

# Screening of tumor microenvironment-related prognostic genes in breast cancer by data mining

Yunhui LI (✉ [liyunhui2737@163.com](mailto:liyunhui2737@163.com))

PLA North Military Command Region General Hospital <https://orcid.org/0000-0003-1466-8776>

Na REN

PLA Army General Hospital

---

## Primary research

**Keywords:** Breast cancer, Tumor microenvironment, Bioinformatics, Immune infiltration, Prognosis

**Posted Date:** February 24th, 2020

**DOI:** <https://doi.org/10.21203/rs.2.24359/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background** Accumulating evidence has demonstrated that the components of tumor immune microenvironment (TME) play important roles in breast cancer (BC) initiation, progression and prognosis.

**Materials and methods** We downloaded the TCGA, GSE12276, GSE58812 and GSE42568 datasets. We calculated the immune scores and tumor immune infiltrating cells (TIICs) of TCGA-BRCA and GEO datasets using ESTIMATE and CIBERSORT algorithm, respectively. Then, the overlapping immune-related differentially expressed genes (DEGs) were screened using R 'limma' package between TCGA and GSE12276 datasets. The GO and KEGG enrichment analysis were used to predict the function and signaling pathways of common DEGs. Finally, we extracted a series of tumor immune microenvironment-related genes, and explore the relationship between these genes and clinical outcomes in TCGA, GSE58812 and GSE42568 datasets.

**Results** Based on the ESTIMATE algorithm, the immune scores were significantly associated with cancer types, as well as overall survival in BC patients. The fractions of some TIICs, such as naïve B cells, memory B cells, CD8 + T cells, resting CD4 + memory T cells, activated CD4 + memory T cells, resting NK cells, monocytes, macrophage M0, M1, M2, resting DCs, activated DCs and resting mast cells, were significantly different between low and high immune score groups (all  $P < 0.05$ ). The DEGs between low and high immune score groups were mainly involved in immune-related biological processes, including adaptive immune response, innate response and inflammatory response. Finally, we found that ACSL5, GIMAP2, HLA-DRA and CLEC10A were significantly associated with prognosis among TCGA, GSE58812 and GSE42568 datasets (all  $P < 0.05$ ).

**Conclusion** These findings provide a more comprehensive understanding of immune cells and immune-related genes within TME as well as prognosis-related genes in BC. Future studies need to perform in vivo and in vitro experiments to clarify the mechanisms of these genes in TME and provide a comprehensive idea to immune therapy.

## Backgrounds

Breast cancer (BC) is the most common malignancy in women worldwide, approximately 2.4 million new cases and 533,000 deaths due to BC in 2015 [1, 2]. Estimated by American Cancer Society, the global incidence of BC will be approximately 3.2 million new cases until to 2050 [3]. With the technological advances and health care in medicine, it is possible to detect the BC early and adopt early treatment, which prevent the BC progression into advanced state. But, the 5-year survival rate of BC is still being dismal due to highly heterogeneous in BC. BC is a complex disease, characterized by a wide range of pathological and histological characteristics, different molecular subtypes, and diverse response to therapy [4]. Therefore, a deep and comprehensive exploration of BC pathogenesis is urgently needed, and effective therapeutic targets are required for BC treatment.

BC progression is a multi-factor process influenced by gene alteration and extracellular stimuli that trigger some key signaling pathways abnormal between cancer cells and the main components of immune system. Accumulating evidence has demonstrated that the components of tumor immune environment play important roles in BC initiation, progression and prognosis [5, 6]. BC is characterized by a highly inflammatory microenvironment, which is supported by the infiltrating innate and adaptive immune cell subpopulations and cytokines [7]. Several reports indicated that the presence and/or absence of immune cell subgroups in tumor microenvironment (TME) were associated with clinical stages and clinical outcomes in BC [8, 9]. Previous studies have suggested that high immune cell infiltration were correlated with favorable prognosis in BC patients [10, 11]. It is necessary to investigate the quality and quantity of immune response in BC, as it may help to identify the driver genes involved in immune response, explore the key signal pathways of anti-tumor immune response, pinpoint the BC patients who could benefit from immunotherapy, understand the tumor-immune cell interactions through numerous molecules, thus, help us to tailor the individual treatment and improve our understanding of tumor-host biology.

Recently, more and more platforms, database and datasets provide the researchers comprehensive multi omics data to make in-depth bioinformatics analysis. In the present study, we downloaded the TCGA-BRCA dataset and BC datasets from Gene Expression Omnibus (GEO) database, including RNA sequencing data and clinical information. Then, we calculated the immune scores and the fractions of tumor immune infiltrating cells (TIICs) of TCGA-BRCA and GEO datasets using ESTIMATE algorithm (Estimation of STromal and Immune cells in MAlignant Tumors using Expression data) and CIBERSORT (Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts) algorithm, respectively. Then, we identified the common immune-related differentially expressed genes (DEGs), and performed the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. Finally, we extracted a series of TME-related genes and explore their association with clinical outcomes, and use GEO datasets to verify the prognosis-related genes for result robustness.

## Materials And Methods

### Datasets

The transcriptome data of BRCA were downloaded from the Cancer Genome Atlas database (TCGA), including RNA-sequencing data and clinical information. The clinical data for BC patients in TCGA BRCA dataset includes clinical stage, TNM stage and survival information. The GEO datasets, GSE12276, GSE58812 and GSE42568, which included the microarray-based expression data of BC and related clinical information was downloaded from NCBI GEO website. All datasets met the following criteria: (1) contained >100 samples; (2) completed prognostic information; (3) contained human breast cancer tissue samples. The GSE12276, GSE58812 and GSE42568 datasets were based on GPL570 platform [(HG-U133\_Plus\_2) Affymetrix Human Genome U133 Plus 2.0 Array], and included 204 BC samples, 107 BC samples and 104 BC samples, respectively.

### ESTIMATE Algorithm

All data normalization was performed using package *limma*, and immune score of all datasets were calculated by applying ESTIMATE algorithm (<https://bioinformatics.mdanderson.org/estimate/>). The differences of the immune score among different clinical features were conducted by *t*-test, and the survival analysis was evaluated by Kaplan-Meier plot, examined by Log-rank test. A  $P < 0.05$  was considered to be significant.

### **Estimation of tumor-infiltrating immune cells**

CIBERSORT, a deconvolution algorithm, was applied to estimate the fractions of 22 immune infiltrating cell subgroups based on RNA sequencing data [12]. The 22 tumor infiltrating immune cells (TIICs) were calculated for each sample from TCGA BRCA dataset. In each sample, the proportion of all TIICs equaled to 1. These TIICs includes naïve B cells, memory B cells, plasma cells, 7 T cell types (CD8<sup>+</sup> T cells, naïve CD4<sup>+</sup> T cells, resting CD4<sup>+</sup> memory T cells, activated CD4<sup>+</sup> memory T cells, follicular helper T cells, Tregs,  $\gamma\delta$  T cells), macrophages (M0, M1, and M2), mast cells (resting and activated mast cells), NK cells (resting and activated NK cells), dendritic cells (resting and activated DC cells), eosinophils and neutrophils. The fractions of TIICs were compared between low- and high- immune score groups in TCGA samples.

