types in GSE28660.