### **Identification of differentially expressed genes**

Based on median immune score, we divided BC samples into high immune score group and low immune score group for TCGA dataset and GSE12276 dataset. The threshold of DEGs were as follows:  $\log_2|\text{fold change (FC)}| > 0.5$ ,  $P < 0.05$ , and false discovery rate (FDR)  $< 0.05$ . The overlapping DEGs would be used for further analysis.

### **Functional enrichment analysis of common DEGs**

In order to reveal the potential functions of common DEGs, GO annotation and KEGG enrichment analysis were performed using DAVID online database (<https://david.ncifcrf.gov/>). The GO terms consist of biological process, cell component, and molecular function. The adjusted  $P$  value  $< 0.05$  was considered statistically significant.

### **Construction of protein-protein interaction (PPI) network**

The potential interactions of common DEGs were analyzed using Search Tool for the Retrieval of Interacting Genes database (<https://string-db.org/>). The highest confidence in STRING database was set as 0.90. Subsequently, the result of STRING analysis was visualized by Cytoscape software (version 3.7.1, <https://cytoscape.org>). The Molecular Complex Detection plug-in (MCODE) was used to explore the most significant modules in PPI network with the degree cut-off = 2, node score cut-off = 0.2, max depth = 100 and  $k$ -score = 2.

### **Survival analysis of DEGs**

The common DEGs were grouped by dichotomies method according the median value of gene expression in TCGA dataset. The Kaplan-Meier curves were used to perform overall survival analyses between high- and low- expression groups. The criteria were set as: HR with 95% CI and Log-rank  $P$  value  $<0.05$ . The significant prognosis-related genes were validated by the GSE58812 and GSE42568 datasets.

## Statistical analysis

Statistical analyses were conducted using SPSS 20.0 software (SPSS 20.0, Chicago, IL) and GraphPad 8.0 (GraphPad Software, La Jolla, CA, USA). The comparison was performed with student  $t$ -test and ANOVA. The overall analysis used Kaplan-Meier plotter and Log-rank test. In the above analyses,  $P$  values  $< 0.05$  indicated statistical significance.

# Results

## Association of immune scores with different clinical characteristics and prognosis

The RNA-sequencing data and clinical information of BC samples were downloaded from TCGA database. A total of 1,097 samples were included in this study. The clinical data were list in Table S1. Based on the ESTIMATE algorithm, the immune scores range from -1724.48 to 3459.35. There was significant difference among different cancer types for immune scores ( $P<0.001$ , Figure 1A). The average immune scores in metaplastic breast cancer (MBC) were the highest, followed by breast invasive lobular carcinoma (ILC), breast mixed ductal and lobular carcinoma (MDLC), breast invasive ductal carcinoma (IDC), paget disease of the nipple (PD), and breast invasive mixed mucinous carcinoma (IMMC). But, we did not find any differences between immune scores and other clinical features (clinical stage:  $P=0.756$ ; metastasis:  $P=0.496$ ; T stage:  $P=0.259$ ; N stage:  $P=0.379$ , Figure 1B, C, D, E). In survival analysis, the BC patients with high immune scores had a longer overall survival time than low immune scores ( $P=0.034$ , Figure 1F).

## Estimation of the fractions of TIICs

According to the median value of immune score, we divided the TCGA samples into low immune score group and high immune score group. The fractions of naïve B cells ( $P=0.001$ ), memory B cells ( $P=0.024$ ), CD8<sup>+</sup> T cells ( $P<0.001$ ), resting CD4<sup>+</sup> memory T cells ( $P<0.001$ ), activated CD4<sup>+</sup> memory T cells ( $P<0.001$ ), resting NK cells ( $P<0.001$ ), monocytes ( $P<0.001$ ), macrophage M1 ( $P<0.001$ ) and resting DCs ( $P=0.013$ ) in high immune score group were significantly higher than low immune score group; while the fractions of macrophage M0 ( $P<0.001$ ), macrophage M2 ( $P<0.001$ ), activated DCs ( $P=0.008$ ), and resting mast cells ( $P<0.001$ ) in high immune score group were significantly lower than low immune score group (Figure 2).

## Identification of immune related DEGs

According to the median value of immune score, a total of 3,104 DEGs, including 1,782 up-regulated genes and 1,322 down-regulated genes, were identified between high immune score and low immune score groups in TCGA database. Next, a total of 1,294 DEGs were obtained from GSE12276 dataset, of which 1,048 genes were up-regulated, and 246 genes were down-regulated. The heatmaps were constructed for TCGA dataset and GSE 12276 dataset (Figure 3A and 3B). Venn diagram analyses showed 858 common DEGs between TCGA and GSE12276 datasets (Figure 3C).

### **GO function and KEGG pathway enrichment analyses**

To further understand the biological functions of common DEGs, we conducted GO and KEGG functional enrichment analyses. We list the top 30 terms for GO functions and KEGG pathways, respectively. The analyses indicated that common DEGs were mainly in a series of biological processes, including adaptive immune response, innate immune response, inflammatory response, chemokine-mediated signaling pathway, cell adhesion, response to lipopolysaccharide, chemotaxis, *et al* (Figure 4A). These DEGs were mainly distributed in plasma membrane, extracellular exosome, extracellular region, and intracellular (Figure 4B). Moreover, the common DEGs have different molecular functions, including receptor binding, carbohydrate binding, serine-type endopeptidase activity, cytokine activity, and chemokine activity (Figure 4C). Subsequently, the common DEGs were enriched in pathways, such as cytokine-cytokine receptor interaction, chemokine signaling pathway, cell adhesion molecules, hematopoietic cell lineage, PI3K-Akt signaling pathway, NF-kappa B signaling pathway, natural killer cell mediated cytotoxicity, Jak-STAT signaling pathway, T cell receptor signaling pathway, and so on (Figure 4D).

### **Construction of PPI network**

The PPI network of DEGs was constructed using STRING database, and visualized by Cytoscape software. The PPI network of DEGs consisted of 227 nodes and 597 edges according to the highest confidence >0.9. Module analyses were performed by MCODE plug-in, which was used to identify the clusters or highly interconnected regions. We list the most significant three modules with high scores in the PPI network (Figure 5A, B, C), and related GO function and KEGG pathway (Table 1).

### **Prognostic evaluation of DEGs**

The common DEGs were evaluated by Kaplan-Meier curves, examined by Log-rank test. A total of 150 genes were significantly associated with over survival in TCGA dataset (all  $P < 0.05$ ). The prognostic values of above genes were validated using GSE58812 dataset and GSE42568 dataset for result robustness. Finally, we found that acyl-CoA synthetase 5 (ACSL5), GTPases of immunity-associated proteins (GIMAP2), HLA-DRA and C-type lectin receptor 10A (CLEC10A) were significantly associated with overall survival among TCGA, GASE58812 and GSE42568 datasets (all  $P < 0.05$ , Figure 6).

## **Discussion**

BC continues to threaten the physical and psychological health of women, representing an important worldwide health problem. The impact of TME on BC patient outcome has been studied in many decades. Mounting evidence has demonstrated that TME is essential for BC behaviors: initiation, progression, migration, invasion, metastasis, and therapy resistance [13]. Recently, advances in transcriptomics study help us to form a comprehensive insight into the characteristics of TME and identify the genomic alteration in varied cancer types. A main finding in this study is that we applied TCGA and NCBI GEO datasets to analyze the correlation between the immune score and clinical characteristic and clinical outcome, estimate the fractions of TIICs, screen the DEGs between low immune score and high immune score groups for BC patients, perform the GO function and KEGG pathway analysis, and finally identify the prognostic values of ACSL5, GIMAP2, HLA-DRA and CLEC10A in BC patients.

ESTIMATE algorithm was introduced to assess the stromal cell and immune cell scores and infer the tumor purity based on gene expression data [14]. The prediction accuracy of ESTIMATE method has been validated in multiple datasets. Our results showed that immune scores were significantly different in varied cancer types, indicating that the degree and modes of immune status differed. Maybe, the existing differences are due to the heterogeneity and different pathological mechanisms of BC. In survival analyses, the BC patients with high immune scores had a longer overall survival time than low immune scores. The recruitment of immune cell to BC site is determined by the interaction between cancer cells and non-cancer cells, cellular components and non-cellular components in TME. Moreover, diverse immune cells in TME could interact with the cancer cells through chemokine and cytokine signaling pathways, and be involved in the prognosis of breast cancer patients by supporting and obstructing therapeutic efficacy [15, 16].

To further understand the components of TIICs in TME, we analyzed the fractions of TIICs between low immune score and high immune score groups using CIBERSORT algorithm. CIBERSORT could quantify 22 immune cell types by assessing the relative expression changes of a set of barcode 547 genes [12], which is considered to be the most accurate method for estimating the fractions of TIICs. We found there were significant differences in many immune cell subgroups, including naïve B cells, memory B cells, CD8<sup>+</sup> T cells, resting CD4<sup>+</sup> memory T cells, activated CD4<sup>+</sup> memory T cells, resting NK cells, monocytes, macrophage M0, M1, M2, resting DCs, activated DCs, and resting mast cells. T cells play important roles in cell-mediated adaptive immunity. Growing evidence has demonstrated that T cells protect the body from infections through recognition by T cell receptors (TCRs). T cells could eliminate the tumor cells in the tumor-immunity process, and adoptively-transferred T cells were used as anti-tumor therapeutic approach has been explored extensively over the past several years [17]. Emerging clinical data revealed that tumor immunotherapy is likely to become a key strategy for clinical cancer management. While B cells are known for producing protective antibodies, they also perform diverse immune functions by secreting chemokines and cytokines, presenting antigens, and secreting antibody in tumor immunity [18]. B cells infiltration has been detected in many tumors, including lung cancer [19], gastric cancer [20], breast cancer [18], ovarian cancer [21], colorectal cancer [22], hepatocellular carcinoma [23], prostate

cancer [24], and pancreatic adenocarcinoma [25]. Current studies indicated that B cell depletion, selective clearance of B regulatory cells, and regulation of B cell related signaling pathways have the potential to be effective approaches for tumor immunotherapy. The macrophages, referred to as tumor-associated macrophages, are main components in TME, which influence tumor development and progression [26]. Macrophage consists of two different phenotypes under different physiological or pathological conditions: tumor-inhibiting M1 and tumor-promoting M2 [27]. To date, in all cancer types, high infiltration of macrophage M1 predicted a favorable prognosis, whereas high infiltration of macrophage M2 was associated with poor prognosis [28]. In view of the proportion differences of TIICs in BC patients, we speculated that genomic alteration and immune phenotype maybe interact as both cause and result.

To further understand the underlying mechanism, we screen the differentially expressed genes between low and high immune score groups for BC patients, and analyzed the GO function and KEGG pathways. The result showed that the DEGs mainly involved in immune-related biological processes, including adaptive immune response, innate response and inflammatory response. Moreover, the enriched KEGG pathways also includes common signaling pathway, such as PI3K-Akt signaling pathway, NF-kappa B signaling pathway, and Jak-STAT signaling pathway. PI3K-Akt signaling pathways control multiple T cell functions, including proliferation, migration and metabolism, by increasing cytokine expression [29, 30]. The recent study reported that CC-chemokine ligand 2 (CCL2), secreted by tumor associated macrophages, activates PI3K/Akt pathway and promote endocrine resistance in breast cancer [31]. In addition, Lerebours F, et al. found that NF-kappaB pathway plays a major role in inflammatory breast cancer and contributes to the aggressiveness of its phenotype [32]. Macrophage migration inhibitory factor functions as a critical link between immune response and carcinogenesis via activation of HMGB1/TLR4/NF kappa B axis [33]. Furthermore, JAK/STAT signaling pathway is well-known for its key roles in a wide physiology and pathological environment. Activated JAK/STAT pathway in tumor exhibits oncogenic activity in BC, but JAK/STAT inhibition in TME promote therapeutic resistance by producing more protumorigenic inflammatory factors [34]. Therefore, our study highlights the complex mechanism of BC, and provides insight on potential therapy targets in BC.

Finally, we found ACSL5, GIMAP2, HLA-DRA and CLEC10A were correlated with overall survival in BC patients, validated by GSE58812 and GSE42568 datasets. ACSL5, as a lipid metabolic enzyme, activate the most abundant long-chain fatty acids [35]. Previous studies indicated that ACSL5 have diverse roles in different cancers. Chen WC, et al. found that ACSL5 was significantly up-regulated in bladder, lung, pancreatic, prostate, bladder, and esophageal cancer, but down-regulated in breast, ovarian and liver cancer [36]. Yen MC also reported that high ACSL5 was associated with favorable prognosis in BC patients. Similar with above studies, our analyses suggested that ACSL5 was a potential prognostic marker, and positively correlated with overall survival. GIMAPs family is a conserved clade of guanine nucleotide-binding (G) proteins family, including 7 members, located on chromosome 7 [37]. Several studies have demonstrated that GIMAPs play important roles in lymphocyte development, maintenance, and differentiation [38–41]. Human leukocyte antigen (HLA)-DR-alpha (DRA) (HLA-DRA) belongs to HLA class II gene, which are necessary for the induction and regulation of the adaptive immune response [42]. HLA-DRA appears to be a key gene in the presentation of tumor-associated antigens to T lymphocytes,

generating an effective anti-tumor immune response. Some studies reported that low expression of HLA class II genes was associated with aggressive phenotype in many cancers [43]. CLEC10A belongs to the family of C-type lectin receptors (CLRs). Heger L, et al. found that CLEC10A is mainly expressed on the majority of CD1c<sup>+</sup> dendritic cell [44]. Despite these intriguing results, the function of ACSL5, GIMAP2, HLA-DRA and CLEC10A in BC has remained poorly understood.

## Conclusion

Taken together, our study indicated a close association between immune score in TME and cancer types as well as clinical outcome in BC. We found significant differences of TIICs in TME between low and high immune groups. Moreover, we screened a list of immune-related DEGs and explored that ACSL5, GIMAP2, HLA-DRA, and CLEC10A were correlated with prognosis in BC patients. Future studies need to perform *in vivo* and *in vitro* experiments to clarify the mechanisms of these genes in TME and provide a comprehensive idea to immune therapy.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### Competing interests

All authors declare no conflict of interest.

### Funding

Not applicable

### Authors' contributions

LYH the conceived and designed the study; RN collected the data; LYH and RN analyzed the data; and LYH wrote the paper.

### Acknowledgement

Thanks the researchers for providing TCGA and GEO datasets.

## References

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin.* 2017; 67(1): 7-30.
2. Rodriguez Bautista R, Ortega Gomez A, Hidalgo Miranda A, Zentella Dehesa A, Villarreal-Garza C, Avila-Moreno F, Arrieta O. Long non-coding RNAs: implications in targeted diagnoses, prognosis, and improved therapeutic strategies in human non- and triple-negative breast cancer. *Clin Epigenetics.* 2018; 10: 88.
3. Tao Z, Shi A, Lu C, Song T, Zhang Z, Zhao J. Breast Cancer: Epidemiology and Etiology. *Cell Biochem Biophys.* 2015; 72(2): 333-8.
4. Viale G. The current state of breast cancer classification. *Ann Oncol.* 2012; 23 Suppl 10: x207-10.
5. Tekpli X, Lien T, Rossevold AH, Nebdal D, Borgen E, Ohnstad HO, Kyte JA, Vallon-Christersson J, Fongaard M, Due EU, Svartdal LG, Sveli MAT, Garred O, Osbreac, Frigessi A, Sahlberg KK, Sorlie T, Russnes HG, Naume B, Kristensen VN. An independent poor-prognosis subtype of breast cancer defined by a distinct tumor immune microenvironment. *Nat Commun.* 2019; 10(1): 5499.
6. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med.* 2013; 19(11): 1423-37.
7. Esquivel-Velazquez M, Ostoa-Saloma P, Palacios-Arreola MI, Nava-Castro KE, Castro JI, Morales-Montor J. The role of cytokines in breast cancer development and progression. *J Interferon Cytokine Res.* 2015; 35(1): 1-16.
8. Baharlou R, Atashzar MR, Vasmehjani AA, Rahimi E, Khoshmirsafa M, Seif F, Mahdiyar M. Reduced levels of T-helper 17-associated cytokines in the serum of patients with breast cancer: indicators for following the course of disease. *Cent Eur J Immunol.* 2016; 41(1): 78-85.
9. de Groot AF, Blok EJ, Charehbili A, Engels CC, Smit V, Dekker-Ensink NG, Putter H, Meershoek-Klein Kranenbarg E, van de Velde CJH, Liefers GJ, Nortier JWR, Kuppen PJK, van der Burg SH, Kroep JR. Strong CD8+ lymphocyte infiltration in combination with expression of HLA class I is associated with better tumor control in breast cancer patients treated with neoadjuvant chemotherapy. *Breast Cancer Res Treat.* 2019; 175(3): 605-15.
10. Manuel M, Tredan O, Bachelot T, Clapisson G, Courtier A, Parmentier G, Rabeony T, Grives A, Perez S, Mouret JF, Perol D, Chabaud S, Ray-Coquard I, Labidi-Galy I, Heudel P, Pierga JY, Caux C, Blay JY, Pasqual N, Menetrier-Caux C. Lymphopenia combined with low TCR diversity (divpenia) predicts poor overall survival in metastatic breast cancer patients. *Oncoimmunology.* 2012; 1(4): 432-40.
11. Papatestas AE, Lesnick GJ, Jenkins G, Aufses AH, Jr. The prognostic significance of peripheral lymphocyte counts in patients with breast carcinoma. *Cancer.* 1976; 37(1): 164-8.
12. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015; 12(5): 453-7.

13. Lim B, Woodward WA, Wang X, Reuben JM, Ueno NT. Inflammatory breast cancer biology: the tumour microenvironment is key. *Nat Rev Cancer*. 2018; 18(8): 485-99.
14. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB, Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013; 4: 2612.
15. Dougan M, Dougan SK. Targeting Immunotherapy to the Tumor Microenvironment. *J Cell Biochem*. 2017; 118(10): 3049-54.
16. Wu T, Dai Y. Tumor microenvironment and therapeutic response. *Cancer Lett*. 2017; 387: 61-8.
17. Petersen CT, Krenciute G. Next Generation CAR T Cells for the Immunotherapy of High-Grade Glioma. *Front Oncol*. 2019; 9: 69.
18. Garaud S, Buisseret L, Solinas C, Gu-Trantien C, de Wind A, Van den Eynden G, Naveaux C, Lodewyckx JN, Boisson A, Duvillier H, Craciun L, Ameye L, Veys I, Paesmans M, Larsimont D, Piccart-Gebhart M, Willard-Gallo K. Tumor infiltrating B-cells signal functional humoral immune responses in breast cancer. *JCI Insight*. 2019; 5.
19. Ho KH, Chang CJ, Huang TW, Shih CM, Liu AJ, Chen PH, Cheng KT, Chen KC. Gene landscape and correlation between B-cell infiltration and programmed death ligand 1 expression in lung adenocarcinoma patients from The Cancer Genome Atlas data set. *PLoS One*. 2018; 13(12): e0208459.
20. Hennequin A, Derangere V, Boidot R, Apetoh L, Vincent J, Orry D, Fraisse J, Causeret S, Martin F, Arnould L, Beltjens F, Ghiringhelli F, Ladoire S. Tumor infiltration by Tbet+ effector T cells and CD20+ B cells is associated with survival in gastric cancer patients. *Oncoimmunology*. 2016; 5(2): e1054598.
21. Milne K, Kobel M, Kalloger SE, Barnes RO, Gao D, Gilks CB, Watson PH, Nelson BH. Systematic analysis of immune infiltrates in high-grade serous ovarian cancer reveals CD20, FoxP3 and TIA-1 as positive prognostic factors. *PLoS One*. 2009; 4(7): e6412.
22. Berntsson J, Nodin B, Eberhard J, Micke P, Jirstrom K. Prognostic impact of tumour-infiltrating B cells and plasma cells in colorectal cancer. *Int J Cancer*. 2016; 139(5): 1129-39.
23. Gamelo M, Tan A, Her Z, Yeong J, Lim CJ, Chen J, Lim KH, Weber A, Chow P, Chung A, Ooi LL, Toh HC, Heikenwalder M, Ng IO, Nardin A, Chen Q, Abastado JP, Chew V. Interaction between tumour-infiltrating B cells and T cells controls the progression of hepatocellular carcinoma. *Gut*. 2017; 66(2): 342-51.
24. Woo JR, Liss MA, Muldong MT, Palazzi K, Strasner A, Ammirante M, Varki N, Shabaik A, Howell S, Kane CJ, Karin M, Jamieson CA. Tumor infiltrating B-cells are increased in prostate cancer tissue. *J Transl Med*. 2014; 12: 30.
25. Castino GF, Cortese N, Capretti G, Serio S, Di Caro G, Minerì R, Magrini E, Grizzi F, Cappello P, Novelli F, Spaggiari P, Roncalli M, Ridolfi C, Gavazzi F, Zerbi A, Allavena P, Marchesi F. Spatial distribution of B cells predicts prognosis in human pancreatic adenocarcinoma. *Oncoimmunology*. 2016; 5(4): e1085147.

26. Sumitomo R, Hirai T, Fujita M, Murakami H, Otake Y, Huang CL. M2 tumor-associated macrophages promote tumor progression in non-small-cell lung cancer. *Exp Ther Med*. 2019; 18(6): 4490-8.
27. Li X, Liu R, Su X, Pan Y, Han X, Shao C, Shi Y. Harnessing tumor-associated macrophages as aids for cancer immunotherapy. *Mol Cancer*. 2019; 18(1): 177.
28. de Groot AE, Pienta KJ. Epigenetic control of macrophage polarization: implications for targeting tumor-associated macrophages. *Oncotarget*. 2018; 9(29): 20908-27.
29. Abu-Eid R, Samara RN, Ozbun L, Abdalla MY, Berzofsky JA, Friedman KM, Mkrtichyan M, Khleif SN. Selective inhibition of regulatory T cells by targeting the PI3K-Akt pathway. *Cancer Immunol Res*. 2014; 2(11): 1080-9.
30. Finlay D, Cantrell D. Phosphoinositide 3-kinase and the mammalian target of rapamycin pathways control T cell migration. *Ann N Y Acad Sci*. 2010; 1183: 149-57.
31. Li D, Ji H, Niu X, Yin L, Wang Y, Gu Y, Wang J, Zhou X, Zhang H, Zhang Q. Tumor-associated macrophages secrete CC-chemokine ligand 2 and induce tamoxifen resistance by activating PI3K/Akt/mTOR in breast cancer. *Cancer Sci*. 2020; 111(1): 47-58.
32. Lerebours F, Vacher S, Andrieu C, Espie M, Marty M, Lidereau R, Bieche I. NF-kappa B genes have a major role in inflammatory breast cancer. *BMC Cancer*. 2008; 8: 41.
33. Lv W, Chen N, Lin Y, Ma H, Ruan Y, Li Z, Li X, Pan X, Tian X. Macrophage migration inhibitory factor promotes breast cancer metastasis via activation of HMGB1/TLR4/NF kappa B axis. *Cancer Lett*. 2016; 375(2): 245-55.
34. Irey EA, Lassiter CM, Brady NJ, Chuntova P, Wang Y, Knutson TP, Henzler C, Chaffee TS, Vogel RI, Nelson AC, Farrar MA, Schwertfeger KL. JAK/STAT inhibition in macrophages promotes therapeutic resistance by inducing expression of protumorigenic factors. *Proc Natl Acad Sci U S A*. 2019; 116(25): 12442-51.
35. Neess D, Bek S, Engelsby H, Gallego SF, Faergeman NJ. Long-chain acyl-CoA esters in metabolism and signaling: Role of acyl-CoA binding proteins. *Prog Lipid Res*. 2015; 59: 1-25.
36. Chen WC, Wang CY, Hung YH, Weng TY, Yen MC, Lai MD. Systematic Analysis of Gene Expression Alterations and Clinical Outcomes for Long-Chain Acyl-Coenzyme A Synthetase Family in Cancer. *PLoS One*. 2016; 11(5): e0155660.
37. Schwefel D, Frohlich C, Eichhorst J, Wiesner B, Behlke J, Aravind L, Daumke O. Structural basis of oligomerization in septin-like GTPase of immunity-associated protein 2 (GIMAP2). *Proc Natl Acad Sci U S A*. 2010; 107(47): 20299-304.
38. Schwefel D, Frohlich C, Daumke O. Purification, crystallization and preliminary X-ray analysis of human GIMAP2. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2010; 66(Pt 6): 725-9.
39. Pascall JC, Webb LMC, Eskelinen EL, Innocentin S, Attaf-Bouabdallah N, Butcher GW. GIMAP6 is required for T cell maintenance and efficient autophagy in mice. *PLoS One*. 2018; 13(5): e0196504.
40. Endale M, Aksoylar HI, Hoebe K. Central role of gimap5 in maintaining peripheral tolerance and T cell homeostasis in the gut. *Mediators Inflamm*. 2015; 2015: 436017.

41. Heinonen MT, Kanduri K, Lahdesmaki HJ, Lahesmaa R, Henttinen TA. Tubulin- and actin-associating GIMAP4 is required for IFN-gamma secretion during Th cell differentiation. *Immunol Cell Biol.* 2015; 93(2): 158-66.
42. Waldburger JM, Suter T, Fontana A, Acha-Orbea H, Reith W. Selective abrogation of major histocompatibility complex class II expression on extrahematopoietic cells in mice lacking promoter IV of the class II transactivator gene. *J Exp Med.* 2001; 194(4): 393-406.
43. Leite FA, Lira RC, Fedatto PF, Antonini SR, Martinelli CE, Jr., de Castro M, Neder L, Ramalho LN, Tucci S, Jr., Mastelaro MJ, Seidinger AL, Cardinalli IA, Yunes JA, Brandalise SR, Tone LG, Scrideli CA. Low expression of HLA-DRA, HLA-DPA1, and HLA-DPB1 is associated with poor prognosis in pediatric adrenocortical tumors (ACT). *Pediatr Blood Cancer.* 2014; 61(11): 1940-8.
44. Heger L, Balk S, Luhr JJ, Heidkamp GF, Lehmann CHK, Hatscher L, Purbojo A, Hartmann A, Garcia-Martin F, Nishimura SI, Cesnjevar R, Nimmerjahn F, Dudziak D. CLEC10A Is a Specific Marker for Human CD1c(+) Dendritic Cells and Enhances Their Toll-Like Receptor 7/8-Induced Cytokine Secretion. *Front Immunol.* 2018; 9: 744.

## Table

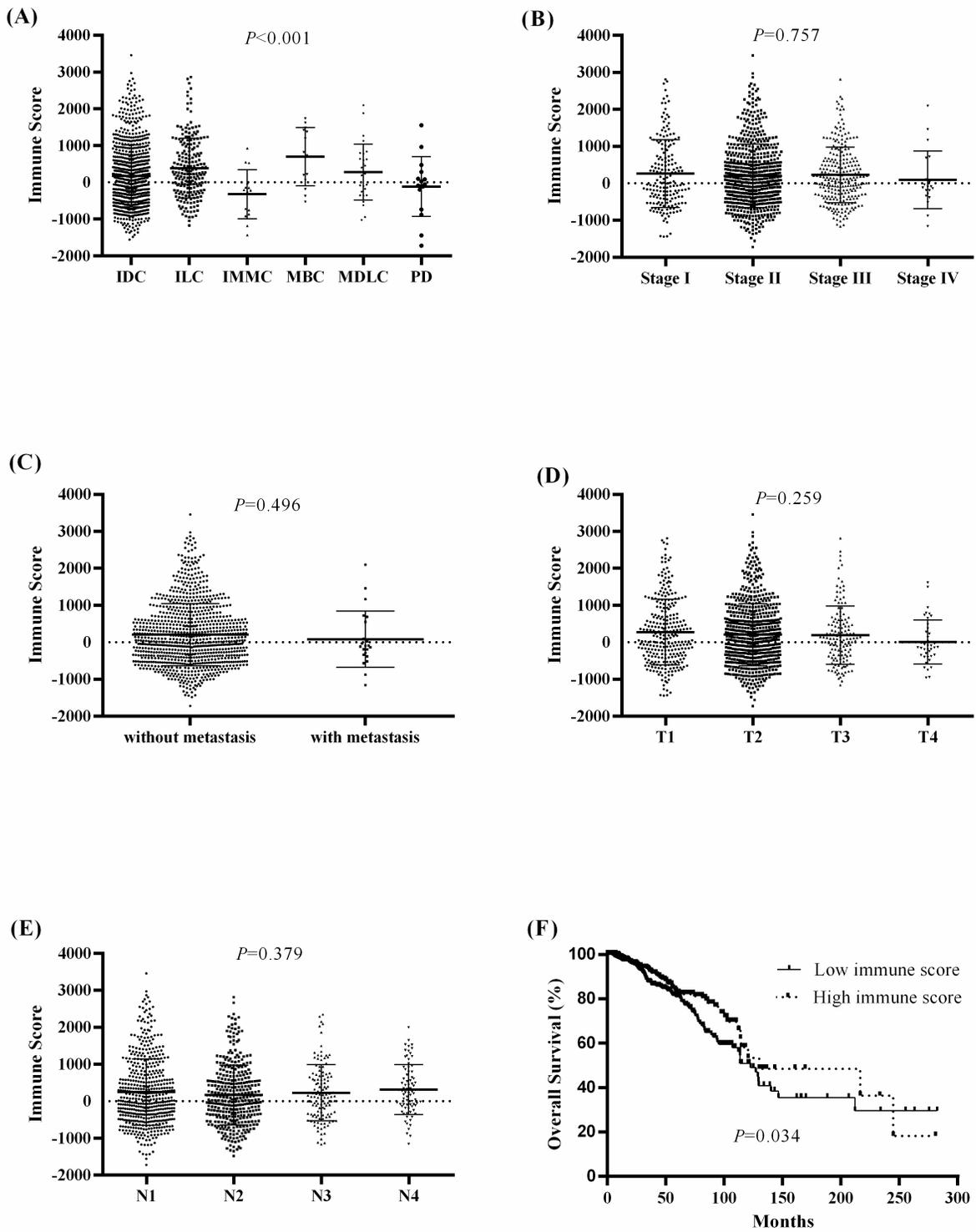
Table 1. GO Function and KEGG pathway enrichment analysis for three modules

Term	Count	<i>P</i> value
<b>Module 1</b>		
GO:0007186 G protein-coupled receptor signaling pathway (BP)	44	<0.001
GO:0042379 chemokine receptor binding	15	<0.001
GO:0044459 plasma membrane part	26	<0.001
hsa04062 Chemokine signaling pathway	27	<0.001
<b>Module 2</b>		
GO:0006955 immune response	31	<0.001
GO:0042605 peptide antigen binding	4	<0.001
GO:0035579 specific granule membrane	16	<0.001
hsa04514 Cell adhesion molecules (CAMs)	8	<0.001
<b>Module 3</b>		
GO:0050870 positive regulation of T cell activation	12	<0.001
GO:0016773 phosphotransferase activity	8	<0.001
GO:0098552 side of membrane	9	<0.001
hsa04660 T cell receptor signaling pathway	12	<0.001

## Supplementary Figure Information

Figure S1. Protein-protein interaction network (PPI) of DEGs from STRING database. The PPI network of DEGs consisted of 227 nodes and 597 edges according to the highest confidence >0.9.

## Figures



**Figure 1**

The correlation between immune scores and clinical characteristics. (A) Cancer types; (B) Clinical stage; (C) Metastasis; (D) T stage; (E) N stage; (F) Overall survival.

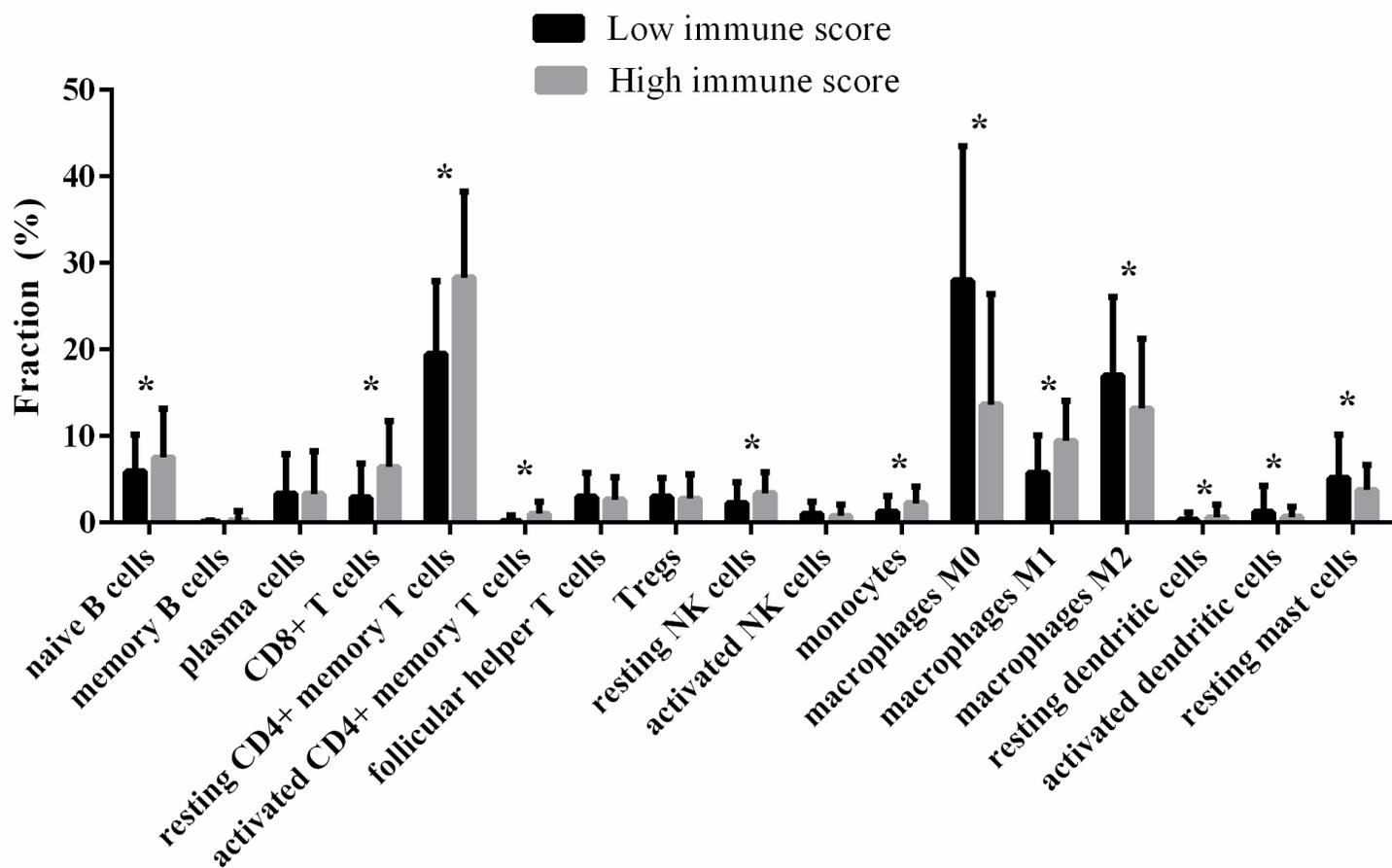
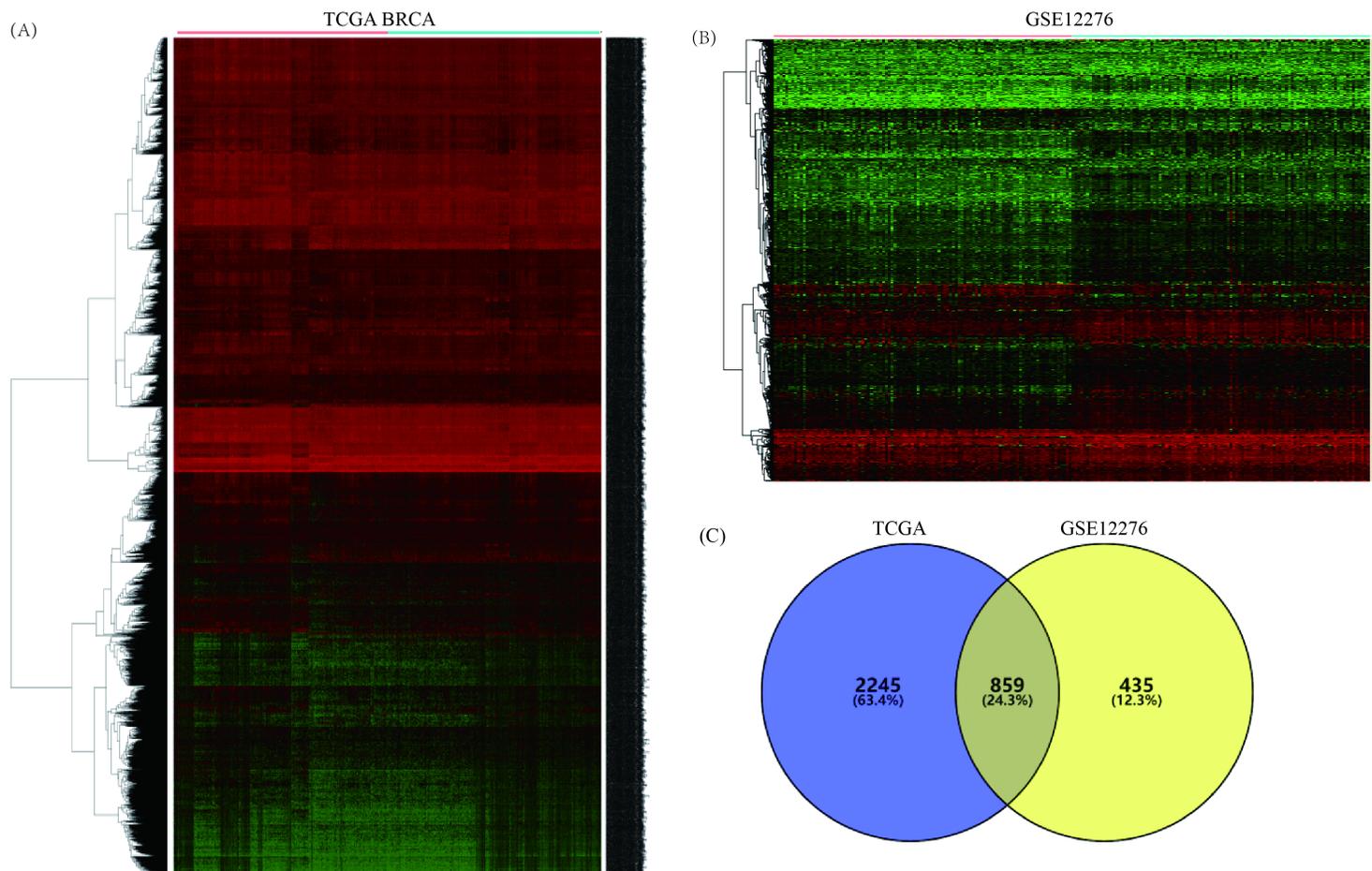


Figure 2

Comparison of tumor immune infiltrating cells between low and high immune score groups in BC patients



**Figure 3**

The differentially expressed genes between low and high immune score groups for TCGA and GSE12276 datasets. (A) Heatmap of DEGs in TCGA dataset; (B) Heatmap of DEGs in GSE12276 dataset; (C) Overlapping DEGs between TCGA and GSE12276 datasets.

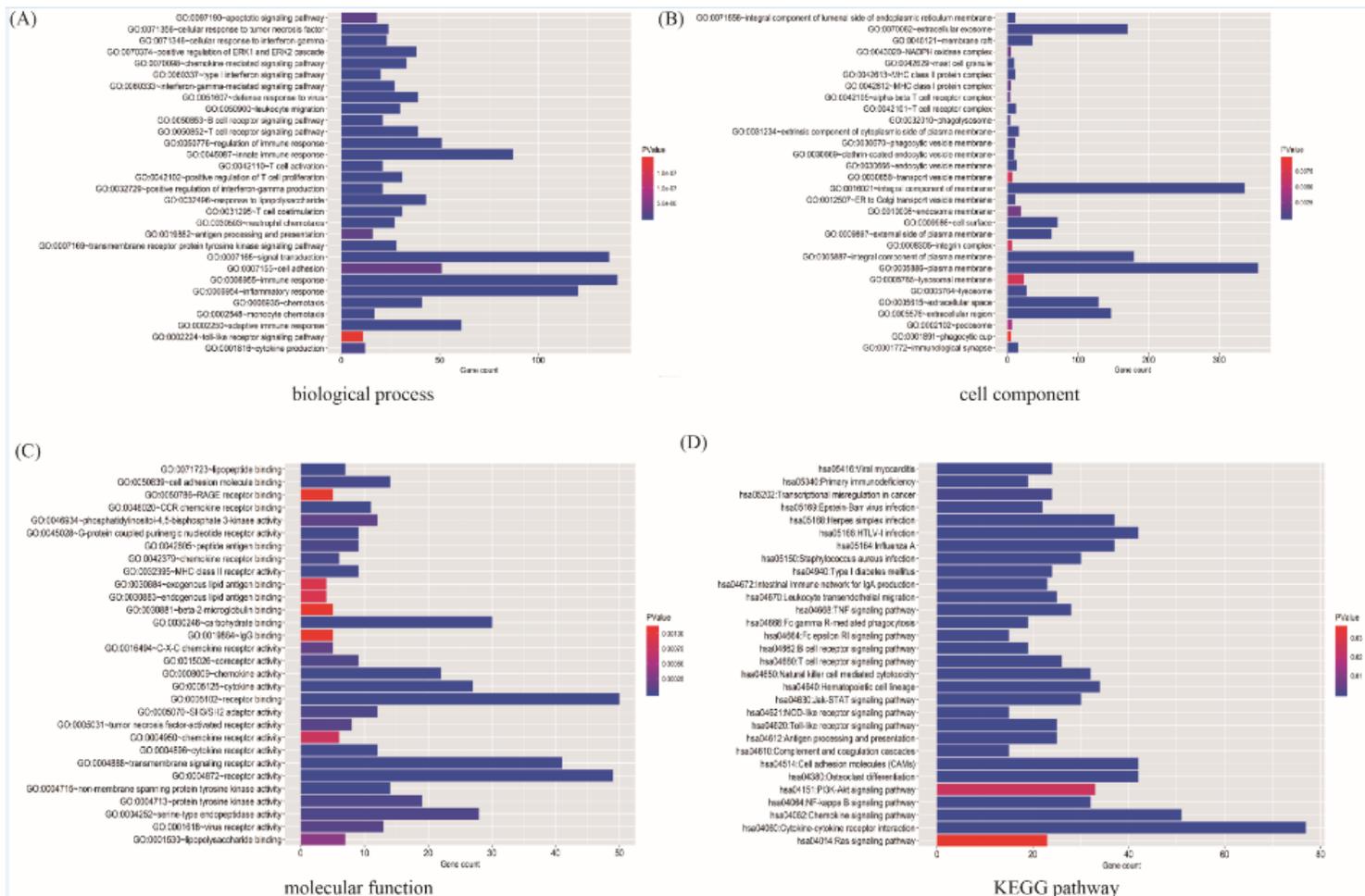


Figure 4

GO function and KEGG pathway enrichment analyses of common DEGs. (A) Biological process; (B) Cell component; (C) Molecular function; (D) KEGG pathways.

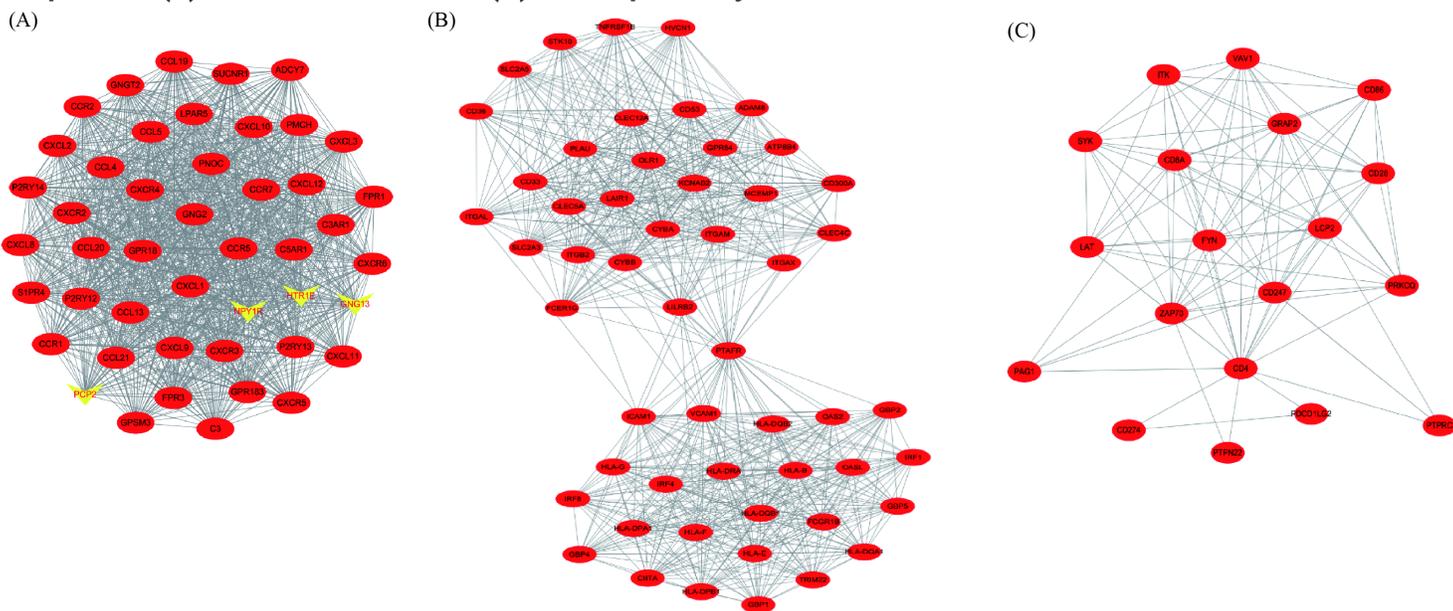
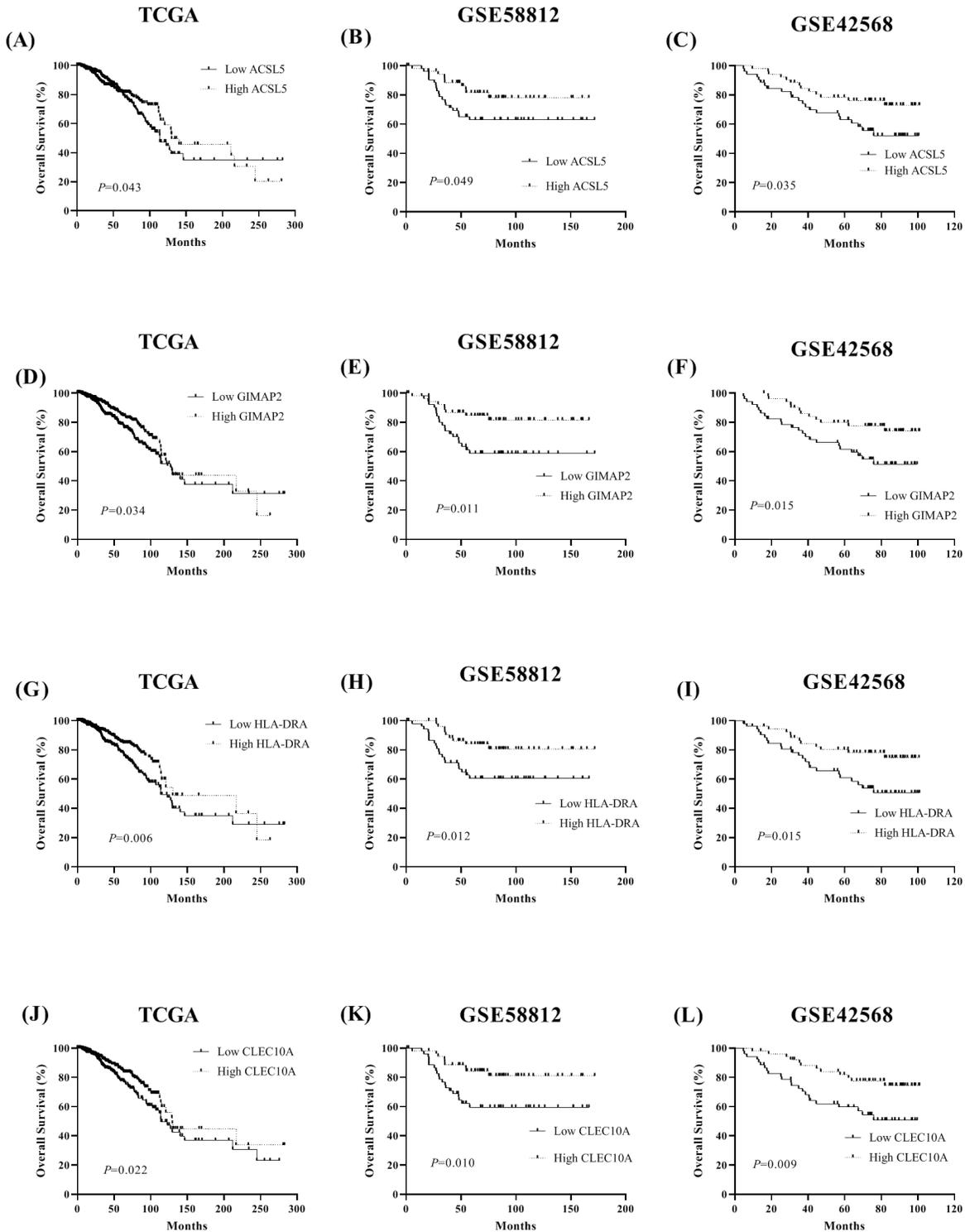


Figure 5

The top three modules of common DEGs using Cytoscape software. Red represents up-regulated genes; yellow presents down-regulated genes. (A) Module 1; (B) Module 2; (C) Module 3



**Figure 6**

Kaplan-Meier survival curves of DEGs. (A-C) ACSL5 in TCGA, GSE58812 and GSE42568 datasets; (D-F) GIMAP2 in TCGA, GSE58812 and GSE42568 datasets; (G-I) HLA-DRA in TCGA, GSE58812 and GSE42568 datasets; (J-L) CLEC10A in TCGA, GSE58812 and GSE42568 datasets.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.tif](#)
- [TableS1.docx](#)